SIAM J. NUMER. ANAL.
Vol. 48, No. 1, pp. 198–223

A POSTERIORI ERROR ESTIMATION BASED ON POTENTIAL AND FLUX RECONSTRUCTION FOR THE HEAT EQUATION*

ALEXANDRE ERN† AND MARTIN VOHRALÍK‡

Abstract. We derive a posteriori error estimates for the discretization of the heat equation in a unified and fully discrete setting comprising the discontinuous Galerkin, various finite volume, and mixed finite element methods in space and the backward Euler scheme in time. Extensions to conforming and nonconforming finite element spatial discretizations are also outlined. Our estimates are based on a H^1-conforming reconstruction of the potential, continuous and piecewise affine in time, and a locally conservative $\mathbf{H}(\mathrm{div})$-conforming reconstruction of the flux, piecewise constant in time. They yield a guaranteed and fully computable upper bound on the error measured in the energy norm augmented by a dual norm of the time derivative. Local-in-time lower bounds are also derived; for nonconforming methods on time-varying meshes, the lower bounds require a mild parabolic-type constraint on the meshsize.

Key words. heat equation, unified framework, a posteriori estimate, discontinuous Galerkin, finite volumes, mixed finite elements, conforming finite elements, nonconforming finite elements

AMS subject classifications. 65N15, 65N30, 76S05

DOI. 10.1137/090759008

1. Introduction.

We consider the heat equation

$$\partial_t u - \Delta u = f \quad \text{a.e. in } Q := \Omega \times (0, t_{\mathrm{F}}), \tag{1.1a}$$

$$u = 0 \quad \text{a.e. on } \partial\Omega \times (0, t_{\mathrm{F}}), \tag{1.1b}$$

$$u(\cdot, 0) = u^0 \quad \text{a.e. in } \Omega, \tag{1.1c}$$

where $\Omega \subset \mathbb{R}^d$, $d \geq 2$, is a polyhedral domain, t_{F} is the finite simulation time, f is the source term, and u^0 is the initial datum. We assume that $f \in L^2(Q)$ and $u^0 \in L^2(\Omega)$. In what follows, u is called the potential and $-\nabla u$ the flux.

The purpose of this work is to derive guaranteed (that is, containing no undetermined constants) and fully computable a posteriori error estimates for the discretization of (1.1a)–(1.1c) by locally conservative methods in space. We consider full discretizations obtained using a backward Euler scheme in time and allow for varying time steps and varying space meshes. The main focus is on nonconforming methods in space, such as discontinuous Galerkin, cell-centered and face-centered finite volumes, and mixed finite elements. Our framework also covers conforming, locally conservative methods such as vertex-centered finite volumes, and under mild modifications, conforming and nonconforming finite elements.

Following the approach proposed by Verfürth for conforming finite elements [43], the error is measured in a (broken) energy norm augmented by a dual norm of the time derivative. This yields error upper bounds that are global-in-space and in time, and

*Received by the editors May 13, 2009; accepted for publication (in revised form) January 8, 2010; published electronically April 2, 2010. This work was partly supported by the Groupement MoMaS (PACEN/CNRS, ANDRA, BRGM, CEA, EdF, IRSN).

http://www.siam.org/journals/sinum/48-1/75900.html

†Université Paris-Est, CERMICS, École des Ponts, 77455 Marne la Vallée cedex 2, France (ern@cermics.enpc.fr).

‡UPMC Université Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, 75005, Paris, France and CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 75005, Paris, France (vohralik@ann.jussieu.fr).

error lower bounds that are local-in-time and global-in-space. The estimators themselves are local-in-time and local-in-space and can be used in a space-time adaptive time-marching algorithm.

The present estimates are based on defining a H^1-conforming reconstruction of the potential, continuous and piecewise affine in time, and a locally conservative $\mathbf{H}(\mathrm{div})$-conforming reconstruction of the flux, piecewise constant in time. A salient feature of this approach is that it allows for a unified setting: the a posteriori error analysis is performed under two simple conditions on the potential and flux reconstructions without any specific reference to the underlying discretization scheme in space. Given a certain scheme, it suffices to verify these two conditions to apply the present analysis. One condition exploits the local conservativity of the scheme through the flux reconstruction, while the other links locally the mean values of the potential reconstruction to those of the discrete solution. The potential reconstruction is not needed for conforming methods (e.g., vertex-centered finite volumes), while the flux reconstruction is not needed for flux-conforming methods (e.g., cell-centered finite volumes and mixed finite elements).

The present parabolic potential and flux reconstructions are inspired from those derived in the context of a posteriori error estimates for elliptic problems in [15, 16, 44, 45, 46, 47, 48]; $\mathbf{H}(\mathrm{div})$-conforming flux reconstruction in elliptic problems can be traced back at least to the Prager–Synge equality [35] and the hypercircle method (cf. [41]), and has been used for a posteriori error estimation in, e.g., [27, 23, 37, 29, 2, 11] and the references therein. The idea to derive parabolic a posteriori error estimates from elliptic estimates on each time level is rather natural. In fact, the residual-based a posteriori error estimates for conforming finite elements derived in [34, 43, 10] take this form. We also mention [30, 28, 20] for the so-called elliptic reconstruction technique allowing for optimal error estimates in higher order norms for conforming finite elements. An important conceptual difference is that we reconstruct the (vector-valued) flux and that this quantity is discrete, is constructed locally by postprocessing, and is directly used to evaluate the estimator. In [8, 7, 40], various estimators for elliptic problems are extended to the heat equation in a conforming setting to bound the error measured in the L^2-norm in the space-time cylinder plus a time-weighted energy-norm; only error upper bounds are considered. Some other results on a posteriori error estimates for parabolic problems in a conforming setting can also be found in [36] using the so-called functional approach where a flux reconstruction is also considered, but without enforcing any local condition; furthermore, only error upper bounds are derived. Finally, we observe that contrary to conforming finite elements, a posteriori energy-norm error estimates for the heat equation discretized by nonconforming methods are less explored; we mention, in particular, [14] for mixed finite elements, [32] for nonconforming finite elements, [21] for discontinuous Galerkin methods, and [4] for finite volume schemes.

This paper is organized as follows. Section 2 presents the continuous and discrete settings. Section 3 collects the main results, namely the error upper and lower bounds. A space-time adaptive time-marching algorithm is also briefly outlined. Section 4 shows how to apply the present framework to various discretization schemes in space, namely discontinuous Galerkin, mixed finite elements, and various finite volume schemes. Sections 5 and 6 are devoted to the proofs of the error upper and lower bounds, respectively. Finally, Appendix A extends the theory to conforming and nonconforming finite elements.

2. The setting. This section briefly describes the continuous and discrete settings.

2.1. The continuous setting. Since $f \in L^2(Q)$ and $u^0 \in L^2(\Omega)$, the exact solution is such that $u \in X := L^2(0,t_{\mathrm{F}};H^1_0(\Omega))$ with $\partial_t u \in X' = L^2(0,t_{\mathrm{F}};H^{-1}(\Omega))$. For a.e. $t \in (0,t_{\mathrm{F}})$ and for all $v \in H^1_0(\Omega)$, there holds

$$\langle \partial_t u, v\rangle(t) + (\nabla u, \nabla v)(t) = (f,v)(t), \tag{2.1}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $H^1_0(\Omega)$ and $H^{-1}(\Omega)$, and $(\cdot,\cdot)$ denotes the $L^2(\Omega)$-inner product with associated norm denoted by $\|\cdot\|$. In what follows, the space $H^1_0(\Omega)$ is equipped with the H^1-seminorm, and for a region $R \subset \Omega$, $\|\cdot\|_R$ denotes the $L^2(R)$-norm with appropriate Lebesgue measure, whereas $|R|$ denotes the measure of R itself.

For $y \in X$, we introduce the space-time energy norm

$$\|y\|_X^2 := \int_0^{t_{\mathrm{F}}} \|\nabla y\|^2(t)\,\mathrm{d}t. \tag{2.2}$$

Furthermore, for $y \in Y := \{y \in X; \partial_t y \in X'\}$, we also introduce the space-time norm

$$\|y\|_Y := \|y\|_X + \|\partial_t y\|_{X'}, \qquad \|\partial_t y\|_{X'} := \left\{\int_0^{t_{\mathrm{F}}} \|\partial_t y\|^2_{H^{-1}}(t)\,\mathrm{d}t\right\}^{1/2}.$$

Observe that $\|\partial_t y\|_{X'} = \sup_{\varphi\in X;\|\varphi\|_X=1} \int_0^{t_{\mathrm{F}}} \langle \partial_t y, \varphi\rangle(t)\,\mathrm{d}t$.

2.2. The discrete setting. This section collects some useful notation concerning the discrete setting.

2.2.1. Time steps and time-varying meshes. We consider a strictly increasing sequence of discrete times $\{t^n\}_{0\le n\le N}$ such that $t^0 = 0$ and $t^N = t_{\mathrm{F}}$, together with a set of meshes $\{\mathcal{T}^n\}_{0\le n\le N}$. The discrete times and the meshes are constructed by a space-time adaptive time-marching algorithm, e.g., that which is outlined in section 3.3.

For all $1 \le n \le N$, we define $I_n := (t^{n-1}, t^n]$ and $\tau^n := t^n - t^{n-1}$. For all $0 \le n \le N$, we assume that $\mathcal{T}^n$ covers exactly the polyhedral domain Ω. For all $T \in \mathcal{T}^n$, h_T denotes the diameter of T, and we let $h^n := \max_{T\in\mathcal{T}^n} h_T$ denote the maximum meshsize of $\mathcal{T}^n$. For simplicity, we also assume that the meshes are simplicial and matching (in the sense that they do not contain so-called hanging nodes). Extensions to general polygonal and nonmatching meshes are possible, but technical; see [17] for an example of flux reconstruction on such meshes. Furthermore, the initial mesh $\mathcal{T}^0$ is used to approximate the initial condition, while for all $1 \le n \le N$, $\mathcal{T}^n$ corresponds to the mesh used to march in time from t^{n-1} to t^n; see Figure 2.1. The meshes can be refined or coarsened as time evolves; precise assumptions on the meshes are stated in section 3. Typically, $\mathcal{T}^n$ is obtained from $\mathcal{T}^{n-1}$ by refining some elements and coarsening some other ones. For all $1 \le n \le N$, we denote by $\mathcal{T}^{n-1,n}$ a common refinement of $\mathcal{T}^{n-1}$ and $\mathcal{T}^n$.

Let W be a vector space of functions defined on Ω. Then, $P^1_\tau(W)$ denotes the vector space of functions defined on Q such that $v(\cdot,t)$ takes values in W and is continuous and piecewise affine in time. Functions in $P^1_\tau(W)$ are uniquely defined by the $(N+1)$ functions $\{v^n := v(\cdot,t^n)\}_{0\le n\le N}$ in W. Similarly, $P^0_\tau(W)$ denotes the vector space of functions defined on Q such that $v(\cdot,t)$ takes values in W and is piecewise constant in time; for $1 \le n \le N$, we then set $v^n := v(\cdot,t)|_{I_n}$. Functions in $P^0_\tau(W)$ are uniquely defined by the N functions $\{v^n\}_{1\le n\le N}$ in W. Furthermore,

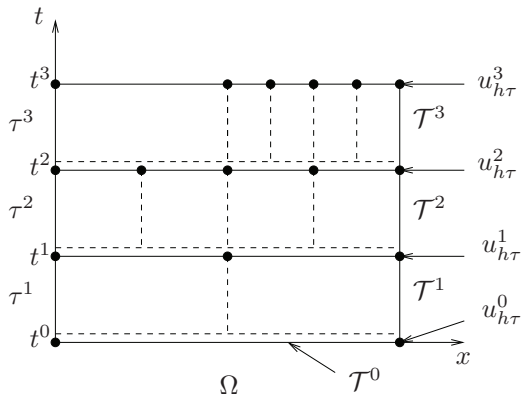

FIG. 2.1. *Time-dependent meshes and discrete solutions.*

we observe that if $v \in P^1_\tau(W)$, then $\partial_t v \in P^0_\tau(W)$ is such that for all $1 \le n \le N$,

$$\partial_t v^n := \partial_t v|_{I_n} = \frac{1}{\tau^n}(v^n - v^{n-1}). \tag{2.3}$$

2.2.2. The discrete solution. For all $0 \le n \le N$, the (yet unspecified) approximate solution at time t^n, say $u^n_{h\tau}$, is such that $u^n_{h\tau} \in V^n_h$, where $V^n_h := V_h(\mathcal{T}^n)$ is a discrete space built on the mesh $\mathcal{T}^n$. The spaces V^n_h consist of piecewise polynomial functions whose degree is fixed. In what follows, Π^n_0 denotes the L^2-orthogonal projection onto piecewise constant functions on $\mathcal{T}^n$, while $\Pi_{V^n_h}$ denotes the L^2-orthogonal projection onto V^n_h.

We introduce the space-time function $u_{h\tau} : Q \to \mathbb{R}$, which is continuous and piecewise affine in time and such that for all $1 \le n \le N$ and $t \in I_n$,

$$u_{h\tau}(\cdot, t) := (1 - \varrho(t))u^{n-1}_{h\tau} + \varrho(t)u^n_{h\tau}, \qquad \varrho(t) = \frac{1}{\tau^n}(t - t^{n-1}).$$

More generally, for any function $v : Q \to \mathbb{R}$ that is continuous in time, we set $v^n := v(\cdot, t^n) : \Omega \to \mathbb{R}$ for all $0 \le n \le N$.

2.2.3. Broken gradients and broken X-norm. Because of possible nonconformities in space approximation, the discrete solution $u_{h\tau}$ may not be in the energy space X. We thus need to extend the definition (2.2) of the X-norm. We introduce for all $0 \le n \le N$ the broken gradient operator ∇^n such that for a function v that is smooth within each mesh element in $\mathcal{T}^n$, $\nabla^n v \in [L^2(\Omega)]^d$ is defined as $(\nabla^n v)|_T := \nabla(v|_T)$ for all $T \in \mathcal{T}^n$. The broken gradient operator $\nabla^{n-1,n}$ on the mesh $\mathcal{T}^{n-1,n}$ is defined similarly. Then, for all y in the affine space $-u_{h\tau} + X$,

$$\|y\|^2_X := \sum_{n=1}^N \int_{I_n} \|\nabla^{n-1,n} y\|^2(t)\,\mathrm{d}t = \sum_{n=1}^N \int_{I_n} \sum_{T \in \mathcal{T}^{n-1,n}} \|\nabla y\|^2_T(t)\,\mathrm{d}t.$$

Since $\|\partial_t y\|_{X'}$ is always well defined, the quantity $\|y\|_Y$ is now well defined for all $y \in -u_{h\tau} + X$. In what follows, y is either the exact solution or the potential reconstruction minus the discrete solution. By abuse of notation, we will speak about the $\|u - u_{h\tau}\|_Y$ norm, even if it can be a seminorm only.

