SIAM J. NUMER. ANAL.
Vol. 50, No. 6, pp. 2845–2872

A POSTERIORI ERROR CONTROL FOR FULLY DISCRETE CRANK–NICOLSON SCHEMES*

E. BÄNSCH†, F. KARAKATSANI‡, AND CH. MAKRIDAKIS§

Abstract. We derive residual-based a posteriori error estimates of optimal order for fully discrete approximations for linear parabolic problems. The time discretization uses the Crank–Nicolson method, and the space discretization uses finite element spaces that are allowed to change in time. The main tool in our analysis is the comparison with an appropriate reconstruction of the discrete solution, which is introduced in the present paper.

Key words. a posteriori error estimators, Crank–Nicolson method, parabolic problem

AMS subject classification. 65N15

DOI. 10.1137/110839424

1. Introduction.

Adaptive algorithms, for instance, for time-dependent problems, are linked to error control through a posteriori estimates. Despite the effort given to these problems in recent years, many key issues remain unexplored. One of them is the effect of mesh modification on the estimation and algorithm design in fully discrete finite element approximations for simple model problems. Mesh modification is a necessity in the design of adaptive schemes with moving interesting characteristics. In the present paper, we derive the first optimal order a posteriori estimates in $L^\infty(L^2)$ for fully discrete Crank–Nicolson (CN) schemes for linear parabolic problems allowing mesh modification. In a companion paper [4] we study the qualitative analytical and computational behavior of the schemes and the estimators.

Our analysis leads to the interesting conclusion that even simple *refinement can spoil CN schemes.* As a consequence we introduce a version of the CN scheme consistent with mesh redistribution. The definition of the new version of the fully discrete CN scheme is motivated by the a posteriori analysis presented herein. Although the behavior of the scheme is more robust during mesh change, it appears that the estimators might become sensitive in certain cases. In [4] we present detailed computational experiments which show that the a posteriori estimators are of optimal order and include terms capturing the spatial and the temporal errors separately.

We claim that the observed behavior during mesh modification is related to the well-known fact that CN is sensitive to nonsmooth data effects. It is known that CN is a sensitive scheme and belongs to the border of stable time discretization methods for diffusion problems. Among its known properties is its lack of smoothing effect; see [14, 19]. Spectral arguments and computational results in [4] suggest that this lack of

*Received by the editors July 5, 2011; accepted for publication (in revised form) July 19, 2012; published electronically November 6, 2012.
http://www.siam.org/journals/sinum/50-6/83942.html

†Applied Mathematics III, University of Erlangen, D-91058 Erlangen, Germany (baensch@am.uni-erlangen.de).

‡BCAM - Basque Center for Applied Mathematics, E-48009 Bilbao, Basque Country, Spain (fkarakatsani@bcamath.org). This author's research was supported by the Chair of Applied Mathematics III of the University of Erlangen through a Research Assistantship from 2007–2010.

§Department of Applied Mathematics, University of Crete, 71409 Heraklion-Crete, Greece, and Institute of Applied and Computational Mathematics, FORTH, 71110 Heraklion-Crete, Greece (makr@tem.uoc.gr). This author's research was partially supported by the FP7-REGPOT project Archimedes Center for Modeling, Analysis and Computations.

smoothing of the CN scheme is present and influences the behavior of the a posteriori estimators. Therefore, it will be interesting to study nonstandard projections into the new finite element space with improved smoothing properties compared to the standard interpolant. Our analysis is designed to allow such a possibility, but this is a subtle issue which depends heavily on the qualitative properties of the PDE at hand and requires further investigation.

Our analysis is based on the methodology developed in [16, 12] for space discrete and fully discrete schemes and in [2, 3] for time discrete schemes. A key point is the definition of an auxiliary function which we call *reconstruction* of the approximation U. For the problem considered, the choice of *reconstruction* is not obvious. As far as the time reconstruction is concerned, we follow the approach of [3], which includes the reconstructions based on approximations on one time level (two-point estimators) as in [2] as well as the reconstructions based on approximations on two time levels (three-point estimators) as in [13]. The *elliptic reconstruction* [16] is also involved in the analysis leading to the derivation of optimal order estimators in $L^\infty(L^2)$. The mesh change effect is also taken into account.

Fully discrete a posteriori estimates for CN time discretization methods were derived previously in [13, 21]. The estimators in [13] are valid only without mesh change, and they are of optimal order in $L^2(H^1)$ but not in $L^\infty(L^2)$. The estimators in [21] are not second order in time; see [17]. Compared to existing results, apart from including the possibility of mesh change, our analysis provides optimal order estimators in $L^\infty(L^2)$ for higher order in time fully discrete schemes. We would also like to emphasize the fact that our approach does not hinge in an essential way on the parabolic character of the problem and thus is in principle applicable to other evolution problems. In fact, in [11, 10], the ideas presented herein were useful to a posteriori analysis for the linear Schrödinger equation and for generalized KdV-type equations.

A posteriori bounds for CN methods applied to the linear Schrödinger equation were derived by Dörfler [5]. Alternative estimators for the discretization methods and the problem at hand based on the direct comparison of u and the numerical solution U could be derived using parabolic duality as in [9, 6, 8, 7].

The rest of this paper is organized as follows. In section 2 we introduce the problem setting and the notation, and we discuss the standard and the modified CN discretizations allowing mesh change. In section 3 we present the space-time reconstruction. Specific choices of the reconstructions leading to two-point and three-point estimators are given. Section 4 is devoted to the main error analysis. In section 5 we state the final estimates in the case of two as well as in the case of three-point estimators. Further, each term appearing in the main theorem of section 4 is analyzed completing the proof of the estimates. For completeness we include the main conclusions of [4] regarding the qualitative and computational behavior of the estimators derived in this paper in section 6.

2. CN methods. Let Ω be a convex polygonal domain in $\mathbb{R}^d$, $d = 2, 3$, and let $T > 0$. Our assumption on the domain is due to the requirement that elliptic regularity estimates hold. If this is not the case, the terms affected can still be treated by using more involved elliptic estimators; see [12]. We consider the following heat equation: Find $u \in L^\infty(0,T;H_0^1(\Omega))$, with $\partial_t u \in L^2(0,T;L^2(\Omega))$, satisfying

$$\begin{cases} \langle u_t, \phi\rangle + a(u,\phi) = \langle f, \phi\rangle \quad \forall \phi \in H_0^1(\Omega) \ \text{ on } (0,T], \\ u(0) = u^0, \end{cases} \tag{2.1}$$

where $f \in L^2(0,T;L^2(\Omega))$ and $u^0 \in H_0^1(\Omega)$. We denote by $\langle\cdot,\cdot\rangle$ either the inner product in $L^2(\Omega)$ or the duality pairing between $H_0^1(\Omega)$ and its dual $H^{-1}(\Omega)$, and we denote by $a(\cdot,\cdot)$ the bilinear form in $H_0^1(\Omega)$ defined as

$$a(v,w) = \langle \nabla v, \nabla w\rangle \quad \forall v,w \in H_0^1(\Omega). \tag{2.2}$$

For $\mathcal{D} \subset \mathbb{R}^d$ we denote by $\|\cdot\|_{\mathcal{D}}$ the norm in $L^2(\mathcal{D})$, and we denote by $\|\cdot\|_{r,\mathcal{D}}$ and $|\cdot|_{r,\mathcal{D}}$ the norm and the seminorm, respectively, in the Sobolev space $H^r(\mathcal{D}), r \in \mathbb{Z}^+$. In view of the Poincaré inequality, we consider $|\cdot|_{1,\mathcal{D}}$ to be the norm in $H_0^1(\mathcal{D})$ and denote by $|\cdot|_{-1,\mathcal{D}}$ the norm in $H^{-1}(\mathcal{D})$. Further, to simplify the notation, we shall omit the subscript $\mathcal{D}$ in the notation of function spaces and norms whenever $\mathcal{D} = \Omega$ throughout the rest of this paper.

Let $0 = t^0 < t^1 < \cdots < t^N = T$ be a partition of $[0,T]$, and let $I_n := (t^{n-1}, t^n]$, $k_n := t^n - t^{n-1}$, $t^{n-\frac{1}{2}} := t^{n-1} + k_n/2$. In what follows, we shall also denote by $u^m(x)$ and $f^m(x)$ the values $u(x,t^m)$ and $f(x,t^m)$. In particular, if v is piecewise linear in time, obviously, $v^{n-\frac{1}{2}} = \frac{v^{n-1}+v^n}{2}$, $n = 1,\dots,N$.

For the space discretization we shall use the finite element method; therefore we introduce a family $\{\mathcal{T}_n\}_{n=0}^N$ of *conforming shape-regular triangulations* of the domain Ω corresponding to the time node t^n. Here we assume that the triangulations may be changed in time. Let h_n be the *local mesh-size function* of each given triangulation $\mathcal{T}_n$ defined by

$$h_n(x) := h_K, \quad K \in \mathcal{T}_n \text{ and } x \in K, \tag{2.3}$$

with $h_K := \operatorname{diam}(K)$. For each n and for each $K \in \mathcal{T}_n$, we let $\mathcal{E}_n(K)$ be the set of the sides of K (edges in $d = 2$ or faces in $d = 3$) and $\Sigma_n(K) \subset \mathcal{E}_n(K)$ be the set of the internal sides of K, i.e., the sides which do not belong to the boundary of Ω. Given two successive triangulations $\mathcal{T}_{n-1}$ and $\mathcal{T}_n$, we define the *finest common coarsening* $\hat{\mathcal{T}}^n := \mathcal{T}^{n-1} \wedge \mathcal{T}^n$, whose local mesh sizes are given by $\hat{h}_n := \max(h_n, h_{n-1})$, respectively; see [12] for precise definitions. Notice that essentially $\hat{\mathcal{T}}^n$ is the triangulation of $\mathbb{V}^n \cap \mathbb{V}^{n-1}$. In addition, we introduce the sets $\mathcal{E}_n := \cup_{K\in\mathcal{T}_n}\mathcal{E}_n(K)$ and $\Sigma_n := \cup_{K\in\mathcal{T}_n}\Sigma_n(K)$. We shall also use the sets $\hat{\Sigma}_n := \Sigma_n \cap \Sigma_{n-1}$ and $\check{\Sigma}_n := \Sigma_n \cup \Sigma_{n-1}$. Finally, by $[v]_e$ and $[\mathbf{v}]_e$ we denote the jump of the possibly discontinuous scalar and vector valued function, respectively. The global function $[v]$ defined on Σ_n is just $[v]|_e = [v]_e$.

We associate with each triangulation $\mathcal{T}_n$ the finite element spaces

$$\tilde{\mathbb{V}}_h^n := \{\phi \in H^1(\Omega) : \forall K \in \mathcal{T}_n : \phi|_K \in \mathbb{P}_l(K)\} \quad \text{and} \quad \mathbb{V}_h^n := \tilde{\mathbb{V}}_h^n \cap H_0^1(\Omega), \tag{2.4}$$

where $\mathbb{P}_l(Q)$ is the space of polynomials of degree at most l on Q.

2.1. The fully discrete scheme. The fully discrete schemes obtained by the CN method in time and by the standard finite element method in space satisfy the following discrete equations: let U^0 be a given initial approximation of u^0, and for $1 \le n \le N$, find $U^n \in \mathbb{V}_h^n$ such that

$$\left\langle \frac{U^n - U^{n-1}}{k_n}, \phi_n \right\rangle + a\left(\frac{U^n + U^{n-1}}{2}, \phi_n\right) = \langle f^{n-\frac{1}{2}}, \phi_n\rangle \quad \forall \phi_n \in \mathbb{V}_h^n. \tag{2.5}$$

This scheme can be written in the pointwise form

$$\frac{U^n - P_0^n U^{n-1}}{k_n} + \frac{1}{2}(-\Delta_h^n)U^n + \frac{1}{2}(-\Delta_h^n)U^{n-1} = P_0^n f^{n-\frac{1}{2}}. \tag{2.6}$$

Here $P_0^n : L^2 \to \mathbb{V}_h^n$ is the L^2-projection onto $\mathbb{V}_h^n$, and Δ_h^n is the discrete Laplacian corresponding to the finite element space $\mathbb{V}_h^n$ defined by the following.