2.2.4. Mesh faces. For all $0 \le n \le N$, the mesh faces in $\mathcal{T}^n$ are collected into the set $\mathcal{F}^n$. More specifically, $F \in \mathcal{F}^n$ if F has positive $(d-1)$-dimensional measure and if there are either distinct mesh elements $T^\pm \in \mathcal{T}^n$ (arbitrary but fixed once and for all) such that $F = \partial T^- \cap \partial T^+$ (F is then called an interior face and we write $F \in \mathcal{F}^{\mathrm{i},n}$) or if there is a mesh element $T \in \mathcal{T}^n$ such that $F = \partial T \cap \partial\Omega$ (F is then called a boundary face and we write $F \in \mathcal{F}^{\mathrm{b},n}$). For $T \in \mathcal{T}^n$, it is convenient to introduce the sets

$$\mathcal{F}_T^n := \{F \in \mathcal{F}^n;\ F \subset \partial T\}, \qquad \mathfrak{F}_T^n := \{F \in \mathcal{F}^n;\ F \cap T \neq \emptyset\},$$
$$\mathfrak{T}_T^n := \{T' \in \mathcal{T}^n;\ T' \cap T \neq \emptyset\};$$

that is, $\mathcal{F}_T^n$ collects the faces of T, $\mathfrak{F}_T^n$ collects the faces having a nonempty intersection with T, and $\mathfrak{T}_T^n$ collects the elements having a nonempty intersection with T. We will also use the set $\mathfrak{F}_T^{\mathrm{i},n} := \mathfrak{F}_T^n \cap \mathcal{F}^{\mathrm{i},n}$.

For $F \in \mathcal{F}^{\mathrm{i},n}$ and a smooth-enough function v that is possibly two-valued on F, we define its average and jump at F as

$$\{\!\{v\}\!\} := \frac{1}{2}(v^- + v^+), \qquad [\![v]\!] := v^- - v^+,$$

where $v^\pm = v|_{T^\pm}$, and we define $\mathbf{n}_F$ as the unit normal to F pointing from T^- towards T^+. For $F \in \mathcal{F}^{\mathrm{b},n}$, the above definitions are extended by setting $\{\!\{v\}\!\} := [\![v]\!] := v|_F$, while $\mathbf{n}_F$ coincides with the unit outward normal to Ω. For a subset $\mathcal{F} \subset \mathcal{F}^n$, we define the jump seminorms

$$|[\![v]\!]|_{\pm\frac{1}{2},\mathcal{F}} := \left\{ \sum_{F\in\mathcal{F}} h_F^{\pm 1} \|[\![v]\!]\|_F^2 \right\}^{1/2},$$

where h_F denotes the diameter of F. Finally, for $F \in \mathcal{F}^{\mathrm{i},n}$, we define the jump in the normal derivative of v as $\mathbf{n}\cdot[\![\nabla^n v]\!] := \mathbf{n}_F\cdot(\nabla v|_{T^-} - \nabla v|_{T^+})$.

3. Main results. This section collects the main results of this paper concerning the error upper and lower bounds. We also outline a space-time adaptive time-marching algorithm to be used in conjunction with the present estimates.

3.1. Error upper bound. The approximation error $u - u_{h\tau}$ will be measured in the Y-norm, while the error upper bound will be formulated in terms of a potential reconstruction s and a flux reconstruction $\boldsymbol{\theta}$.

3.1.1. Assumptions on the potential and flux reconstructions. We assume that

$$s \in P_\tau^1(H_0^1(\Omega)), \qquad \boldsymbol{\theta} \in P_\tau^0(\mathbf{H}(\mathrm{div},\Omega)). \tag{3.1}$$

The potential reconstruction s is determined by the $(N+1)$ functions $s^n \in H_0^1(\Omega)$ associated with the discrete times $\{t^n\}_{0\le n\le N}$, while the flux reconstruction $\boldsymbol{\theta}$ is determined by the N functions $\boldsymbol{\theta}^n \in \mathbf{H}(\mathrm{div},\Omega)$ associated with the time intervals $\{I_n\}_{1\le n\le N}$.

The potential and flux reconstructions are assumed to satisfy two important additional conditions. First, the mean values on mesh elements of the potential reconstruction s are related to those of the discrete solution $u_{h\tau}$. Specifically, we assume that for all $0 \le n \le N$,

$$(s^n, 1)_{T'} = (u_{h\tau}^n, 1)_{T'} \qquad \forall\, T' \in \mathcal{T}^{n,n+1}, \tag{3.2}$$

with the convention that $\mathcal{T}^{N,N+1} := \mathcal{T}^N$. An important consequence of (3.2) is the following lemma.

LEMMA 3.1 (mean values of time derivatives). *Let* (3.1) *and* (3.2) *hold. Then, for all* $1 \le n \le N$,

$$(\partial_t s^n, 1)_T = (\partial_t u_{h\tau}^n, 1)_T \qquad \forall T \in \mathcal{T}^n. \tag{3.3}$$

Proof. Observe first that by definition, for all $1 \le n \le N$,

$$\partial_t (s - u_{h\tau})^n = \frac{1}{\tau^n}[(s^n - u_{h\tau}^n) - (s^{n-1} - u_{h\tau}^{n-1})].$$

Furthermore, by (3.2), s^n and $u_{h\tau}^n$ have the same mean values on all the elements of $\mathcal{T}^n$ and of $\mathcal{T}^{n+1}$. Similarly, s^{n-1} and $u_{h\tau}^{n-1}$ have the same mean values on all the elements of $\mathcal{T}^{n-1}$ and of $\mathcal{T}^n$. Hence, $(s^n - s^{n-1})$ and $(u_{h\tau}^n - u_{h\tau}^{n-1})$ have the same mean values on all the elements of $\mathcal{T}^n$, yielding (3.3). ∎

Second, we assume that the flux reconstruction $\boldsymbol{\theta}$ satisfies the following local conservation property: For all $1 \le n \le N$,

$$(\widetilde{f}^n - \partial_t u_{h\tau}^n - \nabla\cdot\boldsymbol{\theta}^n, 1)_T = 0 \qquad \forall T \in \mathcal{T}^n. \tag{3.4}$$

Here, we have set $\widetilde{f}^n := \frac{1}{\tau^n}\int_{I_n} f(\cdot,t)\,\mathrm{d}t$ (it is also possible to take $\widetilde{f}^n := f^n$ if $f \in C^0(0,t_{\mathrm{F}};L^2(\Omega))$). We define accordingly $\widetilde{f} \in P_\tau^0(L^2(\Omega))$ such that $\widetilde{f}|_{I_n} := \widetilde{f}^n$ for all $1 \le n \le N$.

The actual design of the potential reconstruction s is independent of the given numerical scheme and is described in section 3.2.2. The actual design of the flux reconstruction $\boldsymbol{\theta}$ exploits the local conservation properties of the numerical scheme in space; examples are given in section 4.

3.1.2. The error estimators. For all $1 \le n \le N$ and $T \in \mathcal{T}^n$, we define the *residual estimator* and the *diffusive flux estimator*, respectively, as

$$\eta_{\mathrm{R},T}^n := C_{\mathrm{P}} h_T \|\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n\|_T, \tag{3.5}$$

$$\eta_{\mathrm{DF},T}^n(t) := \|\nabla s(t) + \boldsymbol{\theta}^n\|_T, \quad t \in I_n, \tag{3.6}$$

where $C_{\mathrm{P}} := \frac{1}{\pi}$ is the constant in the Poincaré inequality stating that on each $T \in \mathcal{T}^n$, $1 \le n \le N$ (recall that T is convex as simplex), there holds [33, 9]

$$\|\varphi - \Pi_0^n \varphi\|_T \le C_{\mathrm{P}} h_T \|\nabla\varphi\|_T \qquad \forall \varphi \in H^1(T). \tag{3.7}$$

Observe that the Poincaré inequality is applied locally on each $T \in \mathcal{T}^n$ and not globally on Ω. This is possible owing to the local conservation property (3.4) of the flux reconstruction, as reflected in the proof of Lemma 5.2. Furthermore, still for all $1 \le n \le N$ and $T \in \mathcal{T}^n$, we define the following *nonconformity estimators*:

$$\eta_{\mathrm{NC},1,T}^n(t) := \|\nabla^{n-1,n}(s - u_{h\tau})(t)\|_T, \quad t \in I_n, \tag{3.8}$$

$$\eta_{\mathrm{NC},2,T}^n := C_{\mathrm{P}} h_T \|\partial_t (s - u_{h\tau})^n\|_T. \tag{3.9}$$

We observe that the four above estimators are local-in-space and in time. Finally, we define the *initial condition estimator* as

$$\eta_{\mathrm{IC}} := 2^{1/2}\|s^0 - u^0\|. \tag{3.10}$$

3.1.3. Guaranteed and fully computable upper bound. We are now in a position to state our main result concerning the error upper bound.

THEOREM 3.2 (error upper bound). *Assume* (3.1), (3.2), *and* (3.4). *Let* $\eta^n_{\mathrm{R},T}$, $\eta^n_{\mathrm{DF},T}$, $\eta^n_{\mathrm{NC},1,T}$, $\eta^n_{\mathrm{NC},2,T}$, *and* η_{IC} *be defined by* (3.5)–(3.10). *Then,*

$$\|u-u_{h\tau}\|_Y \le 3\left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta^n_{\mathrm{R},T}+\eta^n_{\mathrm{DF},T}(t))^2\,\mathrm{d}t\right\}^{1/2} + \eta_{\mathrm{IC}} + 3\|f-\widetilde{f}\|_{X'} + \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta^n_{\mathrm{NC},1,T})^2(t)\,\mathrm{d}t\right\}^{1/2} + \left\{\sum_{n=1}^N \tau^n \sum_{T\in\mathcal{T}^n} (\eta^n_{\mathrm{NC},2,T})^2\right\}^{1/2}.$$

Remark 3.3 (conforming methods). For conforming methods, $u^n_{h\tau}\in H^1_0(\Omega)$ for $0\le n\le N$, and we set $s^n := u^n_{h\tau}$. Hence, the nonconformity estimators $\eta^n_{\mathrm{NC},1,T}$ and $\eta^n_{\mathrm{NC},2,T}$ vanish.

Remark 3.4 (flux-conforming methods). The cell-centered finite volume method of section 4.2, as well as the lowest order mixed finite element method of section 4.3, are flux-conforming methods that directly produce $u^n_{h\tau}$ such that $-\nabla^n u^n_{h\tau} =: \boldsymbol{\theta}^n \in \mathbf{H}(\mathrm{div},\Omega)$ for all $1\le n\le N$. Consequently, we can drop in such a case the flux reconstruction $\boldsymbol{\theta}^n$ and only work with $u^n_{h\tau}$.

Remark 3.5 (time oscillation of the source term). The quantity $\|f-\widetilde{f}\|_{X'}$ can be viewed as an error estimator related to the time-oscillation of the source term.

In view of the space-time adaptive time-marching algorithm considered in section 3.3, it is important to distinguish the space and time errors. To this purpose, similarly to [34, 43, 10], the diffusive flux estimator $\eta^n_{\mathrm{DF},T}(t)$ is split into two contributions using the triangle inequality. We define for all $1\le n\le N$

$$(\eta^n_{\mathrm{sp}})^2 := \sum_{T\in\mathcal{T}^n} 3\left\{\tau^n(9(\eta^n_{\mathrm{R},T}+\eta^n_{\mathrm{DF},1,T})^2+(\eta^n_{\mathrm{NC},2,T})^2)+\int_{I_n}(\eta^n_{\mathrm{NC},1,T})^2(t)\,\mathrm{d}t\right\}, \tag{3.11}$$

$$(\eta^n_{\mathrm{tm}})^2 := \sum_{T\in\mathcal{T}^n} 3\tau^n\|\nabla(s^n-s^{n-1})\|^2_T, \tag{3.12}$$

where

$$\eta^n_{\mathrm{DF},1,T} := \|\nabla s^n+\boldsymbol{\theta}^n\|_T. \tag{3.13}$$

THEOREM 3.6 (error upper bound distinguishing space and time errors). *Under the assumptions of Theorem* 3.2, *there holds*

$$\|u-u_{h\tau}\|_Y \le \left\{\sum_{n=1}^N (\eta^n_{\mathrm{sp}})^2\right\}^{1/2} + \left\{\sum_{n=1}^N (\eta^n_{\mathrm{tm}})^2\right\}^{1/2} + \eta_{\mathrm{IC}} + 3\|f-\widetilde{f}\|_{X'}.$$

Remark 3.7 (evaluation of η^n_{sp}). Since $s-u_{h\tau}$ is piecewise affine in time, the evaluation of the time integral of $\eta^n_{\mathrm{NC},1,T}$ in the expression of η^n_{sp} can be simplified into a sum involving the two quantities $\|\nabla^m(s^m-u^m_{h\tau})\|$ for $m\in\{n-1,n\}$; see Lemma 6.1.

3.2. Error lower bound. The goal of this section is to derive local-in-time upper bounds on the error estimators appearing in Theorem 3.6 in terms of the error $u-u_{h\tau}$, the data, and possibly some jump seminorms of the discrete solution $u_{h\tau}$.

These bounds are derived for a specific choice of the potential reconstruction s based on a suitable averaging interpolate and a specific approximation property for the flux reconstruction $\boldsymbol{\theta}$.

3.2.1. Assumptions on the space-time meshes. We assume the following conditions:

(M1) Shape regularity: the meshes $\{\mathcal{T}^n\}_{0\le n\le N}$ are shape regular uniformly in n.

(M2) Transition condition: for all $1 \le n \le N$, the commonly refined meshes $\mathcal{T}^{n-1,n}$ are also shape regular uniformly in n and such that

$$\Xi := \sup_{1\le n\le N} \sup_{T\in\mathcal{T}^{n-1}\cup\mathcal{T}^n} \sup_{T'\in\mathcal{T}^{n-1,n};T'\subset T} \frac{h_T}{h_{T'}} < +\infty. \tag{3.14}$$

(M3) For all $1 \le n \le N$, $(h^n)^2 \le \Upsilon\tau^n$ for a constant $\Upsilon > 0$.

We suppose for simplicity that all the meshes $\{\mathcal{T}^n\}_{0\le n\le N}$ are refinements of a base simplicial mesh $\mathcal{T}^*$, their shape regularity parameter is bounded by that of $\mathcal{T}^*$, and $\mathcal{T}^n$, $n \ge 1$, is a combination of refinement and coarsening of $\mathcal{T}^{n-1}$. For instance, the newest vertex bisection procedure [38, 31, 39] maintains these properties.

Remark 3.8 (transition condition). In the transition condition (3.14), $T \in \mathcal{T}^{n-1}$ in the second supremum restricts the refinement, while $T \in \mathcal{T}^n$ restricts the coarsening, in the sense that mesh changes in time cannot be too abrupt. In fact, we only need the restriction on coarsening for conforming methods, as classical in the context of conforming finite element methods; see [43]. For nonconforming methods, the proof of Lemmas 6.3 and 6.4 uses, in addition, the restriction on refinement.

Henceforth, C denotes a generic constant whose value can change at each occurrence and which can depend on the regularity of the meshes, the transition constant Ξ in (3.14), the constant Υ in (M3), and the polynomial degree used to build the approximation spaces V_h^n, but is independent of the meshsizes and the time steps. The inequality $A \le CB$ is often abbreviated as $A \lesssim B$.

3.2.2. Design of the potential reconstruction. Let $0 \le n \le N$. The averaging interpolation operator $\mathcal{I}_{\rm av}^n$ on the mesh $\mathcal{T}^n$ is classically constructed as the H_0^1-conforming Lagrange interpolate by prescribing at interior interpolation nodes averaged values of the piecewise discontinuous function to interpolate; see [24] for the h-analysis and [13] for the hp-analysis. We set

$$s^n := \mathcal{I}_{\rm av}^n(u_{h\tau}^n) + \sum_{T'\in\mathcal{T}^{n,n+1}} \alpha_{T'}^n b_{T'}, \tag{3.15}$$

where for all $T' \in \mathcal{T}^{n,n+1}$, $b_{T'}$ denotes the standard (time-independent) bubble function supported on T', defined as the product of the barycentric coordinates of T', and scaled so that its maximal value is 1. Observe that for all $0 \le n \le N$, $\mathcal{I}_{\rm av}^n(u_{h\tau}^n) \in H_0^1(\Omega)$ and observe that also the bubble part belongs to $H_0^1(\Omega)$, so that indeed $s^n \in H_0^1(\Omega)$ for all $0 \le n \le N$ and thus $s \in P_\tau^1(H_0^1(\Omega))$ as required in (3.1). For all $0 \le n \le N$ and for all $T' \in \mathcal{T}^{n,n+1}$, we set

$$\alpha_{T'}^n := \frac{1}{(b_{T'},1)_{T'}}(u_{h\tau}^n - \mathcal{I}_{\rm av}^n(u_{h\tau}^n),1)_{T'}. \tag{3.16}$$

This important choice guarantees that (3.2) holds.

3.2.3. Approximation property of the flux reconstruction. We assume that for all $1 \le n \le N$ and for all $T \in \mathcal{T}^n$,

$$\|\nabla u_{h\tau}^n + \boldsymbol{\theta}^n\|_T \lesssim \left\{ \sum_{T' \in \mathfrak{T}_T} h_{T'}^2 \|\widetilde{f}^n - \partial_t u_{h\tau}^n + \Delta u_{h\tau}^n\|_{T'}^2 \right\}^{1/2} + |\mathbf{n}\cdot[\![\nabla^n u_{h\tau}^n]\!]|_{+\frac{1}{2},\mathfrak{F}_T^{\mathrm{i},n}} + |[\![u_{h\tau}^n]\!]|_{-\frac{1}{2},\mathfrak{F}_T^n}. \tag{3.17}$$

The approximation property (3.17) thus means that the term $\|\nabla u_{h\tau}^n + \boldsymbol{\theta}^n\|_T$ is a lower bound for the classical residual-based estimator known from conforming methods (cf., e.g., [43]), and the jump term known from a posteriori error analysis of nonconforming methods (cf., e.g., [24]).

3.2.4. Optimally efficient lower bound. For all $1 \le n \le N$, we introduce the following jump seminorms (local-in-time and global-in-space)

$$\mathcal{J}_*^n(u_{h\tau})^2 := \tau^n \sum_{T \in \mathcal{T}^n} |[\![u_{h\tau}^n]\!]|_{-\frac{1}{2},\mathfrak{F}_T^n}^2, \tag{3.18}$$

and

$$\mathcal{J}^n(u_{h\tau})^2 := \tau^n \sum_{T \in \mathcal{T}^{n-1}} |[\![u_{h\tau}^{n-1}]\!]|_{-\frac{1}{2},\mathfrak{F}_T^{n-1}}^2 + \tau^n \sum_{T \in \mathcal{T}^n} |[\![u_{h\tau}^n]\!]|_{-\frac{1}{2},\mathfrak{F}_T^n}^2. \tag{3.19}$$

We also localize in time the X- and Y-norms as follows:

$$\|y\|_{X(I_n)}^2 := \int_{I_n} \|\nabla^{n-1,n} y\|^2(t)\,\mathrm{d}t, \qquad \|y\|_{Y(I_n)} := \|y\|_{X(I_n)} + \|\partial_t y\|_{X'(I_n)},$$

with $\|z\|_{X'(I_n)}^2 := \int_{I_n} \|z\|_{H^{-1}}^2(t)\,\mathrm{d}t$. Finally, we define for all $1 \le n \le N$ the space-time data oscillation term

$$(\mathcal{E}_f^n)^2 := \|f - \widetilde{f}\|_{X'(I_n)}^2 + \tau^n \sum_{T \in \mathcal{T}^n} h_T^2 \|\widetilde{f}^n - \Pi_{V_h^n} \widetilde{f}^n\|_T^2, \tag{3.20}$$

where we recall that $\Pi_{V_h^n}$ denotes the L^2-orthogonal projection onto V_h^n.

THEOREM 3.9 (error lower bound). *Assume that the meshes satisfy* (M1)–(M3)*, that the potential reconstruction s is defined by* (3.15)–(3.16)*, and that the flux reconstruction $\boldsymbol{\theta}$ satisfies* (3.17)*. Let $1 \le n \le N$. Let $\mathcal{E}_f^n$ be defined by* (3.20)*, and let the jump seminorm $\mathcal{J}^n(u_{h\tau})$ be defined by* (3.19)*. Finally, let η_{sp}^n and η_{tm}^n be defined by* (3.11)–(3.12)*. Then,*

$$\eta_{\mathrm{sp}}^n + \eta_{\mathrm{tm}}^n \lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{J}^n(u_{h\tau}) + \mathcal{E}_f^n. \tag{3.21}$$

Remark 3.10 (bound on the jumps). Further handling of the jump seminorm $\mathcal{J}^n(u_{h\tau})$ depends on the numerical scheme. For conforming methods, this jump seminorm vanishes. Furthermore, if the jumps have zero mean on each face of the meshes as in mixed finite elements, face-centered finite volumes, and nonconforming finite elements, it can be proven (see section 4.3) that

$$\mathcal{J}^n(u_{h\tau}) \lesssim \|u - u_{h\tau}\|_{X(I_n)}, \tag{3.22}$$

thereby closing the equivalence cycle between the error and the estimators. For discontinuous Galerkin methods, it is possible to bound the jump seminorm $\mathcal{J}_*^n(u_{h\tau})$ by the usual residual-based estimators for the heat equation, provided the jump penalty parameter is large enough; assuming that the ratios τ^n/τ^{n-1} are uniformly bounded, this yields (4.4), which is a global-in-time variant of (3.22); see section 4.1. Finally, for the cell-centered finite volumes examined in section 4.2, we leave the question of proving (3.22) open, since it must be addressed using the specific form of the diffusive fluxes in the finite volume scheme.

Remark 3.11 (condition (M3)). Condition (M3) is only needed for nonconforming methods in the proof of Lemma 6.4, and results from the fact that mesh modifications in time trigger changes in the jumps. On fixed meshes, the condition can be avoided provided an additional jump seminorm $\{\tau^n \sum_{T\in\mathcal{T}^n} h_T^2 |[\![\partial_t u_{h\tau}^n]\!]|^2_{+\frac{1}{2},\mathfrak{F}_T^n}\}^{1/2}$ is added to the right-hand side of (3.21).

3.3. A space-time adaptive time-marching algorithm. We briefly outline here a space-time adaptive time-marching procedure to be used in conjunction with the above analysis. Let $\mathrm{Sol}(\mathtt{P},\mathtt{TS},\mathtt{Mesh})$ denote the discrete solution produced by the numerical scheme starting from a previous approximation `P` over a single time step `TS` and using a mesh `Mesh`. The previous approximation need not be defined on the mesh `Mesh`; the way this is handled depends on the numerical scheme. We discard for simplicity the influence of the time-oscillation of the source term $\|f-\widetilde{f}\|_{X'}$. Then, a possible adaptive algorithm designed to (1) achieve a prescribed tolerance ε for the relative error,

$$\frac{\sum_{n=1}^N \{(\eta_{\mathrm{sp}}^n)^2 + (\eta_{\mathrm{tm}}^n)^2\}}{\sum_{n=1}^N \|u_{h\tau}\|^2_{Z(I_n)}} \le \varepsilon^2, \tag{3.23}$$

where $\|\cdot\|_{Z(I_n)}$ is a certain (semi)norm over $\Omega\times I_n$, and (2) make the calculation efficient through equilibration of the space and time errors is as follows:

1. Initialization:
   (a) choose an initial mesh $\mathcal{T}^0$ with corresponding space $V_h^0 := V(\mathcal{T}^0)$ and an initial approximation $u_{h\tau}^0 \in V_h^0$ such that the initial condition estimator η_{IC} is below a prescribed tolerance (negligible for the sake of simplicity);
   (b) select an initial time step τ^0 and set $n := 1$.
2. Loop in time: While $\sum_i \tau^i < t_{\mathrm{F}}$,
   (a) set $\mathcal{T}^{n*} := \mathcal{T}^{n-1}$ with the corresponding space V_h^{n*} and set $\tau^{n*} := \tau^{n-1}$;
   (b) solve $u_{h\tau}^{n*} := \mathrm{Sol}(u_{h\tau}^{n-1}, \tau^{n*}, \mathcal{T}^{n*})$;
   (c) estimate the space and time errors by η_{sp}^n and η_{tm}^n;
   (d) when η_{sp}^n or η_{tm}^n are too much above or below $\varepsilon\|u_{h\tau}\|_{Z(I_n)}/\sqrt{2}$ or when η_{sp}^n and η_{tm}^n are not of similar size, refine or redefine the time step τ^{n*} and the space mesh $\mathcal{T}^{n*}$ in such a way that conditions (M1)–(M3) hold, and return to step (2b);
   (e) when both η_{sp}^n and η_{tm}^n are slightly below $\varepsilon\|u_{h\tau}\|_{Z(I_n)}/\sqrt{2}$ and η_{sp}^n and η_{tm}^n are of similar size, save the approximate solution, mesh, and time step as $u_{h\tau}^n$, $\mathcal{T}^n$, and τ^n, and set $n := n+1$.

A theoretical analysis showing that, in the above algorithm, the loop in steps (2b)–(2d) terminates successfully and that the final time t_{F} is reached goes beyond the present scope. For a more detailed version of the above algorithm and for some numerical experiments, we refer the reader to [22].

Remark 3.12 (choice of $\|\cdot\|_{Z(I_n)}$ in (3.23)). The natural candidate for $\|\cdot\|_{Z(I_n)}$ in (3.23) is the norm $\|\cdot\|_{Y(I_n)}$. This norm is, however, not computable, so that some approximation is to be used in practice.

Remark 3.13 (time-marching evaluation of s). Condition (3.2) uses the mesh $\mathcal{T}^{n+1}$, which is not yet known at the discrete time t^n. In practice, it suffices to adjust the mean values of s^n only at the elements of the current mesh $\mathcal{T}^n$ (putting temporarily $\mathcal{T}^{n+1} := \mathcal{T}^n$). Then, at the next time step t^{n+1}, and only if there are elements of $\mathcal{T}^{n+1}$ which are refinements of those of $\mathcal{T}^n$, additional bubble functions are added to s^n, as prescribed by (3.15)–(3.16), before evaluating the error at t^{n+1}.

4. Applications. We apply in this section the preceding results to different space discretization schemes. For the upper bound of Theorem 3.2, this is done by specifying the reconstructed flux $\boldsymbol{\theta}$ and checking that the conservation property (3.4) holds true. For the lower bound of Theorem 3.9, we need to check the approximation property (3.17). The reconstructed potential s in nonconforming methods is always given by (3.15)–(3.16). In conforming methods, where $u_{h\tau}^n \in H_0^1(\Omega)$ for all $0 \le n \le N$, we set $s^n := u_{h\tau}^n$, so that (3.2) immediately holds true.

The reconstructed flux $\boldsymbol{\theta}^n$ on all time levels $1 \le n \le N$ will belong to the Raviart–Thomas–Nédélec (RTN) spaces of vector functions on the mesh $\mathcal{T}^n$,

$$\mathbf{RTN}_l(\mathcal{T}^n) := \big\{\mathbf{v}_h \in \mathbf{H}(\mathrm{div},\Omega)\,;\mathbf{v}_h|_T \in \mathbf{RTN}_l(T) \quad \forall T \in \mathcal{T}^n\big\},$$

where $\mathbf{RTN}_l(T) := [\mathbb{P}_l(T)]^d + \mathbf{x}\mathbb{P}_l(T)$, $l \ge 0$. In particular, $\mathbf{v}_h \in \mathbf{RTN}_l(\mathcal{T}^n)$ is such that $\nabla\cdot\mathbf{v}_h \in \mathbb{P}_l(T)$ for all $T \in \mathcal{T}^n$, $\mathbf{v}_h\cdot\mathbf{n}_F \in \mathbb{P}_l(F)$ for all $F \in \mathcal{F}^n$, and such that its normal trace is continuous; cf. [12]. In certain cases, a submesh of $\mathcal{T}^n$ will be used instead for the construction of the flux $\boldsymbol{\theta}$. We use the notation $\mathbb{P}_k(T)$ for the vector space spanned by polynomials of total degree $\le k$ on T and $\mathbb{P}_k(\mathcal{T}^n)$ for the vector space spanned by discontinuous piecewise polynomials of total degree $\le k$ on $\mathcal{T}^n$.

The computational cost of the estimators can be split into two parts. First, the flux $\boldsymbol{\theta}^n$ and/or the potential s^n have to be constructed. In all the applications below, this is done by simply prescribing the corresponding degrees of freedom. For vertex-centered finite volumes of section 4.4 (and closely related conforming finite elements of Appendix A), a solution of a local linear system on a patch around each vertex may be preferable instead for the local construction of $\boldsymbol{\theta}^n$. Second, the estimators (3.5)–(3.10) or (3.11), (3.12), and (3.10) have to be evaluated, and this is manageable at marginal costs using quadrature formulae. Note that all these procedures are local and that the overall cost is linear in the number of mesh elements on each I_n.