DEFINITION 2.1. *The discrete Laplacian $\Delta_h^n : H_0^1(\Omega) \to \mathbb{V}_h^n$ is the operator with the property*

$$\langle -\Delta_h^n v, \phi_n \rangle = a(v, \phi_n) \quad \forall \phi_n \in \mathbb{V}_h^n. \tag{2.7}$$

It turns out that scheme (2.6) might lead to inaccurate approximations; this is due to the term $\Delta_h^n U^{n-1}$, which may cause problems during a mesh modification. Indeed, if, for instance, $\mathcal{T}_n$ is a refinement of $\mathcal{T}_{n-1}$, then the discrete Laplace operator on the *finer* mesh is applied to coarse grid functions leading to oscillatory behavior of the term $\Delta_h^n U^{n-1}$; see section 6.1 for details. To support this claim we present a simple numerical example in one dimension: There, the standard CN scheme (2.6) was applied for 20 time steps with global refinement each 6 time steps. Then the computed solution at the final time has a clear oscillatory behavior; see Figure 1 and section 6.1. Notice that this is particularly interesting since usually errors are not expected during refinement only. From a theoretical point of view, the analysis of the standard scheme (2.6) leads to an additional error term of the form

$$\|(\Delta_h^{n-1} - \Delta_h^n) U^{n-1}\|;$$

see section 3.2 for more details and an alternative estimate. This term is clearly not zero under pure refinement, and in fact it can be seen as the term reflecting the oscillatory behavior from the theoretical viewpoint. The above numerical experiment shows that the presence of this term is not an artifact of the theory.

The above discussion motivated us to introduce the modified CN scheme: For n, $1 \le n \le N$, find $U^n \in \mathbb{V}_h^n$, $1 \le n \le N$, such that

$$\frac{U^n - \widetilde{\Pi}^n U^{n-1}}{k_n} + \frac{1}{2}\Pi^n(-\Delta_h^{n-1})U^{n-1} + \frac{1}{2}(-\Delta_h^n)U^n = f_h^{n-\frac{1}{2}}. \tag{2.8}$$

Here Π^n, $\widetilde{\Pi}^n : \mathbb{V}_h^{n-1} \to \mathbb{V}_h^n$ denote suitable projections or interpolants to be chosen. In addition we use the notation $f_h(t)|_{I_n} = P_0^n f(t)$. The reason for introducing a further operator $\widetilde{\Pi}^n$ is that we would like to study schemes and corresponding estimators including several possible choices for the projection step. It seems that (2.8) indeed resolves the oscillatory behavior of the classical scheme, as the same computational test indicates; see Figure 1 and section 6.1 for further experiments. In addition, as it will become evident by the forthcoming analysis, the only difference in the a posteriori estimation of (2.6) and (2.8) is that the estimators for the first scheme include the additional term corresponding to $(\Delta_h^{n-1} - \Delta_h^n)U^{n-1}$; see section 3.2. Let us note at this point that we do not claim that it is always preferable to use (2.8) when mesh modification occurs. Its implementation requires the solution of one more system with the mass matrix, and thus, depending on the specific application, it might be possible to appropriately tune the mesh parameters in (2.6) yielding approximations free from artifacts. In this case the use of estimators for (2.6) derived herein might prove to be particularly useful.

The scheme (2.8) is in fact natural: One may think that the fact that the discrete Laplacian changes with time introduces an *"artificial time dependence"* of the form

$$y_t + A(t)y = 0\,,$$

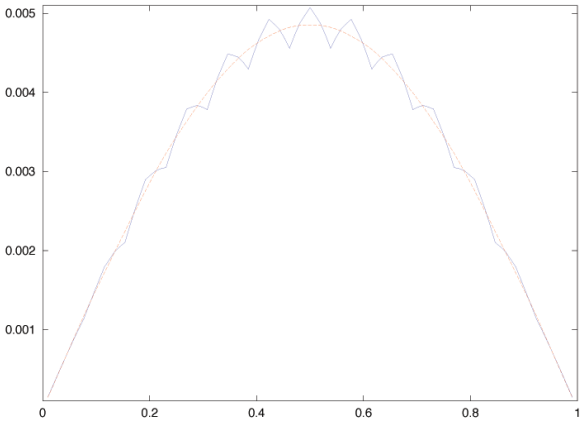


FIG. 1. *Comparison of the standard CN scheme* (2.6) *(blue dotted line) and the modified method* (2.8) *(red dotted line); solutions after* 20 *time steps with global refinement each* 6 *time steps of the homogeneous heat equation with* $u^0(x) = x(1-x),\ x \in [0,1]$, $T = 0.4$, *constant time steps* $k_n = 0.02$, *and initial spatial step* $h_0 = 0.025$.

say, to the space discrete ODE. An application of the trapezoidal method to this problem will yield

$$\frac{y^n - y^{n-1}}{k_n} + \frac{1}{2}A(t^{n-1})y^{n-1} + \frac{1}{2}A(t^n)y^n = 0,$$

and the similarity to (2.8) is evident.

2.2. CN–Galerkin schemes. Next we discuss the relation of the CN scheme to the CN–Galerkin method, i.e., the lowest order *continuous Galerkin* time discrete method. We shall need some more notation. Let $\mathcal{V}_q(I_n; \mathbb{V}_h^n)$ be the space of polynomial functions in time of degree q, i.e., of functions $g : I_n \to \mathbb{V}_h^n$ of the form $g|_{I_n}(t) = \sum_{j=0}^q t^j w_j,\ w_j \in \mathbb{V}_h^n$. Then, the CN–Galerkin approximation to u is defined as follows: We seek $\tilde{U}$, $\tilde{U}|_{I_n}, \in \mathcal{V}_1(I_n; \mathbb{V}_h^n)$ such that

(2.9)
$$\int_{I_n} [\langle \tilde{U}', v\rangle + a(\tilde{U}, v)]\, dt = \int_{I_n} \langle f, v\rangle \quad \forall v \in \mathcal{V}_0(I_n; \mathbb{V}_h^n) \text{ and } \tilde{U}^{n-1+} = \Pi_G^n \tilde{U}^{n-1}$$

for $n = 1, \ldots, N$. Here $\Pi_G^n : \mathbb{V}_h^{n-1} \to \mathbb{V}_h^n$ denote suitable projections at our disposal. Notice that when no mesh change is performed the function $\tilde{U}$ is globally continuous and the method is reduced to the standard form of the continuous Galerkin method, cG(1). It is easily seen, using the exactness of the midpoint rule, that the CN–Galerkin scheme is equivalent to

$$\frac{\tilde{U}^n - \Pi_G^n \tilde{U}^{n-1}}{k_n} + \frac{1}{2}(-\Delta_h^n)\tilde{U}^n + \frac{1}{2}(-\Delta_h^n)\Pi_G^n \tilde{U}^{n-1} = \frac{1}{k_n}\int_{I_n} f_h(t)dt. \tag{2.10}$$

In the case where Π_G^n is the elliptic projection operator on $\mathbb{V}_h^n$, then $(-\Delta_h^n)\Pi_G^n \tilde{U}^{n-1} = (-\Delta_h^n)\tilde{U}^{n-1}$. As in the case of the CN scheme one can modify (2.10) in various ways, which all reduce to the standard scheme in the case of absence of mesh modification with n. In all cases, schemes of the form (2.10) differ from the CN schemes in the way the nonhomogeneous term f is approximated. The analysis presented herein can be applied to CN–Galerkin schemes with appropriate adaptations; see Remark 3.1.

2.3. Projections. It will be useful to collect the notation of all projections/interpolants used in this paper. We start with $P_0^n : L^2 \to \mathbb{V}_h^n$, which is the standard projection onto $\mathbb{V}_h^n$. P_0^n is used in the standard scheme (2.6). The analysis in this paper is carried for the modified scheme (2.8), where Π^n, $\widetilde{\Pi}^n : \mathbb{V}_h^{n-1} \to \mathbb{V}_h^n$ denote suitable projections which are at our disposal. Notice that $\widetilde{\Pi}^n$ is always multiplying Δ_h^{n-1} and our analysis does not require any specific assumptions on the choice of these projections. In the definition of the second reconstruction leading to the three-point estimator in section 3.2, we use another generic projection π^n onto $\mathbb{V}_h^n$. Finally, by $\mathit{\Pi}_G^n : \mathbb{V}_h^{n-1} \to \mathbb{V}_h^n$ we denote a projection used to project $\tilde{U}^{n-1+} = \mathit{\Pi}_G^n \tilde{U}^{n-1}$ in the definition of the continuous Galerkin method in section 2.2. Further, we shall use the standard Clément-type interpolant on $\mathbb{V}_h^n$, $\mathcal{I}^n : H_0^1 \to \mathbb{V}_h^n$ and the Clément-type interpolant on $\mathbb{V}_h^n \cap \mathbb{V}_h^{n-1}$, $\hat{\mathcal{I}}^n : H_0^1 \to \mathbb{V}_h^n \cap \mathbb{V}_h^{n-1}$.

3. Space-time reconstructions. The a posteriori error analysis is based on an appropriate *reconstruction* $\hat{U}$ of the piecewise linear in time interpolant of $\{U^n\}$ denoted by U; this idea was developed in [16, 12, 2, 3] and was summarized in [15]. The *total error* $e := u - U$ may be split into $e = \hat{\rho} + \varepsilon$, where $\hat{\rho}$, ε denote the *parabolic error* and the *reconstruction error* defined by $\hat{\rho} := u - \hat{U}$, $\varepsilon := \hat{U} - U$, respectively. A key ingredient of this approach is the fact that $\hat{U}$ should satisfy the same PDE with the exact solution but be perturbed with an a posteriori term which we would like to have in the final estimate (terms which are not computable but can be bounded a posteriori are also allowed).

As it was observed first in the time discrete case in [2], although U is a second order approximation of u in time, its residual is of suboptimal order; thus, energy techniques cannot lead to optimal order estimators; see also [21, 13, 17]. As far as the time reconstruction is concerned, we follow the approach of [3], which includes the reconstructions based on approximations on one time level (two-point estimators) as in [2] as well as the reconstructions based on approximations on two time levels (three-point estimators) as in [13]. To derive estimators of optimal order in $L^\infty(0,T;L^2(\Omega))$, we have to appropriately define $\hat{U}$ by involving in its derivation the *elliptic reconstruction* operator [16].

To fix notation, let $U : [0,T] \to H_0^1(\Omega)$ be the piecewise linear approximation of u defined by linearly interpolating between the nodal values U^{n-1} and U^n:

$$U(t) := l_0^n(t)U^{n-1} + l_1^n(t)U^n, \quad t \in I_n, \tag{3.1}$$

with

$$l_0^n(t) := \frac{t^n - t}{k_n} \quad \text{and} \quad l_1^n(t) := \frac{t - t^{n-1}}{k_n}\,. \tag{3.2}$$

We also introduce $\Theta : [0,T] \to H_0^1(\Omega)$, defined as

$$\Theta(t) = l_0^n(t)\Pi^n(-\Delta_h^{n-1})U^{n-1} + l_1^n(t)(-\Delta_h^n)U^n, \quad t \in I_n. \tag{3.3}$$

As we mentioned, the definition of the space-time reconstruction of the fully discrete approximate solution U^n, $n = 0, \dots, N$, defined in (2.8) involves the elliptic reconstruction operator $\mathcal{R}^n$ [16]. Since the definition of the elliptic reconstruction depends on the finite element space $\mathbb{V}_h^n$, the operator $\mathcal{R}^n$ also changes with n.