4.1. Discontinuous Galerkin. The discontinuous Galerkin method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n := \mathbb{P}_k(\mathcal{T}^n)$, $k \ge 1$, such that

$$\begin{aligned}&(\partial_t u_{h\tau}^n, v_h) - \sum_{F\in\mathcal{F}^n}\{(\mathbf{n}_F\cdot\{\!\{\nabla^n u_{h\tau}^n\}\!\},[\![v_h]\!])_F + \theta(\mathbf{n}_F\cdot\{\!\{\nabla^n v_h\}\!\},[\![u_{h\tau}^n]\!])_F\}\\ &+ (\nabla^n u_{h\tau}^n,\nabla^n v_h) + \sum_{F\in\mathcal{F}^n}(\alpha_F h_F^{-1}[\![u_{h\tau}^n]\!],[\![v_h]\!])_F = (\widetilde{f}^n, v_h) \qquad \forall v_h \in V_h^n.\end{aligned} \tag{4.1}$$

Here $\theta \in \{-1,0,1\}$, and the quantities α_F are positive parameters. Note that $\partial_t u_{h\tau}^n$ is defined by (2.3) and that $u_{h\tau}^{n-1} \in V_h^{n-1}$ is linked to the mesh $\mathcal{T}^{n-1}$, which may be different from $\mathcal{T}^n$. For the discretization of the initial condition, we set $u_{h\tau}^0 := \Pi_{V_h^0}(u^0)$. As in [15, 26], we use the following definition of $\boldsymbol{\theta}^n \in \mathbf{RTN}_l(\mathcal{T}^n)$,

$l \in \{k-1, k\}$: For all $T \in \mathcal{T}^n$ and all $F \in \mathcal{F}_T^n$, we prescribe the face degrees of freedom of the reconstructed flux $\boldsymbol{\theta}^n$ by the natural discontinuous Galerkin face fluxes,

$$(\boldsymbol{\theta}^n \cdot \mathbf{n}_F, q_h)_F = \big(-\mathbf{n}_F \cdot \{\!\{\nabla^n u_{h\tau}^n\}\!\} + \alpha_F h_F^{-1} [\![u_{h\tau}^n]\!], q_h\big)_F \qquad \forall q_h \in \mathbb{P}_l(F). \tag{4.2}$$

Moreover, when $l \geq 1$, we prescribe for all $T \in \mathcal{T}^n$ the element degrees of freedom of the reconstructed flux $\boldsymbol{\theta}^n$ by

$$(\boldsymbol{\theta}^n, \mathbf{r}_h)_T = -(\nabla^n u_{h\tau}^n, \mathbf{r}_h)_T + \theta \sum_{F \in \mathcal{F}_T^n} \omega_F (\mathbf{n}_F \cdot \mathbf{r}_h, [\![u_{h\tau}^n]\!])_F \qquad \forall \mathbf{r}_h \in [\mathbb{P}_{l-1}(T)]^d, \tag{4.3}$$

where $\omega_F := \frac{1}{2}$ for $F \in \mathcal{F}^{\mathrm{i},n}$ and $\omega_F := 1$ for $F \in \mathcal{F}^{\mathrm{b},n}$. Taking $v_h \in \mathbb{P}_0(\mathcal{T}^n)$ with support on a single mesh element in (4.1) yields (3.4) as in [15, Theorem 3.1]. Furthermore, the approximation property (3.17) can be proven as in [16]; it is only a consequence of (4.2)–(4.3) and there is no elementwise residual in the upper bound.

Bounding the jump seminorm by the energy norm is a known property of discontinuous Galerkin methods in the elliptic case [3, 25]; hereafter, we present an original approach to extend this property to the parabolic case. Let $1 \leq n \leq N$. We integrate by parts the quantity $(\nabla^n u_{h\tau}^n, \nabla^n v_h)$ in (4.1) and rearrange terms to obtain

$$\begin{aligned}\sum_{F \in \mathcal{F}^n} (\alpha_F h_F^{-1} [\![u_{h\tau}^n]\!], [\![v_h]\!])_F = \sum_{T \in \mathcal{T}^n} (\Pi_{V_h^n} \widetilde{f}^n - \partial_t u_{h\tau}^n - \Delta u_{h\tau}^n, v_h)_T \\ - \sum_{F \in \mathcal{F}^{\mathrm{i},n}} (\mathbf{n}_F \cdot [\![\nabla^n u_{h\tau}^n]\!], \{\!\{v_h\}\!\})_F + \sum_{F \in \mathcal{F}^n} \theta (\mathbf{n}_F \cdot \{\!\{\nabla^n v_h\}\!\}, [\![u_{h\tau}^n]\!])_F.\end{aligned}$$

Choosing $v_h = u_{h\tau}^n - \mathcal{I}_{\mathrm{av}}^n(u_{h\tau}^n)$, using the Cauchy–Schwarz inequality, and using the approximation properties of $\mathcal{I}_{\mathrm{av}}^n$ stated in the proof of Lemma 6.3 (details are skipped for brevity), it is inferred that

$$\begin{aligned}\sum_{F \in \mathcal{F}^n} \alpha_F h_F^{-1} \|[\![u_{h\tau}^n]\!]\|_F^2 \lesssim \Bigg\{ \sum_{T \in \mathcal{T}^n} h_T^2 \|\Pi_{V_h^n} \widetilde{f}^n - \partial_t u_{h\tau}^n - \Delta u_{h\tau}^n\|_T^2 \\ + \sum_{F \in \mathcal{F}^{\mathrm{i},n}} h_F \|\mathbf{n}_F \cdot [\![\nabla^n u_{h\tau}^n]\!]\|_F^2 \Bigg\}^{1/2} \times \Bigg\{ \sum_{F \in \mathcal{F}^n} h_F^{-1} \|[\![u_{h\tau}^n]\!]\|_F^2 \Bigg\}^{1/2} + \sum_{F \in \mathcal{F}^n} h_F^{-1} \|[\![u_{h\tau}^n]\!]\|_F^2.\end{aligned}$$

Hence, if the penalty parameters $\{\alpha_F\}_{F \in \mathcal{F}^n}$ are large enough,

$$\sum_{F \in \mathcal{F}^n} h_F^{-1} \|[\![u_{h\tau}^n]\!]\|_F^2 \lesssim \sum_{T \in \mathcal{T}^n} h_T^2 \|\Pi_{V_h^n} \widetilde{f}^n - \partial_t u_{h\tau}^n - \Delta u_{h\tau}^n\|_T^2 + \sum_{F \in \mathcal{F}^{\mathrm{i},n}} h_F \|\mathbf{n}_F \cdot [\![\nabla^n u_{h\tau}^n]\!]\|_F^2.$$

Considering the usual residual-based a posteriori error estimators $\mathcal{E}_{\mathrm{R}}^n$ and $\mathcal{E}_{\mathrm{J}}^n$ defined by (6.1) and (6.2), respectively, we finally arrive at the estimate

$$\mathcal{J}_*^n(u_{h\tau})^2 \lesssim (\mathcal{E}_{\mathrm{R}}^n)^2 + (\mathcal{E}_{\mathrm{J}}^n)^2.$$

Then squaring (6.6) with ϵ sufficiently small, summing over n, and assuming that the ratios τ^n/τ^{n-1} are uniformly bounded, we obtain

$$\sum_{n=1}^N \mathcal{J}_*^n(u_{h\tau})^2 \lesssim \|u - u_{h\tau}\|_Y^2 + \sum_{n=1}^N (\mathcal{E}_f^n)^2 + \tau^1 \sum_{T \in \mathcal{T}^0} |[\![u_{h\tau}^0]\!]|_{-\frac{1}{2}, \mathfrak{F}_T^0}^2, \tag{4.4}$$

which is a global-in-time variant of (3.22).

Remark 4.1 (residual estimator). A consequence of (4.1) and (4.2)–(4.3) is that, for $l = k$, there holds

$$\Pi_{V_h^n}(\partial_t u_{h\tau}^n) + \nabla\cdot\boldsymbol{\theta}^n = \Pi_{V_h^n}\widetilde{f}^n. \tag{4.5}$$

When $\mathcal{T}^{n-1} = \mathcal{T}^n$, $\Pi_{V_h^n}(\partial_t u_{h\tau}^n) = \partial_t u_{h\tau}^n$, so that $\eta_{\mathrm{R},T}^n \le C_{\mathrm{P}} h_T \|\widetilde{f}^n - \Pi_{V_h^n}\widetilde{f}^n\|_T + \eta_{\mathrm{NC},2,T}^n$, where the first term is clearly superconvergent under appropriate assumptions on the smoothness of $\widetilde{f}^n$.

4.2. Cell-centered finite volumes. A general cell-centered finite volume method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ (cf. [19]) reads as follows: Find $\bar{u}_{h\tau}^n \in \bar{V}_h^n := \mathbb{P}_0(\mathcal{T}^n)$ such that

$$\frac{1}{\tau^n}(\bar{u}_{h\tau}^n - u_{h\tau}^{n-1}, 1)_T + \sum_{F\in\mathcal{F}_T^n} S_{T,F}^n = (\widetilde{f}^n, 1)_T \qquad \forall\, T \in \mathcal{T}^n. \tag{4.6}$$

Here $S_{T,F}^n$ is the diffusive flux through the face F of an element T, which depends linearly on $\bar{u}_{h\tau}^n$. For our a posteriori error estimates we do not need the specific form of $S_{T,F}^n$, as long as the conservation property $S_{T^-,F}^n = -S_{T^+,F}^n$ holds for all $F = \partial T^- \cap \partial T^+ \in \mathcal{F}^{\mathrm{i},n}$. A simple example is the so-called two-point scheme; cf. [19]. The functions $u_{h\tau}^{n-1}$, $n \ge 1$, are specified below.

The original finite volume approximation $\bar{u}_{h\tau}^n$ is not suitable for energy norm-type a posteriori error estimates since $\nabla^n \bar{u}_{h\tau}^n = 0$. Following [47], we introduce its postprocessing $u_{h\tau}^n$. To this purpose, we first define $\boldsymbol{\theta}^n \in \mathbf{RTN}_0(\mathcal{T}^n)$ from the finite volume fluxes $S_{T,F}^n$ by $(\boldsymbol{\theta}^n\cdot\mathbf{n}_T, 1)_F := S_{T,F}^n$ for all $T \in \mathcal{T}^n$ and all $F \in \mathcal{F}_T^n$, $1 \le n \le N$. Here, $\mathbf{n}_T$ denotes the outward unit normal to T. Next, for $0 \le n \le N$, we define $u_{h\tau}^n \in V_h^n := \mathbb{P}_{1,2}(\mathcal{T}^n)$, where $\mathbb{P}_{1,2}(\mathcal{T}^n)$ is the space $\mathbb{P}_1(\mathcal{T}^n)$ enriched elementwise by the parabolas $\sum_{i=1}^d x_i^2$, such that for all $T \in \mathcal{T}^n$,

$$-\nabla^n u_{h\tau}^n = \boldsymbol{\theta}^n, \qquad (u_{h\tau}^n, 1)_T = (\bar{u}_{h\tau}^n, 1)_T. \tag{4.7}$$

Such a postprocessing is local and its cost is negligible.

Note that by (4.7), $u_{h\tau}^{n-1}$, $n \ge 2$, can be replaced by the usual term $\bar{u}_{h\tau}^{n-1}$ in (4.6) when the two meshes $\mathcal{T}^{n-1}$ and $\mathcal{T}^n$ are the same. We use $u_{h\tau}^{n-1}$ instead of $\bar{u}_{h\tau}^{n-1}$ for two reasons. First, under the form (4.6), the scheme enters exactly the present framework; a minor modification would be necessary otherwise. Second, since $u_{h\tau}^{n-1}$ is a kind of regularization of the piecewise constant function $\bar{u}_{h\tau}^{n-1}$, we think that it is better suited for being evaluated on the different mesh $\mathcal{T}^n$. Finally, we set $u_{h\tau}^0 := \Pi_{V_h^0}(u^0)$.

With the above constructions, (4.6) immediately yields (3.4) using the Green theorem. By (4.7), $\nabla^n u_{h\tau}^n + \boldsymbol{\theta}^n = 0$, whence (3.17) is trivially satisfied.

4.3. Mixed finite elements. The mixed finite element method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $\boldsymbol{\sigma}_{h\tau}^n \in \mathbf{W}_h^n$, an approximation to the exact flux $-\nabla u(\cdot, t^n)$, and $\bar{u}_{h\tau}^n \in \bar{V}_h^n$, an approximation to the exact potential $u(\cdot, t^n)$, such that

$$(\boldsymbol{\sigma}_{h\tau}^n, \mathbf{w}_h) - (\bar{u}_{h\tau}^n, \nabla\cdot\mathbf{w}_h) = 0 \qquad \forall\, \mathbf{w}_h \in \mathbf{W}_h^n, \tag{4.8a}$$

$$(\nabla\cdot\boldsymbol{\sigma}_{h\tau}^n, v_h) + \frac{1}{\tau^n}(\bar{u}_{h\tau}^n - u_{h\tau}^{n-1}, v_h) = (\widetilde{f}^n, v_h) \quad \forall v_h \in \bar{V}_h^n. \tag{4.8b}$$

The couple $\mathbf{W}_h^n \times \bar{V}_h^n$ can be any of the usual mixed finite element spaces; cf. [12]. In particular, in the RTN method, $\mathbf{W}_h^n = \mathbf{RTN}_l(\mathcal{T}^n)$ and $\bar{V}_h^n = \mathbb{P}_l(\mathcal{T}^n)$, $l \geq 0$. As for the cell-centered finite volume method, $\bar{u}_{h\tau}^n$ is not suitable to be used in the present a posteriori error estimates. As before, we postprocess it to a more regular potential approximation $u_{h\tau}^n$. In the lowest-order RTN case ($\mathbf{W}_h^n = \mathbf{RTN}_0(\mathcal{T}^n)$), following [44], we define $u_{h\tau}^n \in \mathbb{P}_{1,2}(\mathcal{T}^n)$ by (4.7), with $\boldsymbol{\theta}^n$ replaced by $\boldsymbol{\sigma}_{h\tau}^n$. In higher order cases, following [6, 5, 48] we define a postprocessed $u_{h\tau}^n \in V_h^n$, where the spaces V_h^n are specified in [5]. As an example, in the RTN case with l even, V_h^n is $\mathbb{P}_{l+1}(\mathcal{T}^n)$ enriched by bubbles. We then prescribe $u_{h\tau}^n \in V_h^n$ by requiring

$$\Pi_{\mathbf{W}_h^n}(-\nabla^n u_{h\tau}^n) = \boldsymbol{\sigma}_{h\tau}^n, \qquad \Pi_{\bar{V}_h^n}(u_{h\tau}^n) = \bar{u}_{h\tau}^n. \tag{4.9}$$

Here, $\Pi_{\mathbf{W}_h^n}$ and $\Pi_{\bar{V}_h^n}$ are, respectively, the L^2-orthogonal projections onto $\mathbf{W}_h^n$ and $\bar{V}_h^n$. Equation (4.9) represents a higher order polynomial equivalent of (4.7). As in the previous section, $u_{h\tau}^{n-1}$ in (4.8b) can be replaced by $\bar{u}_{h\tau}^{n-1}$ whenever the two meshes $\mathcal{T}^{n-1}$ and $\mathcal{T}^n$ coincide. Finally, we set $u_{h\tau}^0 := \Pi_{V_h^0}(u^0)$.

In order to use our a posteriori error estimates, we set $\boldsymbol{\theta}^n := \boldsymbol{\sigma}_{h\tau}^n$. Taking $v_h \in \mathbb{P}_0(\mathcal{T}^n)$ in (4.8b) and considering (4.9) yields (3.4). In the lowest order RTN case, we have $\nabla^n u_{h\tau}^n + \boldsymbol{\theta}^n = 0$, whence (3.17) is trivial. For higher order cases, the present theory does not seem to lead to a condition of the form (3.17). However, owing to (4.9), $\nabla^n u_{h\tau}^n + \boldsymbol{\theta}^n$ is expected to be negligible and act as a numerical quadrature. Finally, to prove (3.22), we first use the fact that $u_{h\tau}$ is piecewise affine in time to infer

$$\mathcal{J}^n(u_{h\tau})^2 \lesssim \int_{I_n} \sum_{F \in \mathcal{F}^{n-1,n}} h_F^{-1} \|[\![u_{h\tau}]\!]\|_F^2(t)\,\mathrm{d}t,$$

where $\mathcal{F}^{n-1,n}$ collects the faces of the mesh $\mathcal{T}^{n-1,n}$. Observe that the jumps of $u_{h\tau}^n$ on all faces of $\mathcal{F}^n$ have zero mean value. This follows from (4.8a) and (4.9) as in [48]. The same holds for the jumps of $u_{h\tau}^{n-1}$ on all faces of $\mathcal{F}^{n-1}$. Because $\mathcal{T}^n$ is a combination of refinement and coarsening of $\mathcal{T}^{n-1}$, the jumps of $u_{h\tau}(t)$, $t \in I_n$, have zero mean values only on suitable collections of faces of either $\mathcal{F}^{n-1}$ or $\mathcal{F}^n$; we denote this set by $\mathcal{F}^{n-1,n,*}$. By the transition condition (M2), we obtain from the previous bound

$$\mathcal{J}^n(u_{h\tau})^2 \lesssim \int_{I_n} \sum_{\mathcal{F}^{n-1,n,*}} h_F^{-1} \|[\![u_{h\tau}]\!]\|_F^2(t)\,\mathrm{d}t.$$

Then, proceeding as in [1, Theorem 10] for each $t \in I_n$ (details are skipped for brevity), it is inferred that

$$\mathcal{J}^n(u_{h\tau})^2 \lesssim \int_{I_n} \|\nabla^{n-1,n}(u - u_{h\tau})\|^2(t)\,\mathrm{d}t = \|u - u_{h\tau}\|_{X(I_n)}^2.$$

Remark 4.2 (residual estimator). Similarly to (4.5) for discontinuous Galerkin methods, there holds for mixed finite elements $\Pi_{\bar{V}_h^n}(\partial_t u_{h\tau}^n) + \nabla\cdot\boldsymbol{\theta}^n = \Pi_{\bar{V}_h^n} \widetilde{f}^n$.

4.4. Vertex-centered finite volumes. The vertex-centered finite volume method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \leq n \leq N$, and the corresponding mesh $\mathcal{T}^n$ reads as follows: Find $u_{h\tau}^n \in V_h^n := \mathbb{P}_1(\mathcal{T}^n) \cap H_0^1(\Omega)$ such that

$$(\partial_t u_{h\tau}^n, 1)_D - (\nabla u_{h\tau}^n \cdot \mathbf{n}_D, 1)_{\partial D} = (\widetilde{f}^n, 1)_D \qquad \forall D \in \mathcal{D}^{\mathrm{i},n}. \tag{4.10}$$

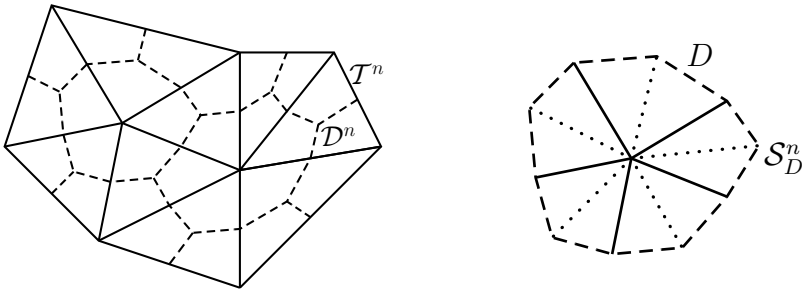

FIG. 4.1. *Simplicial mesh $\mathcal{T}^n$ and the associated vertex-centered dual mesh $\mathcal{D}^n$ (left) and the fine simplicial mesh $\mathcal{S}_D^n$ of $D \in \mathcal{D}^n$ (right).*

Here $\mathcal{D}^n$ is the dual mesh around vertices created using face and element barycenters as indicated in the left part of Figure 4.1; $\mathcal{D}^{\mathrm{i},n}$ corresponds to the interior vertices, $\mathcal{D}^{\mathrm{b},n}$ to the boundary ones, and $\mathbf{n}_D$ denotes the outward unit normal to D. Finally, we set $u_{h\tau}^0 := \Pi_{V_h^0}(u^0)$.

The reconstructed flux $\boldsymbol{\theta}^n$ is prescribed locally on the fine simplicial mesh $\mathcal{S}_D^n$ of each $D \in \mathcal{D}^n$, indicated in the right part of Figure 4.1. Note that the mesh $\mathcal{S}^n$, formed by the meshes $\mathcal{S}_D^n$ for all $D \in \mathcal{D}^n$, is a matching simplicial submesh of both $\mathcal{T}^n$ and $\mathcal{D}^n$. Let $D \in \mathcal{D}^n$, and let $\mathbf{RTN}_0^{\mathrm{N}}(\mathcal{S}_D^n)$ denote the space of such $\mathbf{v}_h \in \mathbf{RTN}_0(\mathcal{S}_D^n)$ that $\boldsymbol{\theta}^n{\cdot}\mathbf{n}_F|_F = -\nabla u_{h\tau}^n{\cdot}\mathbf{n}_F|_F$ on all faces F included in ∂D but not in $\partial\Omega$. We will always require that $\boldsymbol{\theta}^n|_D \in \mathbf{RTN}_0^{\mathrm{N}}(\mathcal{S}_D^n)$ for all $D \in \mathcal{D}^n$. Note that, consequently, $\boldsymbol{\theta}^n \in \mathbf{RTN}_0(\mathcal{S}^n)$ and, using (4.10), (3.4) holds true on all $D \in \mathcal{D}^{\mathrm{i},n}$. There are two possibilities to define $\boldsymbol{\theta}^n{\cdot}\mathbf{n}_F|_F$ on the remaining faces of $\mathcal{S}^n$ (that is, those faces located inside some $D \in \mathcal{D}^n$ and those located on the boundary $\partial\Omega$). First, we can directly prescribe $\boldsymbol{\theta}^n{\cdot}\mathbf{n}_F|_F := -\{\!\!\{\nabla u_{h\tau}^n{\cdot}\mathbf{n}_F\}\!\!\}$ on these faces; the present a posteriori error estimates can then be used on the dual meshes $\mathcal{D}^n$ as in [46], but not on the original simplicial ones $\mathcal{T}^n$. Alternatively, following [17, 45], define $\mathbf{RTN}_0^{\mathrm{N},0}(\mathcal{S}_D^n)$ as $\mathbf{RTN}_0^{\mathrm{N}}(\mathcal{S}_D^n)$ but with the normal flux condition $\mathbf{v}_h{\cdot}\mathbf{n}_F = 0$ on all faces F included in ∂D but not in $\partial\Omega$. Let $\mathbb{P}_0^*(\mathcal{S}_D^n)$ be spanned by piecewise constants on $\mathcal{S}_D^n$ with zero mean on D when $D \in \mathcal{D}^{\mathrm{i},n}$; when $D \in \mathcal{D}^{\mathrm{b},n}$, the mean value condition is not imposed. Then we can solve the following local mixed finite element problems on each $\mathcal{S}_D^n$, $D \in \mathcal{D}^n$: Find $\boldsymbol{\theta}^n \in \mathbf{RTN}_0^{\mathrm{N}}(\mathcal{S}_D^n)$ and $q_h \in \mathbb{P}_0^*(\mathcal{S}_D^n)$, the mixed finite element approximations of local Neumann problems on $D \in \mathcal{D}^{\mathrm{i},n}$ and local Neumann/Dirichlet problems on $D \in \mathcal{D}^{\mathrm{b},n}$:

$$(\boldsymbol{\theta}^n + \nabla u_{h\tau}^n, \mathbf{v}_h)_D - (q_h, \nabla{\cdot}\mathbf{v}_h)_D = 0 \qquad \forall\, \mathbf{v}_h \in \mathbf{RTN}_0^{\mathrm{N},0}(\mathcal{S}_D^n), \tag{4.11a}$$

$$(\nabla{\cdot}\boldsymbol{\theta}^n, \phi_h)_D - (\widetilde{f}^n - \partial_t u_{h\tau}^n, \phi_h)_D = 0 \qquad \forall\, \phi_h \in \mathbb{P}_0^*(\mathcal{S}_D^n). \tag{4.11b}$$

This leads to a local linear system solution on all $D \in \mathcal{D}^n$. Note that the right-hand side in (4.11b) is the local residual, since $\Delta u_{h\tau}^n = 0$ locally on each mesh element. The key advantage is that solving these local linear systems to reconstruct the flux more accurately ensures that (3.4) is satisfied on all elements T of the original mesh $\mathcal{T}^n$ (in fact on all elements T of the submesh $\mathcal{S}^n$). Finally, (3.17) follows as in [45, Theorem 4.5].

4.5. Face-centered finite volumes. The face-centered finite volume method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the

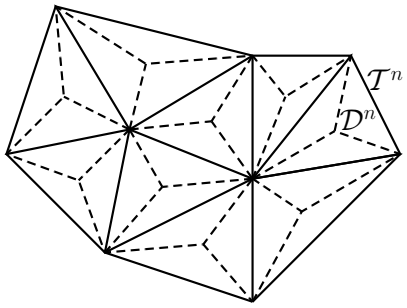

FIG. 4.2. *Simplicial mesh $\mathcal{T}^n$ and the associated face-centered dual mesh $\mathcal{D}^n$.*

corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n$ such that

$$(\partial_t u_{h\tau}^n, 1)_D - (\nabla^n u_{h\tau}^n \cdot \mathbf{n}_D, 1)_{\partial D} = (\widetilde{f}^n, 1)_D \qquad \forall D \in \mathcal{D}^{\mathrm{i},n}. \tag{4.12}$$

Here V_h^n is the Crouzeix–Raviart space of piecewise linear polynomials such that the face jumps have zero mean value, and $\mathcal{D}^n$ is the dual mesh around faces created using vertices and element barycenters as indicated in Figure 4.2; $\mathcal{D}^{\mathrm{i},n}$ corresponds to the interior faces and $\mathcal{D}^{\mathrm{b},n}$ to the boundary ones. We set $u_{h\tau}^0 := \Pi_{V_h^0}(u^0)$. We define $\mathcal{S}_D^n$ for all $D \in \mathcal{D}^{\mathrm{i},n}$ as the union of the two simplices sharing the associated face; $\mathcal{S}_D^n := \{D\}$ for $D \in \mathcal{D}^{\mathrm{b},n}$. We again look for $\boldsymbol{\theta}^n \in \mathbf{RTN}_0(\mathcal{S}^n)$. We set $\boldsymbol{\theta}^n \cdot \mathbf{n}_F|_F := -\nabla^n u_{h\tau}^n \cdot \mathbf{n}_F|_F$ on all faces F of the dual mesh $\mathcal{D}^n$. As in the previous section, we have several ways of defining the remaining fluxes, and we easily get (3.4) using the Green theorem from (4.12). Finally, (3.17) can be shown using the techniques of [45] and (3.22) follows as in section 4.3.

5. Proof of the error upper bound. The goal of this section is to prove Theorems 3.2 and 3.6. The proof is decomposed into several steps.

5.1. Abstract $\|\cdot\|_Y$-norm error estimate. We begin with an abstract $\|\cdot\|_Y$-norm error estimate that can be formulated in a rather general setting for the potential reconstruction. In particular, assumptions (3.1), (3.2), and (3.4) are not yet required.

Let $s \in X \cap H^1(0, t_{\mathrm{F}}; L^2(\Omega))$, so that $\partial_t s \in L^2(0, t_{\mathrm{F}}; L^2(\Omega))$ and hence $s \in Y$. We define the residual $\mathcal{R}(s) \in X'$ such that for all $\varphi \in X$,

$$\langle \mathcal{R}(s), \varphi \rangle_{X',X} := \int_0^{t_{\mathrm{F}}} \{(f - \partial_t s, \varphi) - (\nabla s, \nabla \varphi)\}(t)\,\mathrm{d}t. \tag{5.1}$$

LEMMA 5.1 (abstract $\|\cdot\|_Y$-norm error estimate). *Let $s \in X \cap H^1(0, t_{\mathrm{F}}; L^2(\Omega))$. Then,*

$$\|u - u_{h\tau}\|_Y \le \|s - u_{h\tau}\|_Y + 3\|\mathcal{R}(s)\|_{X'} + 2^{1/2}\|s^0 - u^0\|. \tag{5.2}$$

Proof. We first bound $\|u - s\|_Y$. Since $u - s$ is in Y, there holds (see, e.g., [18, Theorem 5.9.3])

$$\tfrac{1}{2}\|u - s\|^2(T) = \tfrac{1}{2}\|u^0 - s^0\|^2 + \int_0^{t_{\mathrm{F}}} \langle \partial_t(u - s), u - s \rangle(t)\,\mathrm{d}t.$$

As a result,

$$\begin{aligned}\|u-s\|_X^2 &\le \tfrac{1}{2}\|u-s\|^2(T) + \|u-s\|_X^2\\ &= \tfrac{1}{2}\|u^0-s^0\|^2 + \int_0^{t_{\mathrm{F}}} \{\langle \partial_t(u-s), u-s\rangle + (\nabla(u-s), \nabla(u-s))\}(t)\,\mathrm{d}t\\ &= \tfrac{1}{2}\|u^0-s^0\|^2 + \int_0^{t_{\mathrm{F}}} \{(f-\partial_t s, u-s) - (\nabla s, \nabla(u-s))\}(t)\,\mathrm{d}t,\end{aligned}$$

where we have used (2.1). Using the definition (5.1) of $\mathcal{R}(s)$ yields

$$\|u-s\|_X^2 \le \|\mathcal{R}(s)\|_{X'}\|u-s\|_X + \tfrac{1}{2}\|s^0-u^0\|^2.$$

Since $x^2 \le ax + b^2$ implies $x \le a+b$ for positive a and b, it is inferred that

$$\|u-s\|_X \le \|\mathcal{R}(s)\|_{X'} + 2^{-1/2}\|s^0-u^0\|.$$

Now let $\varphi \in X$ with $\|\varphi\|_X = 1$ and observe that

$$\langle \partial_t(u-s), \varphi\rangle_{X',X} = \int_0^{t_{\mathrm{F}}} \{(f-\partial_t s, \varphi) - (\nabla s, \nabla\varphi) - (\nabla(u-s), \nabla\varphi)\}(t)\,\mathrm{d}t,$$

whence

$$\|\partial_t(u-s)\|_{X'} \le \|\mathcal{R}(s)\|_{X'} + \|u-s\|_X,$$

so that $\|u-s\|_Y \le 3\|\mathcal{R}(s)\|_{X'} + 2^{1/2}\|s^0-u^0\|$. The triangle inequality concludes the proof. ∎

5.2. Computable upper bound on $\|\mathcal{R}(s)\|_{X'}$. The dual norm $\|\mathcal{R}(s)\|_{X'}$ in the abstract error estimate (5.2) is not easily computable. We are now going to infer a computable upper bound for this quantity, introducing the flux reconstruction $\boldsymbol{\theta}$ and making use of assumptions (3.1), (3.2), and (3.4).