DEFINITION 3.1 (elliptic reconstruction). *For fixed $v_n \in \mathbb{V}_h^n$, we define the elliptic reconstruction $\mathcal{R}^n v_n \in H_0^1$ of v_n as the solution of the following variational problem:*

$$a(\mathcal{R}^n v_n, \psi) = \langle (-\Delta_h^n)v_n, \psi \rangle \quad \forall \psi \in H_0^1. \tag{3.4}$$

It can easily be seen that the elliptic reconstruction $\mathcal{R}^n$ satisfies the Galerkin orthogonality property

$$a(\mathcal{R}^n v_n - v_n, \chi_n) = 0 \quad \forall \chi_n \in \mathbb{V}_h^n. \tag{3.5}$$

Before we proceed further with the definition of the space-time reconstruction, we notice that the scheme (2.8) may be written in the following compact form:

$$\frac{U^n - \widetilde{\Pi}^n U^{n-1}}{k_n} + F^{n-\frac{1}{2}} = 0, \tag{3.6}$$

where

$$F^{n-\frac{1}{2}} = \frac{1}{2}\Pi^n(-\Delta_h^{n-1})U^{n-1} + \frac{1}{2}(-\Delta_h^n)U^n - P_0^n f^{n-\frac{1}{2}}. \tag{3.7}$$

Definition 3.2 (space-time reconstruction). *We define the space-time reconstruction, $\hat{U} : [0,T] \to H_0^1$, as follows:*

$$\begin{aligned}\hat{U}(t) := \mathcal{R}^{n-1}U^{n-1} + \frac{\mathcal{R}^n\widetilde{\Pi}^n U^{n-1} - \mathcal{R}^{n-1}U^{n-1}}{k_n}(t - t^{n-1}) \\ - \int_{t^{n-1}}^t \mathcal{R}^n \hat{F}(s)\, ds, \quad t \in I_n;\end{aligned} \tag{3.8}$$

here the function $\hat{F}(\cdot)$ is defined such that $\hat{F}(\cdot)|_{I_n}$ is a linear polynomial in time interpolating $F^{n-\frac{1}{2}}$, namely

$$\hat{F}(t^{n-\frac{1}{2}}) = F^{n-\frac{1}{2}}. \tag{3.9}$$

It is easily seen that the function $\hat{U}$ satisfies the following relation:

$$\hat{U}_t(t) + \mathcal{R}^n\hat{F}(t) = \frac{\mathcal{R}^n\widetilde{\Pi}^n U^{n-1} - \mathcal{R}^{n-1}U^{n-1}}{k_n}, \quad t \in I_n. \tag{3.10}$$

In addition, we observe that $\hat{U}$ interpolates the values $\mathcal{R}^{n-1}U^{n-1}$ and $\mathcal{R}^n U^n$. The first claim can easily be seen. Moreover, applying first the midpoint rule to evaluate the integral in (3.8) and then recalling (2.8), we obtain

$$\begin{aligned}\hat{U}(t^n) &= \mathcal{R}^n\widetilde{\Pi}^n U^{n-1} - \int_{t^{n-1}}^{t^n} \mathcal{R}^n\hat{F}(s)\, ds \\ &= \mathcal{R}^n\{\widetilde{\Pi}^n U^{n-1} - k_n F(t^{n-\frac{1}{2}})\} = \mathcal{R}^n U^n.\end{aligned} \tag{3.11}$$

The following result will also be useful in the next sections.

Lemma 3.3 (the difference $\hat{U} - \omega$). *Let $\omega : [0,T] \to H_0^1$ be the piecewise linear in time function defined by linearly interpolating between the values $\mathcal{R}^{n-1}U^{n-1}$ and $\mathcal{R}^n U^n$,*

$$\omega(t) := l_0^n(t)\mathcal{R}^{n-1}U^{n-1} + l_1^n(t)\mathcal{R}^n U^n, \quad t \in I_n, \tag{3.12}$$

with l_0^n and l_1^n defined in (3.2). *Then there holds that*

$$\hat{U}(t) - \omega(t) = -\int_{t^{n-1}}^t \mathcal{R}^n\{\hat{F}(s) - F(t^{n-\frac{1}{2}})\}\, ds. \tag{3.13}$$

Proof. According to (3.12) and (3.10), we have

$$\begin{aligned}\hat{U}_t(t)-\omega_t(t) &= \frac{\mathcal{R}^n\widetilde{\Pi}^n U^{n-1}-\mathcal{R}^{n-1}U^{n-1}}{k_n}-\mathcal{R}^n\hat{F}(t)-\frac{\mathcal{R}^nU^n-\mathcal{R}^{n-1}U^{n-1}}{k_n}\\ &= \mathcal{R}^n\left\{\frac{\widetilde{\Pi}^nU^{n-1}-U^n}{k_n}-\hat{F}(t)\right\}\\ &= \mathcal{R}^n\{F(t^{n-\frac{1}{2}})-\hat{F}(t)\}. \qquad \square\end{aligned} \tag{3.14}$$

Next we specify two choices of $\hat{F}$ leading to two different reconstructions and thus to different estimators. The first one leads to two-point estimators [2] and the second one to three-point estimators [13]. In addition we provide explicit formulas for the quantities

$$F(t^{n-\frac{1}{2}})-\hat{F}(t) \quad \text{and} \quad \hat{F}(t)-\Theta(t),$$

which appear in several terms in the general a posteriori estimate of Theorem 4.3.

3.1. Reconstruction 1: Two-point estimator. It is easily seen that the CN method (2.8) can be written as follows:

$$\frac{U^n-\widetilde{\Pi}^nU^{n-1}}{k_n}+\Theta(t^{n-\frac{1}{2}})=P_0^nf(t^{n-\frac{1}{2}}). \tag{3.15}$$

Let $F(t):=\Theta(t)-P_0^nf(t)$, and let $\hat{I}$ be the piecewise linear interpolant chosen as

$$\hat{I}(v)|_{I_n}\in\mathbb{P}_1(I_n), \quad \hat{I}(v)(t^{n-\frac{1}{2}})=v(t^{n-\frac{1}{2}}), \quad \hat{I}(v)(t^{n-1})=v(t^{n-1}). \tag{3.16}$$

Then

$$\hat{F}(t)=\hat{I}(\Theta(t)-P_0^nf(t))=\Theta(t)-P_0^n\varphi(t), \tag{3.17}$$

where $\varphi(t)=\hat{I}(f(t))$. Moreover, there holds that

$$\begin{aligned}\hat{F}(t^{n-\frac{1}{2}})) &= \hat{I}(\Theta(t^{n-\frac{1}{2}})-P_0^nf(t^{n-\frac{1}{2}}))\\ &= F^{n-\frac{1}{2}}.\end{aligned} \tag{3.18}$$

In preparation of the a posteriori estimators we shall now calculate the terms on the right-hand side in Theorem 4.3 below, depending on the above special choice of F and $\hat{I}$.

LEMMA 3.4 (calculation of $\hat{F}(t)-F^{n-1/2}$). *We have*

$$\hat{F}(t)-F^{n-1/2}=2\,(t-t^{n-1/2})w_n, \tag{3.19}$$

where w_n is given by

$$w_n:=\frac{1}{2}\partial_t\Theta(t)-\frac{P_0^n[f(t^{n-\frac{1}{2}})-f(t^{n-1})]}{k_n}. \tag{3.20}$$

Proof. In view of (3.17), we have

$$\hat{F}(t)-F^{n-1/2}=\Theta(t)-\Theta(t^{n-1/2})-P_0^n[\varphi(t)-\varphi(t^{n-1/2})]. \tag{3.21}$$

Now, in view of (3.3), it is easily seen that

$$\begin{aligned}\Theta(t)-\Theta(t^{n-1/2}) &= l_0^n(t)\Pi^n(-\Delta_h^{n-1})U^{n-1}+l_1^n(t)(-\Delta_h^n)U^n\\ &\quad-\frac{1}{2}\left(\Pi^n(-\Delta_h^{n-1})U^{n-1}+(-\Delta_h^n)U^n\right)\\ &= \frac{1}{2}(l_1^n(t)-l_0^n(t))\Big((-\Delta_h^n)U^n-\Pi^n(-\Delta_h^{n-1})U^{n-1}\Big)\\ &= (t-t^{n-1/2})\,\frac{1}{k_n}\Big((-\Delta_h^n)U^n-\Pi^n(-\Delta_h^{n-1})U^{n-1}\Big).\end{aligned}$$

The result claimed follows by combining the last two relations with the definition of φ. ∎

Furthermore, we have

$$\hat{F}(t)-\Theta(t)=-P_0^n\varphi(t)=-P_0^n\hat{I}(f(t)). \tag{3.22}$$

Remark 3.1 (reconstruction for the CN–Galerkin). Estimators for the time-discrete CN–Galerkin method were derived in [2, 3]. Here we consider the modified scheme

$$\frac{U^n-\widetilde{\Pi}^nU^{n-1}}{k_n}+\frac{1}{2}(-\Delta_h^n)U^n+\frac{1}{2}\Pi^{n-1}(-\Delta_h^{n-1})U^{n-1}=\frac{1}{k_n}\int_{I_n}f_h(t)dt. \tag{3.23}$$

It can be written in the compact form (3.6) with

$$F^{n-\frac{1}{2}}=\frac{1}{2}\Pi^n(-\Delta_h^{n-1})U^{n-1}+\frac{1}{2}(-\Delta_h^n)U^n-P_0^n\frac{1}{k_n}\int_{I_n}f(t)dt. \tag{3.24}$$

Then we define $\hat{F}$ through (3.17) but with $\varphi(t)=P_{1,t}f(t)$, with $P_{1,t}f|_{I_n}\in\mathbb{P}_1(I_n)$ the $L^2(I_n)$-projection onto $\mathbb{P}_1(I_n)$. Then the key property (3.9) still holds. In fact,

$$P_{1,t}f(t^{n-1/2})=\frac{1}{k_n}\int_{I_n}P_{1,t}f(s)ds=\frac{1}{k_n}\int_{I_n}f(s)ds.$$

Then, (3.19) is valid but with w_n given by

$$w_n:=\frac{1}{2}\partial_t\Theta(t)-\frac{6}{k_n^3}\int_{I_n}P_0^n\,f(s)(s-t^{n-1/2})\,ds. \tag{3.25}$$

3.2. Reconstruction 2: Three-point estimator. Our scheme (2.8) can be rewritten in the form

$$\frac{U^n-\widetilde{\Pi}^nU^{n-1}}{k_n}+F^{n-\frac{1}{2}}=0. \tag{3.26}$$

We shall also need the projected version of the same equation at the previous interval:

$$\pi^n\frac{U^{n-1}-\widetilde{\Pi}^{n-1}U^{n-2}}{k_{n-1}}+\pi^nF^{n-\frac{3}{2}}=0. \tag{3.27}$$

Here π^n is any projection to $\mathbb{V}_h^n$ at our disposal and

$$F^{n-\frac{3}{2}}=\frac{1}{2}(-\Delta_h^{n-1})U^{n-1}+\frac{1}{2}\Pi^{n-1}(-\Delta_h^{n-2})U^{n-2}-P_0^{n-1}f^{n-\frac{3}{2}}.$$

Then we define the extended piecewise linear interpolant $\hat{F}$ as

$$\hat{F}(t) := l^n_{1/2}(t)\, F^{n-\frac{1}{2}} + l^n_{-1/2}(t)\, \pi^n F^{n-\frac{3}{2}}, \quad t \in I_n, \tag{3.28}$$

where

$$l^n_{1/2}(t) := \frac{2\,(t - t^{n-\frac{3}{2}})}{k_n + k_{n-1}}, \quad l^n_{-1/2}(t) := \frac{2\,(t^{n-\frac{1}{2}} - t)}{k_n + k_{n-1}}. \tag{3.29}$$

Obviously, $\hat{F}(t) \in \mathbb{V}^n_h$ for each $t \in I_n$, $\hat{F}|_{I_n}$ is a linear function of t, and

$$\hat{F}(t^{n-\frac{1}{2}}) = F^{n-\frac{1}{2}}. \tag{3.30}$$

LEMMA 3.5 (calculation of $\hat{F}(t) - F^{n-1/2}$). *We have*

$$\hat{F}(t) - F^{n-1/2} = 2\,(t - t^{n-1/2})\tilde{w}_n, \tag{3.31}$$

where $\tilde{w}_n$ is given by

$$\tilde{w}_n := \frac{1}{k_n + k_{n-1}} \left[\left(\frac{U^n - \widetilde{\Pi}^n U^{n-1}}{k_n} \right) - \pi^n \left(\frac{U^{n-1} - \widetilde{\Pi}^{n-1} U^{n-2}}{k_{n-1}} \right) \right]. \tag{3.32}$$

Proof. We have

$$\begin{aligned} \hat{F}(t) - F^{n-1/2} &= l^n_{1/2}(t) F^{n-1/2} + l^n_{-1/2} \pi^n\, F^{n-3/2} - F^{n-1/2} \\ &= l^n_{-1/2}(t)(\pi^n F^{n-3/2} - F^{n-1/2}) = \\ &= l^n_{-1/2}(t) \left\{ \pi^n \left(\frac{U^{n-1} - \widetilde{\Pi}^{n-1} U^{n-2}}{k_{n-1}} \right) - \left(\frac{U^n - \widetilde{\Pi}^n U^{n-1}}{k_n} \right) \right\}. \end{aligned} \tag{3.33}$$

∎

LEMMA 3.6 (calculation of $\hat{F}(t) - \Theta(t)$). *If we denote*

$$\hat{\varphi}(t) := l^n_{1/2}(t)\, P^n_0 f^{n-\frac{1}{2}} + l^n_{-1/2}(t)\, \pi^n\, P^{n-1}_0 f^{n-\frac{3}{2}}, \quad t \in I_n, \tag{3.34}$$

we have

$$\begin{aligned} \hat{F}(t) - \Theta(t) = -\hat{\varphi}(t) &+ l^n_{-1/2}(t)\, \frac{1}{2} \left[\frac{k_{n-1}}{k_n} (-\Delta^n_h) U^n \right. \\ &\left. - \left(2\Pi^n + \frac{k_{n-1}}{k_n} \Pi^n - \pi^n \right) (-\Delta^{n-1}_h) U^{n-1} + \pi^n\, \Pi^{n-1} (-\Delta^{n-2}_h) U^{n-2} \right]. \end{aligned} \tag{3.35}$$

Proof. Note that the part of $\hat{F}$ involving f is indeed

$$\hat{\varphi}(t) = l^n_{1/2}(t)\, P^n_0 f^{n-\frac{1}{2}} + l^n_{-1/2}(t)\, \pi^n\, P^{n-1}_0 f^{n-\frac{3}{2}}, \quad t \in I_n. \tag{3.36}$$

Also,

$$\begin{aligned} \hat{F}(t^{n-\frac{1}{2}}) - \Theta(t^{n-1/2}) &= \frac{1}{2}(-\Delta^n_h) U^n + \frac{1}{2} \Pi^n (-\Delta^{n-1}_h) U^{n-1} - P^n_0 f^{n-\frac{1}{2}} - \Theta(t^{n-1/2}) \\ &= -P^n_0 f^{n-\frac{1}{2}}. \end{aligned} \tag{3.37}$$

In addition, we express Θ in I_n in terms of $l^n_{1/2}$ and $l^n_{-1/2}$:

$$\Theta(t) := l^n_{1/2}(t)\,\Theta(t^{n-\frac{1}{2}}) + l^n_{-1/2}(t)\,\widetilde{\Theta}^{n-\frac{3}{2}}, \quad t \in I_n, \tag{3.38}$$

where (see (3.3))

$$\widetilde{\Theta}^{n-\frac{3}{2}} := l^n_0(t^{n-\frac{3}{2}})\,\Pi^n(-\Delta_h^{n-1})U^{n-1} + l^n_1(t^{n-\frac{3}{2}})\,(-\Delta_h^n)U^n.$$

Taking into account the above, we conclude that

$$\begin{aligned}\hat{F}(t) - \Theta(t) &= -\hat{\varphi}(t)\\ &\quad + l^n_{-1/2}(t)\left[\pi^n\,\frac{1}{2}(-\Delta_h^{n-1})U^{n-1} + \pi^n\,\frac{1}{2}\Pi^{n-1}(-\Delta_h^{n-2})U^{n-2} - \widetilde{\Theta}^{n-\frac{3}{2}}\right].\end{aligned} \tag{3.39}$$

The desired result follows upon noticing that

$$l^n_0(t^{n-\frac{3}{2}}) = 1 + \frac{k_{n-1}}{2\,k_n}, \quad l^n_1(t^{n-\frac{3}{2}}) = -\frac{k_{n-1}}{2\,k_n}. \tag{3.40}$$

□

Remark 3.2. If we choose $\pi^n = \Pi^n$, the expression $\hat{F}(t) - \Theta(t)$ becomes

$$\begin{aligned}\hat{F}(t) - \Theta(t) &= -\hat{\varphi}(t)\\ &\quad + l^n_{-1/2}(t)\,\frac{k_{n-1}}{2}\left[\frac{(-\Delta_h^n)U^n - \Pi^n(-\Delta_h^{n-1})U^{n-1}}{k_n}\right.\\ &\quad \left. - \frac{\Pi^n(-\Delta_h^{n-1})U^{n-1} - \Pi^n\,\Pi^{n-1}(-\Delta_h^{n-2})U^{n-2}}{k_{n-1}}\right].\end{aligned} \tag{3.41}$$

Thus this difference gives rise to discrete second derivatives in time of the projected discrete Laplacian of U. Note that this corresponds formally to a $k^2(-\Delta)u_{tt}$ term.

4. Error analysis. In addition to the errors $\hat{\rho} = u - \hat{U}$ and $\varepsilon = \hat{U} - U$ previously introduced, we let ρ be the *parabolic error* defined by

$$\rho := u - \omega. \tag{4.1}$$

Furthermore, the reconstruction error ε may be split into $\varepsilon = \epsilon + \sigma$, where ϵ is the *elliptic reconstruction error* defined by $\epsilon := \omega - U$ and σ is the *time reconstruction error* defined by $\sigma := \hat{U} - \omega$. Hence, the error e can be split further as follows:

$$e = \hat{\rho} + \sigma + \epsilon. \tag{4.2}$$

The proof of the estimate relies on two main ingredients:

(a) the direct estimation of ε via the estimate of σ and ϵ, and

(b) the estimate of $\hat{\rho}$ using PDE stability estimates.

Note that σ will account for the time discretization error and ϵ for the space discretization error.

4.1. Estimates for the scheme (2.8). A crucial step in the proof is to establish the equation that $\hat{\rho}$ satisfies. The result is stated in the following two lemmas.

LEMMA 4.1. *For each $\psi \in H^1_0$, we have*

$$\begin{aligned}\langle \hat{\rho}_t(t), \psi\rangle + a(\rho(t), \psi) &= \langle \mathcal{R}^n \hat{F}(t) - \Theta(t), \psi\rangle + l^n_0(t)\langle(\Pi^n - I)(-\Delta_h^{n-1})U^{n-1}, \psi\rangle\\ &\quad - k_n^{-1}\langle \mathcal{R}^n\widetilde{\Pi}^n U^{n-1} - \mathcal{R}^{n-1}U^{n-1}, \psi\rangle + \langle f(t), \psi\rangle, \quad t \in I_n.\end{aligned} \tag{4.3}$$

Proof. For each $n = 1, \dots, N$ and for each $\psi \in H_0^1$, in view of (3.10), we obtain

$$\langle \hat{U}_t(t), \psi \rangle + \langle \mathcal{R}^n \hat{F}(t), \psi \rangle = k_n^{-1} \langle \mathcal{R}^n \widetilde{\Pi}^n U^{n-1} - \mathcal{R}^{n-1} U^{n-1}, \psi \rangle. \tag{4.4}$$

Subtracting the above relation from (2.1) and using Definition 3.2, we conclude that

$$\begin{aligned} \langle \hat{\rho}_t(t), \psi \rangle + a(\rho(t), \psi) = &-a(l_0^n(t)\mathcal{R}^{n-1}U^{n-1} + l_1^n(t)\mathcal{R}^n U^n, \psi) + \langle \mathcal{R}^n \hat{F}(t), \psi \rangle \\ &- k_n^{-1} \langle \mathcal{R}^n \widetilde{\Pi}^n U^{n-1} - \mathcal{R}^{n-1} U^{n-1}, \psi \rangle + \langle f(t), \psi \rangle. \end{aligned}$$

According to the definition of the elliptic reconstruction (3.4), the last relation leads to

$$\begin{aligned} \langle \hat{\rho}_t(t), \psi \rangle + a(\rho(t), \psi) = &-\langle l_0^n(t)(-\Delta_h^{n-1})U^{n-1} + l_1^n(t)(-\Delta_h^n)U^n, \psi \rangle + \langle \mathcal{R}^n \hat{F}(t), \psi \rangle \\ &- k_n^{-1} \langle \mathcal{R}^n \widetilde{\Pi}^n U^{n-1} - \mathcal{R}^{n-1} U^{n-1}, \psi \rangle + \langle f(t), \psi \rangle, \end{aligned} \tag{4.5}$$

from which, in view of (3.3), the result claimed follows. ∎

LEMMA 4.2 (rearrangement of the mesh-change terms). *With the notation of Lemma* 4.1*, for each* $t \in I_n$*, the error equation* (4.3) *can be rewritten in the form*

$$\langle \hat{\rho}_t(t), \psi \rangle + a(\rho(t), \psi) = \langle R_h, \psi \rangle \quad \forall \psi \in H_0^1, \tag{4.6}$$

where

$$\begin{aligned} R_h := &(\mathcal{R}^n - I)(\hat{F}(t) - F^{n-1/2}) + (\hat{F}(t) - \Theta(t)) - \frac{\widetilde{\Pi}^n U^{n-1} - U^{n-1}}{k_n} \\ &+ l_0^n(t)(\Pi^n - I)(-\Delta_h^{n-1})U^{n-1} - \frac{(\mathcal{R}^n - I)U^n - (\mathcal{R}^{n-1} - I)U^{n-1}}{k_n} + f(t). \end{aligned} \tag{4.7}$$

Proof. According to (3.6), the claim follows from (4.3) by observing that

$$\begin{aligned} &\langle (\mathcal{R}^n - I)F^{n-1/2}), \psi \rangle - k_n^{-1} \langle \mathcal{R}^n \widetilde{\Pi}^n U^{n-1} - \mathcal{R}^{n-1} U^{n-1}, \psi \rangle \\ &\qquad = -k_n^{-1} \langle \mathcal{R}^n U^n - \mathcal{R}^{n-1} U^{n-1}), \psi \rangle + k_n^{-1} \langle U^n - \widetilde{\Pi}^n U^{n-1}, \psi \rangle \end{aligned} \tag{4.8}$$

for each $\psi \in H_0^1$. ∎

We can thus conclude that $\hat{\rho}$ satisfies a parabolic equation whose right-hand side can be controlled a posteriori. Indeed, (4.6) yields

$$\langle \hat{\rho}_t(t), \psi \rangle + a(\hat{\rho}(t), \psi) = \langle R_h, \psi \rangle + a(\sigma(t), \psi) \quad \forall \psi \in H_0^1 \tag{4.9}$$

since $\hat{\rho}(t) - \rho(t) = -(\hat{U} - \omega) = -\sigma$. The terms on the right-hand side of (4.9) either are direct a posteriori terms or involve spatial error operators of the form $\mathcal{R}^j - I$.