LEMMA 5.2 (computable upper bound on $\|\mathcal{R}(s)\|_{X'}$). *Assume* (3.1), (3.2), *and* (3.4). *Let* $\eta_{\mathrm{R},T}^n$ *and* $\eta_{\mathrm{DF},T}^n$ *be defined by* (3.5)–(3.6). *Then,*

$$\|\mathcal{R}(s)\|_{X'} \le \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{R},T}^n + \eta_{\mathrm{DF},T}^n(t))^2\,\mathrm{d}t\right\}^{1/2} + \|f-\widetilde{f}\|_{X'}.$$

Proof. Let $\varphi \in X$ with $\|\varphi\|_X = 1$. Then, adding and subtracting $(\boldsymbol{\theta}, \nabla\varphi)$ in the integrand for a.e. $t \in (0, t_{\mathrm{F}})$ and using the Green theorem yields

$$\begin{aligned}\langle \mathcal{R}(s), \varphi\rangle_{X',X} &= \int_0^{t_{\mathrm{F}}} \{(f-\partial_t s - \nabla\cdot\boldsymbol{\theta}, \varphi) - (\nabla s + \boldsymbol{\theta}, \nabla\varphi)\}(t)\,\mathrm{d}t\\ &= \int_0^{t_{\mathrm{F}}} \{(f-\widetilde{f}, \varphi) + (\widetilde{f} - \partial_t s - \nabla\cdot\boldsymbol{\theta}, \varphi) - (\nabla s + \boldsymbol{\theta}, \nabla\varphi)\}(t)\,\mathrm{d}t\\ &=: T_1 + T_2 + T_3.\end{aligned}$$

Clearly, $|T_1| \le \|f-\widetilde{f}\|_{X'}\|\varphi\|_X = \|f-\widetilde{f}\|_{X'}$. Moreover, owing to (3.1), there holds $s \in P_\tau^1(H_0^1(\Omega))$ and $\boldsymbol{\theta} \in P_\tau^0(\mathbf{H}(\mathrm{div},\Omega))$, so that

$$T_2 = \sum_{n=1}^N \int_{I_n} (\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, \varphi(t))\,\mathrm{d}t.$$

For all $1 \le n \le N$, owing to (3.3) (which is a consequence of (3.2)) and (3.4),

$$(\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, 1)_T = 0 \qquad \forall T \in \mathcal{T}^n.$$

Hence, for a.e. $t \in I_n$,

$$\begin{aligned}(\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, \varphi(t)) &= (\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, \varphi(t) - \Pi_0^n\varphi(t)) \\ &\le \sum_{T\in\mathcal{T}^n} C_{\mathrm{P}} h_T \|\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n\|_T \|\nabla\varphi\|_T(t),\end{aligned}$$

where we have used the Poincaré inequality (3.7) on each $T \in \mathcal{T}^n$. Moreover,

$$T_3 \le \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} \|\nabla s(t) + \boldsymbol{\theta}^n\|_T \|\nabla\varphi\|_T(t)\,\mathrm{d}t.$$

Collecting the above estimates yields using the Cauchy–Schwarz inequality

$$|T_2 + T_3| \le \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{R},T}^n + \eta_{\mathrm{DF},T}^n(t))^2\,\mathrm{d}t\right\}^{1/2}.$$

The conclusion is now straightforward. □

5.3. Computable upper bound on $\|s - u_{h\tau}\|_Y$. The next step is to derive a computable upper bound on $\|s - u_{h\tau}\|_Y$ since this quantity also involves a dual norm.

LEMMA 5.3 (computable upper bound on $\|s - u_{h\tau}\|_Y$). *Assume* (3.1) *and* (3.2). *Let* $\eta_{\mathrm{NC},1,T}^n$ *and* $\eta_{\mathrm{NC},2,T}^n$ *be defined by* (3.8)–(3.9). *Then,*

$$\|s - u_{h\tau}\|_Y \le \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC},1,T}^n)^2(t)\,\mathrm{d}t\right\}^{1/2} + \left\{\sum_{n=1}^N \tau^n \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC},2,T}^n)^2\right\}^{1/2}.$$

Proof. It is clear that

$$\|s - u_{h\tau}\|_X = \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC},1,T}^n)^2(t)\,\mathrm{d}t\right\}^{1/2}.$$

Now let $\varphi \in X$ with $\|\varphi\|_X = 1$. Observe that since both s and $u_{h\tau}$ are piecewise affine and continuous in time,

$$\langle\partial_t(s - u_{h\tau}), \varphi\rangle_{X',X} = \sum_{n=1}^N \int_{I_n} (\partial_t(s - u_{h\tau})^n, \varphi(t))\,\mathrm{d}t.$$

For all $1 \le n \le N$, it is inferred from (3.3) that the quantity $\partial_t(s - u_{h\tau})^n$ has zero mean on each element $T \in \mathcal{T}^n$. Hence, using the Poincaré inequality (3.7) yields

$$\begin{aligned}\langle\partial_t(s - u_{h\tau}), \varphi\rangle_{X',X} &= \sum_{n=1}^N \int_{I_n} (\partial_t(s - u_{h\tau})^n, \varphi(t) - \Pi_0^n\varphi(t))\,\mathrm{d}t \\ &\le \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} \eta_{\mathrm{NC},2,T}^n \|\nabla\varphi\|_T(t)\,\mathrm{d}t,\end{aligned}$$

whence the desired estimate is inferred using the Cauchy–Schwarz inequality. □

5.4. Proof of Theorem 3.2. We observe that using the definition (3.10) for η_{IC}, Lemma 5.1 yields $\|u-u_{h\tau}\|_Y \le \|s-u_{h\tau}\|_Y + 3\|\mathcal{R}(s)\|_{X'} + \eta_{\mathrm{IC}}$, and we use the bounds on $\|\mathcal{R}(s)\|_{X'}$ and $\|s-u_{h\tau}\|_Y$ derived, respectively, in Lemmas 5.2 and 5.3 to conclude the proof.

5.5. Proof of Theorem 3.6. Letting for all $1 \le n \le N$, $T \in \mathcal{T}^n$, and $t \in I_n$,

$$\eta^n_{\mathrm{DF},2,T}(t) := \|\nabla s(t) - \nabla s^n\|_T,$$

and recalling the definitions (3.6) and (3.13) of $\eta^n_{\mathrm{DF},T}(t)$ and $\eta^n_{\mathrm{DF},1,T}$, respectively, the triangle inequality yields $\eta^n_{\mathrm{DF},T}(t) \le \eta^n_{\mathrm{DF},1,T} + \eta^n_{\mathrm{DF},2,T}(t)$. A further triangle inequality with $A^2 := \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta^n_{\mathrm{R},T} + \eta^n_{\mathrm{DF},T}(t))^2 \,\mathrm{d}t$ leads to

$$\begin{aligned} A &\le \left\{ \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta^n_{\mathrm{R},T} + \eta^n_{\mathrm{DF},1,T} + \eta^n_{\mathrm{DF},2,T}(t))^2 \,\mathrm{d}t \right\}^{1/2} \\ &\le \left\{ \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta^n_{\mathrm{R},T} + \eta^n_{\mathrm{DF},1,T})^2 \,\mathrm{d}t \right\}^{1/2} + \left\{ \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} \eta^n_{\mathrm{DF},2,T}(t)^2 \,\mathrm{d}t \right\}^{1/2}. \end{aligned}$$

Both terms on the right-hand side can be readily integrated in time yielding

$$A \le \left\{ \sum_{n=1}^N \sum_{T\in\mathcal{T}^n} \tau^n (\eta^n_{\mathrm{R},T} + \eta^n_{\mathrm{DF},1,T})^2 \right\}^{1/2} + \left\{ \sum_{n=1}^N \sum_{T\in\mathcal{T}^n} \frac{1}{3}\tau^n \|\nabla(s^n - s^{n-1})\|_T^2 \right\}^{1/2}.$$

As a result, owing to Theorem 3.2,

$$\begin{aligned} \|u-u_{h\tau}\|_Y \le{}& 3 \left\{ \sum_{n=1}^N \sum_{T\in\mathcal{T}^n} \tau^n (\eta^n_{\mathrm{R},T} + \eta^n_{\mathrm{DF},1,T})^2 \right\}^{1/2} \\ &+ 3 \left\{ \sum_{n=1}^N \sum_{T\in\mathcal{T}^n} \frac{1}{3}\tau^n \|\nabla(s^n - s^{n-1})\|_T^2 \right\}^{1/2} + \eta_{\mathrm{IC}} + 3\|f-\widetilde{f}\|_{X'} \\ &+ \left\{ \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta^n_{\mathrm{NC},1,T})^2(t) \,\mathrm{d}t \right\}^{1/2} + \left\{ \sum_{n=1}^N \sum_{T\in\mathcal{T}^n} \tau^n (\eta^n_{\mathrm{NC},2,T})^2 \right\}^{1/2}, \end{aligned}$$

whence the conclusion follows by regrouping on the right-hand side the first, third, and fourth sums.

6. Proof of the error lower bound. The goal of this section is to prove Theorem 3.9. The proof is decomposed into several steps. Since the error lower bound is local-in-time, we keep the integer $1 \le n \le N$ fixed in this section. It is readily seen that

$$\eta^n_{\mathrm{sp}} + \eta^n_{\mathrm{tm}} \lesssim \eta^n_{\mathrm{R}} + \eta^n_{\mathrm{DF}} + \eta^n_{\mathrm{NC},1} + \eta^n_{\mathrm{NC},2},$$

where

$$(\eta_{\mathrm{R}}^n)^2 := \tau^n \sum_{T\in\mathcal{T}^n} h_T^2 \|\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n\|_T^2,$$

$$(\eta_{\mathrm{DF}}^n)^2 := \tau^n \|\nabla s^n + \boldsymbol{\theta}^n\|^2 + \int_{I_n} \|\nabla s(t) - \nabla s^n\|^2 \,\mathrm{d}t,$$

$$(\eta_{\mathrm{NC},1}^n)^2 := \int_{I_n} \|\nabla^{n-1,n}(s - u_{h\tau})\|^2(t)\,\mathrm{d}t,$$

$$(\eta_{\mathrm{NC},2}^n)^2 := \tau^n \sum_{T\in\mathcal{T}^n} h_T^2 \|\partial_t(s - u_{h\tau})^n\|_T^2,$$

and we bound the four quantities η_{R}^n, η_{DF}^n, $\eta_{\mathrm{NC},1}^n$, and $\eta_{\mathrm{NC},2}^n$.

6.1. Equivalent expression for $\boldsymbol{\eta_{\mathrm{NC},1}^n}$. It is convenient to simplify the expression for $\eta_{\mathrm{NC},1}^n$.

LEMMA 6.1 (equivalent expression for $\eta_{\mathrm{NC},1}^n$). *Letting $v := s - u_{h\tau}$, there holds*

$$\tau^n \tfrac{1}{6}(\|\nabla^{n-1}v^{n-1}\|^2 + \|\nabla^n v^n\|^2) \le (\eta_{\mathrm{NC},1}^n)^2 \le \tau^n \tfrac{1}{2}(\|\nabla^{n-1}v^{n-1}\|^2 + \|\nabla^n v^n\|^2).$$

Proof. Let $T \in \mathcal{T}^{n-1,n}$. Since v is piecewise affine in time and smooth in T, it is inferred that

$$\begin{aligned}\int_{I_n} \|\nabla v\|_T^2(t)\,\mathrm{d}t &= \tau^n \int_0^1 \|\nabla v^{n-1} + \tau\nabla(v^n - v^{n-1})\|_T^2\, d\tau \\ &= \tau^n \tfrac{1}{3}(\|\nabla v^{n-1}\|_T^2 + \|\nabla v^n\|_T^2 + (\nabla v^{n-1}, \nabla v^n)_T).\end{aligned}$$

Hence,

$$\tau^n \tfrac{1}{6}(\|\nabla v^{n-1}\|_T^2 + \|\nabla v^n\|_T^2) \le \int_{I_n} \|\nabla v\|_T^2(t)\,\mathrm{d}t \le \tau^n \tfrac{1}{2}(\|\nabla v^{n-1}\|_T^2 + \|\nabla v^n\|_T^2).$$

Summing over $T \in \mathcal{T}^{n-1,n}$ yields the assertion. ∎

6.2. Bounds on $\boldsymbol{\eta_{\mathrm{R}}^n}$ and $\boldsymbol{\eta_{\mathrm{DF}}^n}$. We introduce the following quantities:

$$(\mathcal{E}_{\mathrm{R}}^n)^2 := \tau^n \sum_{T\in\mathcal{T}^n} h_T^2 \|\Pi_{V_h^n}\widetilde{f}^n - \partial_t u_{h\tau}^n + \Delta u_{h\tau}^n\|_T^2, \tag{6.1}$$

$$(\mathcal{E}_{\mathrm{J}}^n)^2 := \tau^n \sum_{T\in\mathcal{T}^n} |\mathbf{n}\cdot[\![\nabla^n u_{h\tau}^n]\!]|^2_{+\frac{1}{2},\mathfrak{F}_T^{\mathrm{i},n}}, \tag{6.2}$$

and observe that $\mathcal{E}_{\mathrm{R}}^n$ and $\mathcal{E}_{\mathrm{J}}^n$ take the form of usual residual-based a posteriori error estimators for the heat equation; cf. [34, 43, 10]. We also recall that the quantity $\mathcal{E}_f^n$ is defined by (3.20) and $\mathcal{J}_*^n(u_{h\tau})$ by (3.18). Observe that (3.17) implies

$$\tau^n \sum_{T\in\mathcal{T}^n} \|\nabla u_{h\tau}^n + \boldsymbol{\theta}^n\|_T^2 \lesssim (\mathcal{E}_{\mathrm{R}}^n)^2 + (\mathcal{E}_{\mathrm{J}}^n)^2 + (\mathcal{E}_f^n)^2 + \mathcal{J}_*^n(u_{h\tau})^2. \tag{6.3}$$