We now prove the following general error bound. The concrete estimators accounting for the specific choices of reconstructions as well as the choice of the elliptic a posteriori estimators will be given in the next section.

THEOREM 4.3 (estimate in $L^\infty(L^2)$ and $L^2(H^1)$ for the parabolic error). *Let* u *be the exact solution of* (2.1)*, and let* ω *and* $\hat{U}$ *be defined in* (3.12) *and* (3.8)*, respectively. The following estimate holds:*

$$\max_{t \in [0, t^m]} \left\{ \|\hat{\rho}(t)\|^2 + \int_0^t (|\hat{\rho}(s)|_1^2 + |\rho(s)|_1^2)\, ds \right\} \le \|\hat{\rho}(0)\|^2 + \mathcal{J}_m, \tag{4.10}$$

where $\mathcal{J}_m$, $m = 1, \dots, N$, *are defined by*

$$\mathcal{J}_m := \sum_{n=1}^{m} (\mathcal{J}_n^T + 2\mathcal{J}_n^{S,1} + 2\mathcal{J}_n^{S,2} + 2\mathcal{J}_n^C + 2\mathcal{J}_n^D), \tag{4.11}$$

with

$$\mathcal{J}_n^T := \int_{t^{n-1}}^{t^n} |\sigma(s)|_1^2 ds, \tag{4.12}$$

$$\mathcal{J}_n^{S,1} := \int_{t^{n-1}}^{t^n} |\langle (\mathcal{R}^n - I)(\hat{F}(t) - F^{n-1/2}), \hat{\rho}(s)\rangle| ds, \tag{4.13}$$

$$\mathcal{J}_n^{S,2} := \int_{t^{n-1}}^{t^n} \left| \left\langle \frac{(\mathcal{R}^n - I)U^n - (\mathcal{R}^{n-1} - I)U^{n-1}}{k_n}, \hat{\rho}(s) \right\rangle \right| ds, \tag{4.14}$$

$$\mathcal{J}_n^C := \int_{t^{n-1}}^{t^n} \left| \left\langle (\Pi^n - I) l_0^n(s)(-\Delta_h^{n-1})U^{n-1} - (\widetilde{\Pi}^n - I)\frac{U^{n-1}}{k_n}, \hat{\rho}(s) \right\rangle \right| ds, \tag{4.15}$$

$$\mathcal{J}_n^D := \int_{t^{n-1}}^{t^n} |\langle \hat{F}(s) - \Theta(s) + f(s), \hat{\rho}(s)\rangle| \, ds. \tag{4.16}$$

Proof. Setting $\psi = \hat{\rho}$,

$$\frac{1}{2}\frac{d}{dt}\|\hat{\rho}(t)\|^2 + a(\rho(t), \hat{\rho}(t)) = \langle R_h, \hat{\rho}(t)\rangle \,. \tag{4.17}$$

In view of (4.7), it can easily be seen that

$$\int_0^{t^m} \langle R_h, \hat{\rho}(s)\rangle \, ds \le \sum_{n=1}^{m} (\mathcal{J}_n^{S,1} + \mathcal{J}_n^{S,2} + \mathcal{J}_n^C + \mathcal{J}_n^D). \tag{4.18}$$

Recalling that

$$a(\rho(t), \hat{\rho}(t)) = \frac{1}{2}|\rho(t)|_1^2 + \frac{1}{2}|\hat{\rho}(t)|_1^2 - \frac{1}{2}|\hat{\rho}(t) - \rho(t)|_1^2,$$

we conclude that

$$\|\hat{\rho}(t)\|^2 + \int_0^t (|\rho(s)|^2 + |\hat{\rho}(s)|^2) \, ds \le \|\hat{\rho}(0)\|^2 + J_m \quad \forall t \in [0, t^m], \tag{4.19}$$

which completes the proof. ∎

4.2. Estimates for the standard CN scheme. The same analysis provides estimates for the case of the standard CN scheme with mesh modification (2.6). In fact, keeping the same notation as before, instead of Lemma 4.1 we would have

(4.20)
$$\begin{aligned}\langle \hat{\rho}_t(t), \psi\rangle + a(\rho(t), \psi) &= \langle \mathcal{R}^n \hat{F}(t) - \Theta(t), \psi\rangle + l_0^n(t)\langle (\Pi^n - I)(-\Delta_h^{n-1})U^{n-1}, \psi\rangle \\ &\quad - k_n^{-1}\langle \mathcal{R}^n \widetilde{\Pi}^n U^{n-1} - \mathcal{R}^{n-1}U^{n-1}, \psi\rangle + \langle f(t), \psi\rangle \\ &\quad - l_0^n(t)\langle (\Delta_h^n - \Delta_h^{n-1})U^{n-1}, \psi\rangle, \quad t \in I_n.\end{aligned}$$

Therefore, the main error equation replacing (4.6) is

$$\langle \hat{\rho}_t(t), \psi\rangle + a(\rho(t), \psi) = \langle \widetilde{R}_h, \psi\rangle \quad \forall \psi \in H_0^1, \tag{4.21}$$

where R_h is defined in (4.7) and

$$\widetilde{R}_h := R_h - l_0^n(t)\,(\Delta_h^n - \Delta_h^{n-1})U^{n-1}\,. \tag{4.22}$$

Thus, by modifying the right-hand side appropriately, we can conclude that Theorem 4.3 is also valid in this case. In particular there holds that

$$\max_{t\in[0,t^m]} \left\{ \|\hat{\rho}(t)\|^2 + \int_0^t (|\hat{\rho}(s)|_1^2 + |\rho(s)|_1^2)\, ds \right\} \le \|\hat{\rho}(0)\|^2 + \widetilde{\mathcal{J}}_m, \tag{4.23}$$

where $\widetilde{\mathcal{J}}_m$, $m = 1, \ldots, N$, are defined by

$$\widetilde{\mathcal{J}}_m := \mathcal{J}_m + \sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} |\langle l_0^n(t)\,(\Delta_h^n - \Delta_h^{n-1})U^{n-1}\,, \hat{\rho}(s)\rangle| ds. \tag{4.24}$$

The terms that contribute to $\mathcal{J}_m$ will be estimated in the next section. The last term of the right-hand side can be bounded by

$$\sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} |\langle l_0^n(t)\,(\Delta_h^n - \Delta_h^{n-1})U^{n-1}\,, \hat{\rho}(s)\rangle| ds \le \sum_{n=1}^{m} \frac{k_n}{2} \|(\Delta_h^n - \Delta_h^{n-1})U^{n-1}\|^2 + \frac{1}{2}\sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} \|\hat{\rho}(s)\|^2 ds. \tag{4.25}$$

A slightly improved estimate can be derived by exploiting the possible orthogonality on $\mathbb{V}_h^n \cap \mathbb{V}_h^{n-1}$. In fact, using (5.3), one can see that

$$|(\Delta_h^n - \Delta_h^{n-1})U^{n-1}|_{-1} \le c_{3,1} \|\hat{h}_n(\Delta_h^n - \Delta_h^{n-1})U^{n-1}\|. \tag{4.26}$$

In this case we conclude that

$$\sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} |\langle l_0^n(t)\,(\Delta_h^n - \Delta_h^{n-1})U^{n-1}\,, \hat{\rho}(s)\rangle| ds \le \sum_{n=1}^{m} c_{3,1}^2 k_n \|\hat{h}_n(\Delta_h^n - \Delta_h^{n-1})U^{n-1}\|^2 + \frac{1}{4}\sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} |\hat{\rho}(s)|_1^2 ds. \tag{4.27}$$

Nevertheless, handling these estimators requires some caution. The numerical experiment in section 2 shows that things can go wrong even in a simple refinement strategy.

5. Final residual-based estimators. In this section we use the specific choices of reconstructions of sections 3.1 and 3.2 and residual-based estimators to estimate the spatial finite element errors. We conclude with two alternative final error estimates for the modified CN scheme. We recall first the stability property and the approximation properties of the Clément-type interpolant introduced in [18].

LEMMA 5.1. *Let $\mathcal{I}^n : H_0^1 \to \mathbb{V}_h^n$ be the Clément-type interpolant. Then we have*

$$|\mathcal{I}^n z|_1 \le c_1 |z|_1. \tag{5.1}$$

Further, for $j \le l+1$, the following approximation properties are satisfied:

$$\begin{aligned} &\|h_n^{-j}(z-\mathcal{I}^n z)\| \le c_{1,j}|z|_j, \\ &\|h_n^{1/2-j}(z-\mathcal{I}^n z)\|_{\Sigma_n} \le c_{2,j}|z|_j, \end{aligned} \tag{5.2}$$

where l is the finite element polynomial degree and the constants c_1, $c_{1,j}$, and $c_{2,j}$ depend only on the shape regularity of the family of triangulations $\{T_n\}_{n=0}^N$.

LEMMA 5.2. *Let $\hat{\mathcal{I}}^n : H_0^1 \to \mathbb{V}_h^n \cap \mathbb{V}_h^{n-1}$ be the Clément-type interpolant relative to the finest common coarsening of $\hat{T}_n$. Then, for $j \le l+1$, the following interpolation properties are valid:*

$$\|\hat{h}_n^{1/2-j}(z-\hat{\mathcal{I}}^n z)\|_{\check{\Sigma}_n \setminus \hat{\Sigma}_n} \le c_{3,j}|z|_j, \tag{5.3}$$

where l is as in Lemma 5.1*, and $c_{3,j}$ depend on the shape regularity of the family of triangulations $\{\mathcal{T}_n\}_{n=0}^N$ and on the number of bisections necessary to pass from $\mathcal{T}_{n-1}$ to $\mathcal{T}_n$.*

Remark 5.1. Notice that we do not make any explicit assumptions in what follows on the relation of $\mathbb{V}_h^n$ and $\mathbb{V}_h^{n-1}$ other than the validity (5.3).

For functions defined in a piecewise sense we shall use the notation

$$\begin{aligned} \|h_n^i(\Delta-\Delta_h^n)U^n\|_{\mathcal{T}_n}^2 &= \sum_{K\in\mathcal{T}_n} \|h_K^i(\Delta-\Delta_h^n)U^n\|_K^2, \\ \|h_n^{i+\frac{1}{2}}[\nabla U^n]\|_{\Sigma_n}^2 &= \sum_{e\in\Sigma_n} \|h_e^{i+\frac{1}{2}}[\nabla U^n]\|_e^2, \quad i=1,2. \end{aligned}$$

Moreover, we summarize the notation of the various estimators used in the following definition.

DEFINITION 5.3 ($L^\infty(L^2)$ error estimators). *Let c_1, $c_{i,j}$ be the constants in Lemmas* 5.1 *and* 5.2*. For C_E being the elliptic regularity constant*

$$|v|_2 \le C_E\|\Delta v\|, \quad v \in H^2(\Omega)\cap H_0^1(\Omega),$$

we denote

$$C_{j,2} = C_E\, c_{j,2}.$$

For $n=1,\dots,N$, we define the elliptic reconstruction error estimator appearing in the definition of both two- and three-point estimators

$$\varepsilon_n = C_{1,2}\|h_n^2(\Delta-\Delta_h^n)U^n\|_{\mathcal{T}_n} + C_{2,2}\|h_n^{3/2}J[\nabla U^n]\|_{\Sigma_n} \tag{5.4}$$

and the space-mesh error estimator that also appears in both two- and three-point estimators

$$\begin{aligned} \gamma_n := {} & C_{1,2}\|\hat{h}_n^2\big[k_n^{-1}(\Delta-\Delta_h^n)U^n - k_n^{-1}(\Delta-\Delta_h^{n-1})U^{n-1}\big]\|_{\hat{\mathcal{T}}_n} \\ & + C_{2,2}\|\hat{h}_n^{3/2}J[\nabla U^n - \nabla U^{n-1}]\|_{\hat{\Sigma}_n} + C_{3,2}\|\hat{h}_n^{3/2}J[\nabla U^n - \nabla U^{n-1}]\|_{\check{\Sigma}_n\setminus\hat{\Sigma}_n}. \end{aligned} \tag{5.5}$$

Let w_n and $\tilde{w}_n$ be as in Lemmas 3.4 *and* 3.5*, respectively. We define the time reconstruction error estimator that corresponds to the two-point reconstruction by*

$$\delta_n := \frac{k_n^2}{4}\big\{\|w_n\| + C_{1,2}\|h_n^2(\Delta-\Delta_h^n)w_n\|_{\mathcal{T}_n} + C_{2,2}\|h_n^{3/2}J[\nabla w_n]\|_{\Sigma_n}\big\} \tag{5.6}$$

and to the three-point reconstruction by

$$\tilde{\delta}_n := \frac{k_n^2}{4}\big\{\|\tilde{w}_n\| + C_{1,2}\|h_n^2(\Delta - \Delta_h^n)\tilde{w}_n\|_{\mathcal{T}_n} + C_{2,2}\|h_n^{3/2}J[\nabla\tilde{w}_n]\|_{\Sigma_n}\big\}. \tag{5.7}$$

Further we define the space error estimator corresponding to the two-point reconstruction

$$\eta_n := k_n\big\{C_{1,2}\|h_n^2(\Delta - \Delta_h^n)w_n\|_{\mathcal{T}_n} + C_{2,2}\|h_n^{3/2}J[\nabla w_n]\|_{\Sigma_n}\big\} \tag{5.8}$$

and to the three-point reconstruction

$$\tilde{\eta}_n := k_n\big\{C_{1,2}\|h_n^2(\Delta - \Delta_h^n)\tilde{w}_n\|_{\mathcal{T}_n} + C_{2,2}\|h_n^{3/2}J[\nabla\tilde{w}_n]\|_{\Sigma_n}\big\}. \tag{5.9}$$

We define the time error estimator in the case of the two-point reconstruction

$$\theta_n := \frac{k_n^2}{30}\big\{c_1|w_n|_1 + c_{1,1}\|h_n(-\Delta_h^n)w_n\|\big\} \tag{5.10}$$

and in the case of the three-point reconstruction

$$\tilde{\theta}_n^2 := \frac{k_n^2}{30}\big\{c_1|\tilde{w}_n|_1 + c_{1,1}\|h_n(-\Delta_h^n)\tilde{w}_n\|\big\}. \tag{5.11}$$

We define the coarsening error estimator

$$\beta_n := \|(\Pi^n - I)((-\Delta_h^{n-1})U^{n-1}\| + \|(\widetilde{\Pi}^n - I)k_n^{-1}U^{n-1}\| \tag{5.12}$$

and the data approximation estimators

$$\begin{aligned}\xi_{n,1} &:= \frac{1}{k_n}\int_{t^{n-1}}^{t^n}\|f(s) - \varphi(s)\|ds,\\ \xi_{n,2} &:= 2\,c_{1,1}\max\big\{\|h_n(I - P_0^n)f^{n-1}\|, \|h_n(I - P_0^n)f^{n-\frac{1}{2}}\|\big\}\end{aligned} \tag{5.13}$$

and

$$\tilde{\xi}_{n,1} := \frac{1}{k_n}\int_{t^{n-1}}^{t^n}\|f(s) - \hat{\varphi}(s)\|ds. \tag{5.14}$$

In the case of the three-point estimator the following additional estimator appears:

$$\begin{aligned}\tilde{\zeta}_n := \frac{k_n}{4(k_n + k_{n-1})}\Big\|\frac{k_{n-1}}{k_n}(-\Delta_h^n)U^n &- \left(\frac{2k_n + k_{n-1}}{k_n}\Pi^n - \pi^n\right)(-\Delta_h^{n-1})U^{n-1}\\ &+ \pi^n\,\Pi^{n-1}(-\Delta_h^{n-2})U^{n-2}\Big\|.\end{aligned} \tag{5.15}$$

The a posteriori bounds are summarized in the following theorem.

THEOREM 5.4 (complete $L^\infty(L^2)$ a posteriori error estimates). *For the reconstruction defined in section* 3.1 *and for* $m = 1, \ldots, N$, *the following* two-level *estimate holds:*

$$\begin{aligned}\max_{t\in[0,t^m]}\|u(t) - U(t)\| \le &\sqrt{2}\,\|u^0 - \mathcal{R}^0u^0\| + \left(2\sum_{n=1}^m k_n\theta_n^2\right)^{1/2} + \big\{\mathcal{E}_{m,1}^2 + \mathcal{E}_{m,2}^2\big\}^{1/2}\\ &+ \max_{0\le n\le m}\delta_n + \max_{0\le n\le m}\varepsilon_n,\end{aligned} \tag{5.16}$$

where

$$\mathcal{E}_{m,1} := 2\sum_{n=1}^{m} k_n(\eta_n + \gamma_n + \beta_n + \xi_{n,1}), \quad \mathcal{E}_{m,2} := \sum_{n=1}^{m} k_n^{1/2}\xi_{n,2}. \tag{5.17}$$

Alternatively, if we use the reconstruction defined in section 3.2 *the following* three-level *estimate holds for* $m = 1, \dots, N$*:*

$$\max_{t\in[0,t^m]} \|u(t) - U(t)\| \le \sqrt{2}\,\|u^0 - \mathcal{R}^0 u^0\| + \left(2\sum_{n=1}^{m} k_n\tilde{\theta}_n^2\right)^{1/2} + \tilde{\mathcal{E}}_{m,1} + \max_{0\le n\le m}\tilde{\delta}_n + \max_{0\le n\le m}\varepsilon_n, \tag{5.18}$$

where

$$\tilde{\mathcal{E}}_{m,1} := 2\sum_{n=1}^{m} k_n(\tilde{\eta}_n + \gamma_n + \beta_n + \tilde{\xi}_{n,1} + \tilde{\zeta}_n). \tag{5.19}$$

The proof of the theorem uses Theorem 4.3. The rest of the section will be devoted to the estimate of the elliptic error $\|\epsilon(t)\|$, and of the terms in the right-hand side of (3.11).

5.1. Elliptic estimators. An upper bound for the elliptic reconstruction error ϵ is stated next; it can be proved by applying standard techniques in a posteriori error analysis for elliptic problems; cf., e.g., [1, 20] and [16, 12].

LEMMA 5.5 ($L^\infty(L^2)$ estimate for the elliptic reconstruction error). *For any* $\varphi \in \mathbb{V}_h^n$ *there holds that*

$$\|(\mathcal{R}^n - I)\varphi_n\| \le C_{1,2}\|h_n^2(\Delta - \Delta_h^n)\varphi_n\|_{\mathcal{T}_n} + C_{2,2}\|h_n^{3/2}J[\nabla\varphi_n]\|_{\Sigma_n}. \tag{5.20}$$

In particular for $m = 1, \dots, N$*, the following estimate holds:*

$$\max_{t\in[0,t^m]} \|\epsilon(t)\| \le \max_{0\le n\le m} \varepsilon_n. \tag{5.21}$$

Proof. The proof of (5.4) is standard using a duality argument. According to (3.12) and (3.1), we get

$$\begin{aligned} \|\epsilon(t)\| &\le l_0^n(t)\|(\mathcal{R}^{n-1} - I)U^{n-1}\| + l_1^n(t)\|(\mathcal{R}^n - I)U^n\| \\ &\le \max\left\{\|(\mathcal{R}^n - I)U^n\|,\, \|(\mathcal{R}^{n-1} - I)U^{n-1}\|\right\}, \quad t \in I_n, \end{aligned} \tag{5.22}$$

and (5.21) follows in view of (5.4). ∎

5.2. Main time estimator. We shall show next an upper bound for the time reconstruction error. Before we proceed with the estimate, in view of (3.13), (3.19), and (3.31), we conclude that

$$\sigma(t) = \hat{U}(t) - \omega(t) = (t - t^{n-1})(t^n - t)\mathcal{R}^n w_n, \quad t \in I_n, \tag{5.23}$$

in the case of the two-point reconstruction and

$$\sigma(t) = (t - t^{n-1})(t^n - t)\mathcal{R}^n \tilde{w}_n, \quad t \in I_n, \tag{5.24}$$

in the case of the three-point reconstruction.

LEMMA 5.6 ($L^\infty(L^2)$ estimate for the time reconstruction error). *For $m = 1, \dots, N$, the following estimates hold:*

$$\max_{t\in[0,t^m]} \|\sigma(t)\| \le \max_{1\le n\le m} \delta_n \tag{5.25}$$

and

$$\max_{t\in[0,t^m]} \|\sigma(t)\| \le \max_{1\le n\le m} \tilde{\delta}_n \tag{5.26}$$

in the cases of the two- and three-point reconstructions, respectively.

Proof. Let v_n be equal either to w_n or to $\tilde{w}_n$ appearing in (5.23) and (5.24), respectively. The relation (5.30) can be written as follows:

$$\|\sigma(t)\|^2 = (t-t^{n-1})(t^n-t)\big\{\langle(\mathcal{R}^n - I)v_n, \sigma(t)\rangle + \langle v_n, \sigma(t)\rangle\big\}, \tag{5.27}$$

i.e.,

$$\|\sigma(t)\| \le |(t-t^{n-1})(t^n-t)|\big\{\|(\mathcal{R}^n - I)v_n\| + \|v_n\|\big\}. \tag{5.28}$$

Thus, using (5.4) we can conclude that

$$\max_{t^{n-1}\le t\le t^n} \|\sigma(t)\| \le \frac{k_n^2}{4}\big\{\|v_n\| + C_{1,2}\|h_n^2(\Delta - \Delta_h^n)v_n\|_{\mathcal{T}_n} + C_{2,2}\|h_n^{3/2}J[\nabla v_n]\|_{\Sigma_n}\big\}, \tag{5.29}$$

and the result of the lemma follows. ∎

Next, we shall bound the similar term $\mathcal{J}_n^T$ in Theorem 4.3 which measures the local time discretization error. Let v_n again be equal either to w_n or to $\tilde{w}_n$ appearing in (5.23) and (5.24), respectively. We have

$$|\sigma(t)|_1^2 = a(\sigma(t), \sigma(t)) = (t-t^{n-1})(t^n-t)a(\mathcal{R}^n v_n, \sigma(t)). \tag{5.30}$$

By the definition of the elliptic reconstruction (3.4), we obtain

$$\begin{aligned} |\sigma(t)|_1^2 &= (t-t^{n-1})(t^n-t)\langle(-\Delta_h^n)^n v_n, \sigma(t)\rangle \\ &\le (t-t^{n-1})(t^n-t)|(-\Delta_h^n)v_n|_{-1}|\sigma(t)|_1, \end{aligned}$$

and thus

$$\begin{aligned} \int_{t^{n-1}}^{t^n} |\sigma(t)|_1^2 dt &\le |(-\Delta_h^n)v_n|_{-1}^2 \int_{t^{n-1}}^{t^n} (t-t^{n-1})^2(t^n-t)^2 dt \\ &= \frac{k_n^5}{30}|(-\Delta_h^n)v_n|_{-1}^2. \end{aligned} \tag{5.31}$$