LEMMA 6.2 (bounds on η_{R}^n and η_{DF}^n). *Assume* (M1) *and the approximation property* (3.17) *for the flux reconstruction $\boldsymbol{\theta}$. Then,*

$$\eta_{\mathrm{R}}^n \lesssim \mathcal{E}_{\mathrm{R}}^n + \mathcal{E}_{\mathrm{J}}^n + \mathcal{E}_f^n + \eta_{\mathrm{NC},2}^n + \mathcal{J}_*^n(u_{h\tau}), \tag{6.4}$$

$$\eta_{\mathrm{DF}}^n \lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}_{\mathrm{R}}^n + \mathcal{E}_{\mathrm{J}}^n + \mathcal{E}_f^n + \eta_{\mathrm{NC},1}^n + \eta_{\mathrm{NC},2}^n + \mathcal{J}_*^n(u_{h\tau}). \tag{6.5}$$

Proof. (i) Bound on η_{R}^n. Observe that

$$\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n = (\widetilde{f}^n - \Pi_{V_h^n}\widetilde{f}^n) - (\partial_t s^n - \partial_t u_{h\tau}^n) + (\Pi_{V_h^n}\widetilde{f}^n - \partial_t u_{h\tau}^n + \Delta u_{h\tau}^n) - (\Delta u_{h\tau}^n + \nabla\cdot\boldsymbol{\theta}^n).$$

Hence, using the triangle inequality and the above definitions leads to

$$(\eta_{\mathrm{R}}^n)^2 \lesssim (\mathcal{E}_{\mathrm{R}}^n)^2 + (\mathcal{E}_f^n)^2 + (\eta_{\mathrm{NC},2}^n)^2 + \tau^n \sum_{T\in\mathcal{T}^n} h_T^2 \|\Delta u_{h\tau}^n + \nabla\cdot\boldsymbol{\theta}^n\|_T^2.$$

We next use an inverse inequality to bound the last term:

$$\tau^n \sum_{T\in\mathcal{T}^n} h_T^2 \|\Delta u_{h\tau}^n + \nabla\cdot\boldsymbol{\theta}^n\|_T^2 \lesssim \tau^n \sum_{T\in\mathcal{T}^n} \|\nabla u_{h\tau}^n + \boldsymbol{\theta}^n\|_T^2.$$

The two above estimates and (6.3) imply (6.4).

(ii) Bound on η_{DF}^n. Define

$$(\eta_{\mathrm{DF}}^n)^2 = \int_{I_n} \|\nabla s(t) - \nabla s^n\|^2 \,\mathrm{d}t + \tau^n \|\nabla s^n + \boldsymbol{\theta}^n\|^2 =: \int_{I_n} A^n(t)\,\mathrm{d}t + \tau^n B^n.$$

Let v be the space-time function on I_n such that $v(t) := s(t) - s^n$, and observe that $v \in X(I_n)$. Elementary algebra yields

$$\begin{aligned} A^n(t) = (\nabla s(t) - \nabla s^n, \nabla v(t)) = & -\{(f,v)(t) - (\partial_t s^n, v(t)) - (\nabla s(t), \nabla v(t))\} \\ & + (f(t) - \widetilde{f}^n, v(t)) \\ & + (\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, v(t)) \\ & - (\nabla s^n + \boldsymbol{\theta}^n, \nabla v(t)) \\ =: & A_1^n(t) + A_2^n(t) + A_3^n(t) + A_4^n(t). \end{aligned}$$

Owing to (2.1), $\int_{I_n} A_1^n(t)\,\mathrm{d}t \le \|u-s\|_{Y(I_n)}\|v\|_{X(I_n)}$. Furthermore, $\int_{I_n} A_2^n(t)\,\mathrm{d}t \le \|f-\widetilde{f}\|_{X'(I_n)}\|v\|_{X(I_n)}$. Using as before the Poincaré inequality (3.7) for $A_3^n(t)$ leads to $\int_{I_n} A_3^n(t)\,\mathrm{d}t \le \eta_{\mathrm{R}}^n \|v\|_{X(I_n)}$. Concerning $A_4^n(t)$, the Cauchy–Schwarz inequality yields $\int_{I_n} A_4^n(t)\,\mathrm{d}t \le (\tau^n B^n)^{1/2}\|v\|_{X(I_n)}$. Collecting these bounds, since $\int_{I_n} A^n(t)\,\mathrm{d}t = \|v\|_{X(I_n)}^2$, leads to

$$\int_{I_n} A^n(t)\,\mathrm{d}t \lesssim \|u-s\|_{Y(I_n)}^2 + \|f-\widetilde{f}\|_{X'(I_n)}^2 + (\eta_{\mathrm{R}}^n)^2 + \tau^n B^n.$$

Using the triangle inequality for the first term on the right-hand side and bounding $\|s-u_{h\tau}\|_{Y(I_n)}$ as in Lemma 5.3 yields

$$\|u-s\|_{Y(I_n)}^2 \lesssim \|u-u_{h\tau}\|_{Y(I_n)}^2 + (\eta_{\mathrm{NC},1}^n)^2 + (\eta_{\mathrm{NC},2}^n)^2.$$

As a result, using the bound (6.4) for η_{R}^n leads to

$$\int_{I_n} A^n(t)\,\mathrm{d}t \lesssim \|u-u_{h\tau}\|_{Y(I_n)}^2 + (\mathcal{E}_{\mathrm{R}}^n)^2 + (\mathcal{E}_{\mathrm{J}}^n)^2 + (\mathcal{E}_f^n)^2 + (\eta_{\mathrm{NC},1}^n)^2 + (\eta_{\mathrm{NC},2}^n)^2 + \mathcal{J}_*^n(u_{h\tau})^2 + \tau^n B^n.$$

Thus,

$$(\eta_{\mathrm{DF}}^n)^2 \lesssim \|u-u_{h\tau}\|_{Y(I_n)}^2+(\mathcal{E}_{\mathrm{R}}^n)^2+(\mathcal{E}_{\mathrm{J}}^n)^2+(\mathcal{E}_f^n)^2+(\eta_{\mathrm{NC},1}^n)^2+(\eta_{\mathrm{NC},2}^n)^2+\mathcal{J}_*^n(u_{h\tau})^2+\tau^n B^n,$$

and it remains to bound the last term. To this purpose, we use the triangle inequality, the lower bound in Lemma 6.1, and (6.3) to infer that

$$\begin{aligned}\tau^n B^n &\lesssim \tau^n\{\|\nabla^n(s^n-u_{h\tau}^n)\|^2+\|\nabla^n u_{h\tau}^n+\boldsymbol{\theta}^n\|^2\}\\ &\lesssim (\eta_{\mathrm{NC},1}^n)^2+(\mathcal{E}_{\mathrm{J}}^n)^2+\mathcal{J}_*^n(u_{h\tau})^2+(\mathcal{E}_f^n)^2+(\mathcal{E}_{\mathrm{R}}^n)^2.\end{aligned}$$

The conclusion is straightforward. □

6.3. Bounds on $\eta_{\mathrm{NC},1}^n$ and $\eta_{\mathrm{NC},2}^n$. The next step is to bound $\eta_{\mathrm{NC},1}^n$ and $\eta_{\mathrm{NC},2}^n$.

LEMMA 6.3 (bound on $\eta_{\mathrm{NC},1}^n$). *Assume* (M1)–(M2) *and that the potential reconstruction s is defined by* (3.15)–(3.16). *Then,*

$$\eta_{\mathrm{NC},1}^n \lesssim \mathcal{J}^n(u_{h\tau}).$$

Proof. Owing to the upper bound in Lemma 6.1, it suffices to prove that for $m\in\{n-1,n\}$ and for all $T\in\mathcal{T}^m$,

$$\|\nabla(s^m-u_{h\tau}^m)\|_T \lesssim |[\![u_{h\tau}^m]\!]|_{-\frac12,\mathfrak{F}_T^m}.$$

The proof is presented for $m=n$; it is similar for $m=n-1$. Letting

$$\iota_{h\tau}^n := \mathcal{I}_{\mathrm{av}}^n(u_{h\tau}^n), \qquad \beta_{h\tau}^n := \sum_{T'\in\mathcal{T}^{n,n+1}} \alpha_{T'}^n b_{T'},$$

it is clear that (3.15) yields the decomposition $s^n-u_{h\tau}^n=(\iota_{h\tau}^n-u_{h\tau}^n)+\beta_{h\tau}^n$, and we bound, on each $T\in\mathcal{T}^n$, the two terms on the right-hand side separately. Classical approximation properties of the averaging interpolation operator yield

$$\begin{aligned}\|\nabla(\iota_{h\tau}^n-u_{h\tau}^n)\|_T &\lesssim |[\![u_{h\tau}^n]\!]|_{-\frac12,\mathfrak{F}_T^n},\\ \|\iota_{h\tau}^n-u_{h\tau}^n\|_T &\lesssim |[\![u_{h\tau}^n]\!]|_{+\frac12,\mathfrak{F}_T^n}.\end{aligned}$$

To estimate the bubble contribution, we observe that for all $T'\in\mathcal{T}^{n,n+1}$ such that $T'\subset T$, using (3.16) and the Cauchy–Schwarz inequality,

$$\|\nabla\beta_{h\tau}^n\|_{T'} = |\alpha_{T'}^n|\,\|\nabla b_{T'}\|_{T'} \le \left(h_{T'}\frac{\|\nabla b_{T'}\|_{T'}|T'|^{1/2}}{(b_{T'},1)_{T'}}\right) h_{T'}^{-1}\|\iota_{h\tau}^n-u_{h\tau}^n\|_{T'}.$$

Owing to the shape regularity of $\mathcal{T}^{n,n+1}$, the factor between parentheses is bounded uniformly, so that

$$\|\nabla\beta_{h\tau}^n\|_{T'} \lesssim h_{T'}^{-1}\|\iota_{h\tau}^n-u_{h\tau}^n\|_{T'}.$$

We now use the transition condition (3.14) (the restriction on refinement is invoked), the shape regularity of the meshes, and the approximation properties of the averaging interpolation operator to infer that

$$\begin{aligned}\|\nabla\beta_{h\tau}^n\|_T^2 &= \sum_{T'\in\mathcal{T}^{n,n+1};T'\subset T}\|\nabla\beta_{h\tau}^n\|_{T'}^2 \lesssim \sum_{T'\in\mathcal{T}^{n,n+1};T'\subset T} h_{T'}^{-2}\|\iota_{h\tau}^n-u_{h\tau}^n\|_{T'}^2\\ &\le \Xi^2 h_T^{-2}\sum_{T'\in\mathcal{T}^{n,n+1};T'\subset T}\|\iota_{h\tau}^n-u_{h\tau}^n\|_{T'}^2 = \Xi^2 h_T^{-2}\|\iota_{h\tau}^n-u_{h\tau}^n\|_T^2 \lesssim |[\![u_{h\tau}^n]\!]|_{-\frac12,\mathfrak{F}_T^n}^2.\end{aligned}$$

The proof is complete. □

LEMMA 6.4 (bound on $\eta^n_{\mathrm{NC},2}$). *Assume* (M1)–(M3) *and that the potential reconstruction* s *is defined by* (3.15)–(3.16). *Then,*

$$\eta^n_{\mathrm{NC},2} \lesssim \mathcal{J}^n(u_{h\tau}).$$

Proof. Let $v := s - u_{h\tau}$. Using the transition condition (3.14) and the triangle inequality yields

$$\sum_{T\in\mathcal{T}^n} h_T^2 \|\partial_t v^n\|_T^2 \lesssim \sum_{T\in\mathcal{T}^{n-1}} h_T^2 (\tau^n)^{-2} \|v^{n-1}\|_T^2 + \sum_{T\in\mathcal{T}^n} h_T^2 (\tau^n)^{-2} \|v^n\|_T^2.$$

Consider the second term on the right-hand side. Proceeding as in the proof of Lemma 6.3 (the restriction on refinement is needed here) leads to the bounds $\|v^n\|_T \lesssim |[\![u^n_{h\tau}]\!]|_{+\frac{1}{2},\mathfrak{F}^n_T}$, so that, owing to the shape regularity of the meshes and the condition (M3),

$$\sum_{T\in\mathcal{T}^n} h_T^2 (\tau^n)^{-2} \|v^n\|_T^2 \lesssim \sum_{T\in\mathcal{T}^n} h_T^4 (\tau^n)^{-2} |[\![u^n_{h\tau}]\!]|^2_{-\frac{1}{2},\mathfrak{F}^n_T} \le \Upsilon^2 \sum_{T\in\mathcal{T}^n} |[\![u^n_{h\tau}]\!]|^2_{-\frac{1}{2},\mathfrak{F}^n_T}.$$

The term at t^{n-1} is treated similarly using again the transition condition in combination with (M3). ∎

6.4. Bounds on $\mathcal{E}^n_{\mathrm{R}}$ and $\mathcal{E}^n_{\mathrm{J}}$. The last step consists of bounding the usual residual-based error estimators $\mathcal{E}^n_{\mathrm{R}}$ and $\mathcal{E}^n_{\mathrm{J}}$.