Moreover, we have

$$\begin{aligned} |-\Delta_h^n w_n|_{-1} &= \sup_{0\ne z\in H_0^1} \frac{\langle -\Delta_h^n v_n, z\rangle}{|z|_1} \\ &= \sup_{0\ne z\in H_0^1} \left\{\frac{\langle -\Delta_h^n v_n, \mathcal{I}^n z\rangle}{|z|_1} + \frac{\langle -\Delta_h^n v_n, z - \mathcal{I}^n z\rangle}{|z|_1}\right\}, \end{aligned} \tag{5.32}$$

where $\mathcal{I}^n z \in \mathbb{V}_h^n$ is the Clément-type interpolant of the function z. Now, according to the definition of the discrete Laplacian Δ_h^n (2.1) and using the stability property of the Clément-type interpolant (5.1), we have

$$\langle -\Delta_h^n v_n, \mathcal{I}^n z\rangle \le c_1|v_n|_1|z|_1. \tag{5.33}$$

Further, using the approximation properties of the Clément-type interpolant (5.2), we obtain

$$\langle -\Delta_h^n w_n, z - \mathcal{I}^n z\rangle \le c_{1,1}\|h_n(-\Delta_h^n)v_n\|\,|z|_1. \tag{5.34}$$

According to (5.33) and (5.34), (5.32) leads to

$$|(-\Delta_h^n)v_n|_{-1} \le c_1|v_n|_1 + c_{1,1}\|h_n(-\Delta_h^n)v_n\|. \tag{5.35}$$

We can thus conclude that

$$\mathcal{J}_n^T \le k_n\,\theta_n^2 \tag{5.36}$$

in the case of the two-point estimator and

$$\mathcal{J}_n^T \le k_n\,\tilde{\theta}_n^2 \tag{5.37}$$

in the three-point reconstruction case.

Remark 5.2. The relation (5.30) can be written as follows:

$$|\sigma(t)|_1^2 = (t - t^{n-1})(t^n - t)\big\{a((\mathcal{R}^n - I)v_n, \sigma(t)) + a(v_n, \sigma(t))\big\}. \tag{5.38}$$

For $t \in I_n$, let $\mathcal{I}^n\sigma(s) \in \mathbb{V}_h^n$ be the Clément-type interpolant of $\sigma(s)$. Then

$$a((\mathcal{R}^n - I)v_n, \sigma(s)) = a((\mathcal{R}^n - I)v_n, (\sigma - \mathcal{I}^n\sigma)(s)). \tag{5.39}$$

Then it is standard to control the elliptic reconstruction error yielding

$$\mathcal{J}_n^T \le \frac{k_n^5}{30}\big\{|v_n|_1 + c_{1,1}\|h_n(\Delta - \Delta_h^n)v_n\|_{\mathcal{T}_n} + c_{2,1}\|h_n^{1/2}J[\nabla v_n]\|_{\Sigma_n}\big\}^2, \tag{5.40}$$

where v_n is equal either to w_n or to $\tilde{w}_n$ depending on the choice of the time reconstruction.

5.3. Spatial error estimate. In order to estimate the term $\mathcal{J}_n^{S,1}$ in Theorem 4.3, which accounts for the space discretization error, we use just (5.4) (here we do not use a sharper negative norm estimate as in [16]). For $t \in I_n$, let $v_n(t) := \hat{F}(t) - F^{n-1/2}$. According to Lemmas 3.4 and 3.5 we have

$$v_n(t) = 2\,(s - t^{n-1/2})w_n, \quad t \in I_n, \tag{5.41}$$

or

$$v_n(t) = 2\,(s - t^{n-1/2})\tilde{w}_n, \quad t \in I_n, \tag{5.42}$$

in the case of the two-point or the three-point reconstruction. In addition,

$$\int_{t^{n-1}}^{t^n} |s - t^{n-1/2}|ds = \frac{k_n^2}{4}. \tag{5.43}$$

Since w_n and $\tilde{w}_n$ are piecewise constant in time and using Definition 5.3, we can thus conclude that

$$\mathcal{J}_n^{S,1} \le Ck_n \max_{s\in[0,t^m]}\|\hat{\rho}(s)\|\,\eta_n \quad \text{or} \quad \mathcal{J}_n^{S,1} \le Ck_n \max_{s\in[0,t^m]}\|\hat{\rho}(s)\|\,\tilde{\eta}_n. \tag{5.44}$$

5.4. Space estimator accounting for mesh changing. We shall estimate next the term $\mathcal{J}_n^{S,2}$ in Theorem 4.3 following [12]. For completeness we provide the main arguments. Let $z : [0,T] \to H_0^1$ be the solution of problem

$$a(\chi, z(t)) = \langle \hat{\rho}(t), \chi \rangle \quad \forall \chi \in H_0^1,\ t \in [0,T], \tag{5.45}$$

and let $\hat{\mathcal{I}}^n z(t) \in \mathbb{V}_h^n \cap \mathbb{V}_h^{n-1},\ t \in I_n$, be its Clément-type interpolant. Since $\hat{\mathcal{I}}^n z(t) \in \mathbb{V}_h^n \cap \mathbb{V}_h^{n-1}$, using first (5.45) and then the orthogonality property of the elliptic reconstruction (3.5) in $\mathbb{V}_h^{n-1} \cap \mathbb{V}_h^n$, we get

$$\begin{aligned} &\langle (\mathcal{R}^n - I)U^n - (\mathcal{R}^{n-1} - I)U^{n-1}, \hat{\rho}(t) \rangle \\ &\quad = a((\mathcal{R}^n - I)U^n - (\mathcal{R}^{n-1} - I)U^{n-1}, (z - \hat{\mathcal{I}}^n z)(t)). \end{aligned} \tag{5.46}$$

For each $s \in I_n$, in view of the definition of the elliptic reconstruction (3.4), integration by parts gives

(5.47)
$$\begin{aligned} &\langle (\mathcal{R}^n - I)U^n - (\mathcal{R}^{n-1} - I)U^{n-1}, \hat{\rho}(t) \rangle \\ &\quad = \sum_{K \in T_n} \left\{ \int_K (\Delta - \Delta_h^n) U^n (z - \hat{\mathcal{I}}^n z)(t) - \int_{\partial K} \nabla U^n \cdot \mathbf{n} \, (z - \hat{\mathcal{I}}^n z)(t) \right\} \\ &\qquad - \sum_{K \in T_{n-1}} \left\{ \int_K (\Delta - \Delta_h^{n-1}) U^{n-1} (z - \hat{\mathcal{I}}^n z)(t) - \int_{\partial K} \nabla U^{n-1} \cdot \mathbf{n} \, (z - \hat{\mathcal{I}}^n z)(t) \right\}, \end{aligned}$$

i.e.,

(5.48)
$$\begin{aligned} &\langle (\mathcal{R}^n - I)U^n - (\mathcal{R}^{n-1} - I)U^{n-1}, \hat{\rho}(t) \rangle \\ &\quad = \sum_{K \in \hat{T}_n} \int_K \big\{ (\Delta - \Delta_h^n) U^n - (\Delta - \Delta_h^{n-1}) U^{n-1} \big\} (z - \hat{\mathcal{I}}^n z)(t) \\ &\qquad - \sum_{K \in T_n} \int_{\partial K} \nabla U^n \cdot \mathbf{n} \, (z - \hat{\mathcal{I}}^n z)(t) \, ds + \sum_{K \in T_{n-1}} \int_{\partial K} \nabla U^{n-1} \cdot \mathbf{n} \, (z - \hat{\mathcal{I}}^n z)(t). \end{aligned}$$

Also,

(5.49)
$$\begin{aligned} &\sum_{K \in \hat{T}_n} \int_K \big\{ (\Delta - \Delta_h^n) U^n - (\Delta - \Delta_h^{n-1}) U^{n-1} \big\} (z - \hat{\mathcal{I}}^n z)(t) \\ &\qquad \le c_{1,2} \|\hat{h}_n^2 \big\{ (\Delta - \Delta_h^n) U^n - (\Delta - \Delta_h^{n-1}) U^{n-1} \big\}\| \, \|\hat{\rho}(t)\|. \end{aligned}$$

For the estimation of the jump term, we first notice that

$$\begin{aligned} &\sum_{K \in T_n} \int_{\partial K} \nabla U^n \cdot \mathbf{n} \, (z - \hat{\mathcal{I}}^n z)(t) - \sum_{K \in T_{n-1}} \int_{\partial K} \nabla U^{n-1} \cdot \mathbf{n} \, (z - \hat{\mathcal{I}}^n z)(t) \\ &\qquad = \sum_{e \in \check{\Sigma}_n} \int_e J[\nabla U^n - \nabla U^{n-1}](z - \hat{\mathcal{I}}^n z)(t) \, . \end{aligned} \tag{5.50}$$

Since

$$\begin{aligned} &\sum_{e \in \check{\Sigma}_n} \int_e J[\nabla U^n - \nabla U^{n-1}](z - \hat{\mathcal{I}}^n z)(t) \\ &\qquad \le \|\hat{h}_n^{3/2} J[\nabla U^n - \nabla U^{n-1}]\|_{\hat{\Sigma}_n} \|\hat{h}_n^{-3/2} (z - \hat{\mathcal{I}}^n z)(t)\|_{\hat{\Sigma}_n} \\ &\qquad\quad + \|\hat{h}_n^{3/2} J[\nabla U^n - \nabla U^{n-1}]\|_{\check{\Sigma}_n \setminus \hat{\Sigma}_n} \|\hat{h}_n^{-3/2} (z - \hat{\mathcal{I}}^n z)(t)\|_{\check{\Sigma}_n \setminus \hat{\Sigma}_n}, \end{aligned}$$

again applying the interpolation inequalities (5.2) and (5.3), we get

(5.51)
$$\sum_{K\in T_n}\int_{\partial K}\nabla U^n\cdot\mathbf{n}\,(z-\hat{\mathcal{I}}^n z)(t)\;-\;\sum_{K\in T_{n-1}}\int_{\partial K}\nabla U^{n-1}\cdot\mathbf{n}\,(z-\hat{\mathcal{I}}^n z)(t),$$
$$\le\left\{C_{2,2}\|\hat{h}_n^{3/2}J[\nabla U^n-\nabla U^{n-1}]\|_{\hat{\Sigma}_n}+C_{3,2}\|\hat{h}_n^{3/2}J[\nabla U^n-\nabla U^{n-1}]\|_{\check{\Sigma}_n\setminus\hat{\Sigma}_n}\right\}\|\hat{\rho}(t)\|.$$

Thus, we can finally conclude that

$$\mathcal{J}_n^{S,2}\le k_n\max_{t\in[0,t^m]}\|\hat{\rho}(t)\|\,\gamma_n. \tag{5.52}$$

5.5. Coarsening error estimate. The term $\mathcal{J}_n^C$ in Theorem 4.3 can obviously be bounded as follows:

$$\mathcal{J}_n^C\le k_n\max_{t\in[0,t^m]}\|\hat{\rho}(t)\|\,\beta_n. \tag{5.53}$$