LEMMA 6.5 (bounds on $\mathcal{E}^n_{\mathrm{R}}$ and $\mathcal{E}^n_{\mathrm{J}}$). *Under the assumptions of Lemmas* 6.2 *and* 6.4, *there holds*

$$\mathcal{E}^n_{\mathrm{R}} + \mathcal{E}^n_{\mathrm{J}} \lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}^n_f + \mathcal{J}^n(u_{h\tau}).$$

Proof. Using the technique of element and edge bubble functions introduced by Verfürth [42] and proceeding as in [43], it can be shown (details are skipped for brevity) that

$$(\mathcal{E}^n_{\mathrm{R}})^2 + (\mathcal{E}^n_{\mathrm{J}})^2 \le C(\|u - u_{h\tau}\|^2_{Y(I_n)} + (\mathcal{E}^n_f)^2) + \epsilon^2 \int_{I_n} \|\nabla^{n-1,n}(u_{h\tau}(t) - u^n_{h\tau})\|^2 \,\mathrm{d}t,$$

where ϵ can be chosen to be as small as needed. The triangle inequality yields

$$\|\nabla^{n-1,n}(u_{h\tau}(t) - u^n_{h\tau})\| \le \|\nabla^{n-1,n}(s - u_{h\tau})(t)\| + \|\nabla s(t) + \boldsymbol{\theta}^n\| + \|\nabla^n u^n_{h\tau} + \boldsymbol{\theta}^n\|.$$

Hence, taking square roots and using (6.3),

$$\mathcal{E}^n_{\mathrm{R}} + \mathcal{E}^n_{\mathrm{J}} \le C(\|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}^n_f) + \epsilon(\eta^n_{\mathrm{DF}} + \eta^n_{\mathrm{NC},1} + \mathcal{E}^n_{\mathrm{J}} + \mathcal{E}^n_{\mathrm{R}} + \mathcal{E}^n_f + \mathcal{J}^n_*(u_{h\tau})).$$

Owing to Lemmas 6.2, 6.3, and 6.4,

$$\begin{aligned}\eta^n_{\mathrm{DF}} + \eta^n_{\mathrm{NC},1} &\lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}^n_{\mathrm{R}} + \mathcal{E}^n_{\mathrm{J}} + \mathcal{E}^n_f + \eta^n_{\mathrm{NC},1} + \eta^n_{\mathrm{NC},2} + \mathcal{J}^n_*(u_{h\tau}) \\ &\lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}^n_{\mathrm{R}} + \mathcal{E}^n_{\mathrm{J}} + \mathcal{E}^n_f + \mathcal{J}^n(u_{h\tau}).\end{aligned}$$

Choosing ϵ sufficiently small to eliminate the terms $\mathcal{E}^n_{\mathrm{R}}$ and $\mathcal{E}^n_{\mathrm{J}}$ from the right-hand side yields

$$\mathcal{E}^n_{\mathrm{R}} + \mathcal{E}^n_{\mathrm{J}} \le C(\|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}^n_f) + \epsilon \mathcal{J}^n(u_{h\tau}), \tag{6.6}$$

whence the conclusion is straightforward. ∎

6.5. Proof of Theorem 3.9. The proof is a direct consequence of Lemmas 6.2, 6.3, 6.4, and 6.5.

Appendix. Conforming and nonconforming finite elements.

This appendix extends the above theory to the conforming and nonconforming finite element methods.

A.1. Conforming finite elements. The conforming finite element method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n := \mathbb{P}_1(\mathcal{T}^n) \cap H_0^1(\Omega)$ such that

$$(\partial_t u_{h\tau}^n, v_h) + (\nabla u_{h\tau}^n, \nabla v_h) = (\widetilde{f}^n, v_h) \qquad \forall v_h \in V_h^n, \tag{A.1}$$

where $u_{h\tau}^0 := \Pi_{V_h^0}(u^0)$. The conforming finite element method given by (A.1) is very close to the vertex-centered finite volume method of section 4.4. In fact, as in [45], we find that (A.1) is equivalent to look for $u_{h\tau}^n \in V_h^n$ such that

$$\begin{aligned} &(\partial_t u_{h\tau}^n, 1)_D - (\nabla u_{h\tau}^n \cdot \mathbf{n}_D, 1)_{\partial D} - (\widetilde{f}^n, 1)_D \\ &= -(\partial_t u_{h\tau}^n - \Pi_0^n(\partial_t u_{h\tau}^n), \psi_V)_{\mathcal{T}_V} + (\partial_t u_{h\tau}^n - \Pi_0^n(\partial_t u_{h\tau}^n), 1)_D \\ &\quad + (\widetilde{f}^n - \Pi_0^n(\widetilde{f}^n), \psi_V)_{\mathcal{T}_V} - (\widetilde{f}^n - \Pi_0^n(\widetilde{f}^n), 1)_D \qquad \forall D \in \mathcal{D}^{\mathrm{i},n}. \end{aligned} \tag{A.2}$$

Here, V is always the vertex associated with the dual volume D, ψ_V is the associated "hat" basis function, and $\mathcal{T}_V$ is the patch of elements of $\mathcal{T}^n$ sharing V. Let $\boldsymbol{\theta}^n \in \mathbf{RTN}_0(\mathcal{S}^n)$ satisfy $\boldsymbol{\theta}^n \cdot \mathbf{n}_F|_F := -\nabla u_{h\tau}^n \cdot \mathbf{n}_F|_F$ on all faces F of $\mathcal{S}^n$ included in ∂D for some $D \in \mathcal{D}^{\mathrm{i},n}$. Then, contrarily to section 4.4, (3.4) is not satisfied on $D \in \mathcal{D}^{\mathrm{i},n}$; indeed, the right-hand side of (A.2), clearly equivalent to a numerical quadrature, has to be taken into account. In the proof of Lemma 5.2, the bound on the term T_2 has to be enriched by

$$(\partial_t u_{h\tau}^n - \widetilde{f}^n - \Pi_0^n(\partial_t u_{h\tau}^n - \widetilde{f}^n), \tilde{\varphi} - \bar{\varphi}),$$

where $\tilde{\varphi}$ is a piecewise linear, vertex-based interpolation of φ on the mesh $\mathcal{T}^n$ and $\bar{\varphi}$ is a piecewise constant one on the mesh $\mathcal{D}^n$. Consequently, one estimator has to be added in Theorem 3.2 to $\eta_{\mathrm{R},T}^n$, namely $C_{\mathrm{I},T} h_T \|\partial_t u_{h\tau}^n - \widetilde{f}^n - \Pi_0^n(\partial_t u_{h\tau}^n - \widetilde{f}^n)\|_T$; here $C_{\mathrm{I},T}$ is a constant arising from the interpolation of $\varphi \in H_0^1(\Omega)$ by $\tilde{\varphi}$ and $\bar{\varphi}$. To work on the mesh $\mathcal{T}^n$, the flux reconstruction is determined by solving local linear systems, similarly to section 4.4. The approximation property (3.17) can be shown using the techniques of [45]. To prove a lower bound for the new estimator is immediate, since $\Delta u_{h\tau}^n = 0$, so that this estimator equals $C_{\mathrm{I},T} h_T \|\partial_t u_{h\tau}^n - \widetilde{f}^n - \Delta u_{h\tau}^n - \Pi_0^n(\partial_t u_{h\tau}^n - \widetilde{f}^n - \Delta u_{h\tau}^n)\|_T$.

A.2. Nonconforming finite elements. The nonconforming finite element method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n$ such that

$$(\partial_t u_{h\tau}^n, v_h) + (\nabla^n u_{h\tau}^n, \nabla^n v_h) = (\widetilde{f}^n, v_h) \qquad \forall v_h \in V_h^n. \tag{A.3}$$

Here V_h^n is defined as in section 4.5 and $u_{h\tau}^0 := \Pi_{V_h^0}(u^0)$. The nonconforming finite element method given by (A.3) is, in fact, very close to the face-centered finite volume method of section 4.5. Our estimates can be adapted to the present setting by proceeding similarly to section A.1.

REFERENCES

[1] Y. Achdou, C. Bernardi, and F. Coquel, *A priori and a posteriori analysis of finite volume discretizations of Darcy's equations*, Numer. Math., 96 (2003), pp. 17–42.

[2] M. Ainsworth, *A synthesis of a posteriori error estimation techniques for conforming, nonconforming and discontinuous Galerkin finite element methods*, in Recent Advances in Adaptive Computation, Contemp. Math. 383, AMS, Providence, RI, 2005, pp. 1–14.

[3] M. Ainsworth, *A posteriori error estimation for discontinuous Galerkin finite element approximation*, SIAM J. Numer. Anal., 45 (2007), pp. 1777–1798.

[4] M. Amara, L. Nadau, and D. Trujillo, *A posteriori error estimator for finite volume schemes*, in VIII Journées Zaragoza-Pau de Mathématiques Appliquées et de Statistiques, Monogr. Semin. Mat. García Galdeano 31, Prensas Univ. Zaragoza, Zaragoza, 2004, pp. 21–30.

[5] T. Arbogast and Z. Chen, *On the implementation of mixed methods as nonconforming methods for second-order elliptic problems*, Math. Comp., 64 (1995), pp. 943–972.

[6] D. N. Arnold and F. Brezzi, *Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 7–32.

[7] I. Babuška, M. Feistauer, and P. Šolín, *On one approach to a posteriori error estimates for evolution problems solved by the method of lines*, Numer. Math., 89 (2001), pp. 225–256.

[8] I. Babuška and S. Ohnimus, *A posteriori error estimation for the semidiscrete finite element method of parabolic differential equations*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 4691–4712.

[9] M. Bebendorf, *A note on the Poincaré inequality for convex domains*, Z. Anal. Anwendungen, 22 (2003), pp. 751–756.

[10] A. Bergam, C. Bernardi, and Z. Mghazli, *A posteriori analysis of the finite element discretization of some parabolic equations*, Math. Comp., 74 (2005), pp. 1117–1138.

[11] D. Braess, V. Pillwein, and J. Schöberl, *Equilibrated residual error estimates are p-robust*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 1189–1197.

[12] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer-Verlag, New York, 1991.

[13] E. Burman and A. Ern, *Continuous interior penalty hp-finite element methods for advection and advection-diffusion equations*, Math. Comp., 76 (2007), pp. 1119–1140.

[14] J. M. Cascón, L. Ferragut, and M. I. Asensio, *Space-time adaptive algorithm for the mixed parabolic problem*, Numer. Math., 103 (2006), pp. 367–392.

[15] A. Ern, S. Nicaise, and M. Vohralík, *An accurate* $\mathbf{H}(\mathrm{div})$ *flux reconstruction for discontinuous Galerkin approximations of elliptic problems*, C. R. Math. Acad. Sci. Paris, 345 (2007), pp. 709–712.

[16] A. Ern, A. F. Stephansen, and M. Vohralík, *Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection–diffusion–reaction problems*, J. Comput. Appl. Math., 234 (2010), pp. 114–130.

[17] A. Ern and M. Vohralík, *Flux reconstruction and a posteriori error estimation for discontinuous Galerkin methods on general nonmatching grids*, C. R. Math. Acad. Sci. Paris, 347 (2009), pp. 441–444.

[18] L. C. Evans, *Partial Differential Equations*, Grad. Stud. Math. 19, AMS, Providence, RI, 1998.

[19] R. Eymard, T. Gallouët, and R. Herbin, *Finite Volume Methods*, in Handbook of Numerical Analysis, Vol. VII, North–Holland, Amsterdam, 2000, pp. 713–1020.

[20] J. de Frutos, B. García-Archilla, and J. Novo, *A posteriori error estimates for fully discrete nonlinear parabolic problems*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 3462–3474.

[21] E. H. Georgoulis and O. Lakkis, *A Posteriori Error Control for Discontinuous Galerkin Methods for Parabolic Problems*, Technical report, The University of Sussex, Sussex, England, available online at http://arxiv.org/abs/0804.4362, 2008.

[22] D. Hilhorst and M. Vohralík, *A Posteriori Error Estimates for Combined Finite Volume–Finite Element Discretizations of Reactive Transport Equations on Nonmatching Grids*, Preprint R10007, Laboratoire Jacques-Louis Lions, Paris, France; HAL Preprint 00453952, submitted.

[23] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek, *Solution of Variational Inequalities in Mechanics*, Appl. Math. Sci. 66, Springer-Verlag, New York, 1988. Translated from the Slovak by J. Jarník.

[24] O. A. Karakashian and F. Pascal, *A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003),

pp. 2374–2399.
[25] O. A. Karakashian and F. Pascal, *Convergence of adaptive discontinuous Galerkin approximations of second-order elliptic problems*, SIAM J. Numer. Anal., 45 (2007), pp. 641–665.
[26] K. Y. Kim, *A posteriori error estimators for locally conservative methods of nonlinear elliptic problems*, Appl. Numer. Math., 57 (2007), pp. 1065–1080.
[27] P. Ladevèze, *Comparaison de modèles de milieux continus*, Ph.D. thesis, Université Pierre et Marie Curie (Paris 6), Paris, France, 1975.
[28] O. Lakkis and C. Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp., 75 (2006), pp. 1627–1658.
[29] R. Luce and B. I. Wohlmuth, *A local a posteriori error estimator based on equilibrated fluxes*, SIAM J. Numer. Anal., 42 (2004), pp. 1394–1414.
[30] C. Makridakis and R. H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 1585–1594.
[31] P. Morin, R. H. Nochetto, and K. G. Siebert, *Convergence of adaptive finite element methods*, SIAM Rev., 44 (2002), pp. 631–658.
[32] S. Nicaise and N. Soualem, *A posteriori error estimates for a nonconforming finite element discretization of the heat equation*, M2AN Math. Model. Numer. Anal., 39 (2005), pp. 319–348.
[33] L. E. Payne and H. F. Weinberger, *An optimal Poincaré inequality for convex domains*, Arch. Rational Mech. Anal., 5 (1960), pp. 286–292.
[34] M. Picasso, *Adaptive finite elements for a linear parabolic problem*, Comput. Methods Appl. Mech. Engrg., 167 (1998), pp. 223–237.
[35] W. Prager and J. L. Synge, *Approximations in elasticity based on the concept of function space*, Quart. Appl. Math., 5 (1947), pp. 241–269.
[36] S. Repin, *Estimates of deviations from exact solutions of initial-boundary value problem for the heat equation*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl., 13 (2002), pp. 121–133.
[37] S. I. Repin, *A posteriori error estimation for nonlinear variational problems by duality theory*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), 243 (1997), pp. 201–214, 342.
[38] A. Schmidt and K. G. Siebert, *ALBERT—software for scientific computations and applications*, Acta Math. Univ. Comenian. (N.S.), 70 (2000), pp. 105–122.
[39] R. Stevenson, *The completion of locally refined simplicial partitions created by bisection*, Math. Comp., 77 (2008), pp. 227–241.
[40] T. Strouboulis, I. Babuška, and D. K. Datta, *Guaranteed a posteriori error estimation for fully discrete solutions of parabolic problems*, Internat. J. Numer. Methods Engrg., 56 (2003), pp. 1243–1259.
[41] J. L. Synge, *The Hypercircle in Mathematical Physics: A Method for the Approximate Solution of Boundary Value Problems*, Cambridge University Press, New York, 1957.
[42] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Teubner-Wiley, Stuttgart, Germany, 1996.
[43] R. Verfürth, *A posteriori error estimates for finite element discretizations of the heat equation*, Calcolo, 40 (2003), pp. 195–212.
[44] M. Vohralík, *A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations*, SIAM J. Numer. Anal., 45 (2007), pp. 1570–1599.
[45] M. Vohralík, *Guaranteed and Fully Robust A Posteriori Error Estimates for Conforming Discretizations of Diffusion Problems with Discontinuous Coefficients*, Preprint R08009, Laboratoire Jacques-Louis Lions, Paris, France; HAL Preprint 00235810, submitted.
[46] M. Vohralík, *A posteriori error estimation in the conforming finite element method based on its local conservativity and using local minimization*, C. R. Math. Acad. Sci. Paris, 346 (2008), pp. 687–690.
[47] M. Vohralík, *Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods*, Numer. Math., 111 (2008), pp. 121–158.
[48] M. Vohralík, *Unified primal formulation-based a priori and a posteriori error analysis of mixed finite element methods*, Math. Comp., accepted.