Remark 5.3. In the case of $\Pi^n=\widetilde{\Pi}^n:=P_0^n$, we can estimate the terms $\mathcal{J}_n^C$ by exploiting the orthogonality property of P_0^n. Indeed, for $t\in I_n$, let $\mathcal{I}^n\hat{\rho}(t)\in\mathbb{V}_h^n\cap\mathbb{V}_h^{n-1}$ be the Clément-type interpolant of $\hat{\rho}(t)$. Then we have

$$\begin{aligned}
&\left\langle(I-P_0^n)\left(\Delta_h^{n-1}U^{n-1}+\frac{U^{n-1}}{k_n}\right),\hat{\rho}(s)\right\rangle\\
&\quad=\left\langle(I-P_0^n)\left(\Delta_h^{n-1}U^{n-1}+\frac{U^{n-1}}{k_n}\right),(\hat{\rho}-\mathcal{I}^n\hat{\rho})(s)\right\rangle\\
&\quad\le\left\|h_n(I-P_0^n)\left(\Delta_h^{n-1}U^{n-1}+\frac{U^{n-1}}{k_n}\right)\right\|\,\|h_n^{-1}(\hat{\rho}-\mathcal{I}^n\hat{\rho})(s)\|\\
&\quad\le c_{1,1}\left\|h_n(I-P_0^n)\left(\Delta_h^{n-1}U^{n-1}+\frac{U^{n-1}}{k_n}\right)\right\|\,|\hat{\rho}(s)|_1.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\int_{t^{n-1}}^{t^n}\left|\left\langle(I-P_0^n)\left(\Delta_h^{n-1}U^{n-1}+\frac{U^{n-1}}{k_n}\right),\hat{\rho}(s)\right\rangle\right|ds\\
&\quad\le c_{1,1}\left\|h_n(I-P_0^n)\left(\Delta_h^{n-1}U^{n-1}+\frac{U^{n-1}}{k_n}\right)\right\|\int_{t^{n-1}}^{t^n}|\hat{\rho}(s)|_1 ds\\
&\quad\le c_{1,1}k_n^{1/2}\left\|h_n(I-P_0^n)\left(\Delta_h^{n-1}U^{n-1}+\frac{U^{n-1}}{k_n}\right)\right\|\left(\int_{t^{n-1}}^{t^n}|\hat{\rho}(s)|_1^2 ds\right)^{1/2}.
\end{aligned} \tag{5.54}$$

5.6. Estimation of the term $\mathcal{J}_n^D$.

Two-point reconstruction. We shall first estimate the term $\mathcal{J}_n^D$ in Theorem 4.3 in the case of the two-point reconstruction.

In view of (3.22), the above mentioned term can be written as

$$\mathcal{J}_n^D=\int_{t^{n-1}}^{t^n}|\langle f(s)-P_0^n\varphi(s),\hat{\rho}(s)\rangle|\,ds, \tag{5.55}$$

and it may be bounded as follows:

$$\begin{aligned}
\mathcal{J}_n^D&=\int_{t^{n-1}}^{t^n}|\langle f(s)-P_0^n\varphi(s),\hat{\rho}(s)\rangle|\,ds\\
&\le\int_{t^{n-1}}^{t^n}\left\{|\langle f(s)-\varphi(s),\hat{\rho}(s)\rangle|+|\langle(I-P_0^n)\varphi(s),\hat{\rho}(s)\rangle|\right\}ds.
\end{aligned} \tag{5.56}$$

Now, we have

$$\int_{t^{n-1}}^{t^n} |\langle f(s)-\varphi(s),\hat\rho(t)\rangle|\,ds \le \max_{t\in[0,t^m]}\|\hat\rho(t)\|\int_{t^{n-1}}^{t^n}\|f(s)-\varphi(s)\|\,ds. \tag{5.57}$$

Further, by again using the orthogonality property of P_0^n, we obtain

$$\begin{aligned}\langle (I-P_0^n)\varphi(s),\hat\rho(s)\rangle &= \langle (I-P_0^n)\varphi(s),(\hat\rho-\mathcal{I}^n\hat\rho)(s)\rangle\\ &\le \|h_n(I-P_0^n)\varphi(s)\|\,\|h_n^{-1}(\hat\rho-\mathcal{I}^n\hat\rho)(s)\|\\ &\le c_{1,1}\|h_n(I-P_0^n)\varphi(s)\|\,|\hat\rho(s)|_1.\end{aligned}$$

Now, we notice that

$$\begin{aligned}\|h_n(I-P_0^n)\varphi(s)\| &= \|2k_n^{-1}(t^{n-\frac12}-t)h_n(I-P_0^n)f^{n-1}\\ &\quad + 2k_n^{-1}(t-t^{n-1})h_n(I-P_0^n)f^{n-\frac12}\|\\ &\le \max\{\|h_n(I-P_0^n)f^{n-1}\|,\|h_n(I-P_0^n)(2f^{n-\frac12}-f^{n-1})\|\}\\ &\le 2\,\max\{\|h_n(I-P_0^n)f^{n-1}\|,\|h_n(I-P_0^n)f^{n-\frac12}\|\}\end{aligned}$$

so that

$$\begin{aligned}\int_{t^{n-1}}^{t^n}|\langle (I-P_0^n)\varphi(s),\hat\rho(s)\rangle|\,ds &\le c_{1,1}\int_{t^{n-1}}^{t^n}\|h_n(I-P_0^n)\varphi(s)\|\,|\hat\rho(s)|_1 ds\\ &\le c_{1,1}\,k_n^{1/2}\,2\,\max\{\|h_n(I-P_0^n)f^{n-1}\|,\|h_n(I-P_0^n)f^{n-\frac12}\|\}\left(\int_{t^{n-1}}^{t^n}|\hat\rho(s)|_1^2\right)^{1/2}.\end{aligned}$$

Together our estimates lead to

$$\mathcal{J}_n^D \le k_n\max_{0\le t\le t^m}\|\hat\rho(t)\|\,\xi_{n,1}+k_n^{1/2}\xi_{n,2}\left(\int_{t^{n-1}}^{t^n}|\hat\rho(s)|_1^2\right)^{1/2}. \tag{5.58}$$

Three-point reconstruction. We shall next estimate the term $\mathcal{J}_n^D$ in Theorem 4.3 in the case of the three-point reconstruction.

According to Lemma 3.6, the above mentioned term may be bounded as follows:

$$\mathcal{J}_n^D \le \int_{t^{n-1}}^{t^n}|\langle f(s)-\hat\varphi(s),\hat\rho(s)\rangle|\,ds+\frac{2}{k_n+k_{n-1}}\int_{t^{n-1}}^{t^n}|s-t^{n-\frac12}|\,|\langle v_n,\hat\rho(s)\rangle|\,ds, \tag{5.59}$$

with

$$\begin{aligned}v_n := \frac12\Bigg[&\frac{k_{n-1}}{k_n}(-\Delta_h^n)U^n\ -\left(2\Pi^n+\frac{k_{n-1}}{k_n}\Pi^n-\pi^n\right)(-\Delta_h^{n-1})U^{n-1}\\ &+\ \pi^n\,\Pi^{n-1}(-\Delta_h^{n-2})U^{n-2}\Bigg].\end{aligned} \tag{5.60}$$

The first term on the right-hand side may be bounded as in the previous paragraph to get

$$\int_{t^{n-1}}^{t^n}|\langle f(s)-\hat\varphi(s),\hat\rho(s)\rangle|\le k_n\max_{0\le t\le t^m}\|\hat\rho(t)\|\,\tilde\xi_{n,1}. \tag{5.61}$$

Moreover, the second term may be estimated as follows:

$$\frac{2}{k_n+k_{n-1}}\int_{t^{n-1}}^{t^n}|s-t^{n-\frac12}|\,|\langle v_n,\hat\rho(s)\rangle|\,ds\le\frac{k_n^2}{2(k_n+k_{n-1})}\max_{0\le t\le t^m}\|\hat\rho(t)\|\,\|v_n\|. \tag{5.62}$$

5.7. Concluding step. We are ready now to complete the proof of the theorem. For the two point reconstruction we have the following lemma.

LEMMA 5.7. *For $m=1,\dots,N$, the following estimate holds:*

$$\max_{t\in[0,t^m]}\|\hat\rho(t)\|+\left(\int_0^{t^m}|\hat\rho(s)|_1^2\,ds\right)^{1/2}\le\sqrt{2}\,\|\hat\rho(0)\|+\left(2\sum_{n=1}^m k_n\theta_n^2\right)^{1/2}+\left\{\mathcal{E}_{m,1}^2+\mathcal{E}_{m,2}^2\right\}^{1/2}, \tag{5.63}$$

with $\mathcal{E}_{m,1}$, $\mathcal{E}_{m,2}$ as defined in Theorem 5.4.

Proof. In view of Theorem 4.3, we can easily show that

$$\max_{t\in[0,t^m]}\|\hat\rho(t)\|^2+\int_0^{t^m}|\hat\rho(s)|_1^2\,ds\le 2\,\|\hat\rho(0)\|^2+2\,\mathcal{J}_m. \tag{5.64}$$

Now, according to the previous steps of the proof, we have

$$\begin{aligned}\mathcal{J}_m&=\sum_{n=1}^m\mathcal{J}_n^T+2\sum_{n=1}^m\mathcal{J}_n^{S,1}+\mathcal{J}_n^{S,2}+\mathcal{J}_n^C+\mathcal{J}_n^D\\&\le\sum_{n=1}^m k_n\theta_n^2+2\max_{t\in[0,t^m]}\|\hat\rho(t)\|\sum_{n=1}^m k_n(\eta_n+\gamma_n+\beta_n+\xi_{n,1})\\&\quad+\sum_{n=1}^m k_n^{1/2}\xi_{n,2}\left(\int_{t^{n-1}}^{t^n}|\hat\rho(s)|_1^2\,ds\right)^{1/2}.\end{aligned} \tag{5.65}$$

Thus, we have

$$\begin{aligned}&\max_{t\in[0,t^m]}\|\hat\rho(t)\|^2+\int_0^{t^m}|\hat\rho(s)|_1^2\,ds\le\,2\|\hat\rho(0)\|^2+2\sum_{n=1}^m k_n\theta_n^2\\&+4\max_{t\in[0,t^m]}\|\hat\rho(t)\|\sum_{n=1}^m k_n(\eta_n+\gamma_n+\beta_n+\xi_{n,1})+2\sum_{n=1}^m k_n^{1/2}\xi_{n,2}\left(\int_{t^{n-1}}^{t^n}|\hat\rho(s)|_1^2\,ds\right)^{1/2}.\end{aligned} \tag{5.66}$$

In order to derive the final estimate, we shall make use of the following result: Let $c\in\mathbb{R}$ and $\mathbf{a}=(a_0,a_1,\dots,a_m)$, $\mathbf{b}=(b_0,b_1,\dots,b_m)\in\mathbb{R}^{m+1}$ such that $|\mathbf{a}|^2\le c^2+\mathbf{a}\cdot\mathbf{b}$; then $|\mathbf{a}|\le|c|+|\mathbf{b}|$. According to (5.66), we apply this fact to the case

$$\begin{aligned}c&=\left(2\,\|\hat\rho(0)\|^2+2\sum_{n=1}^m k_n\theta_n^2\right)^{1/2},\\a_0&=\max_{t\in[0,t^m]}\|\hat\rho(t)\|,\quad a_n=\left(\int_{t^{n-1}}^{t^n}|\hat\rho(s)|_1^2\,ds\right)^{1/2},\ n=1,\dots,m,\\b_0&=4\sum_{n=1}^m k_n(\eta_n+\gamma_n+\beta_n+\xi_{n,1}),\quad b_n=2\,k_n^{1/2}\xi_{n,2},\ n=1,\dots,m,\end{aligned} \tag{5.67}$$

to get the result claimed. ∎

Similarly we can prove the following estimate in the case of the three-point reconstruction.

LEMMA 5.8. *For $m = 1, \dots, N$, the following estimate holds:*

(5.68)
$$\max_{t\in[0,t^m]} \|\hat{\rho}(t)\| + \left(\int_0^{t^m} |\hat{\rho}(s)|_1^2 \, ds\right)^{1/2} \le \sqrt{2}\,\|\hat{\rho}(0)\| + \left(2\sum_{n=1}^{m} k_n \tilde{\theta}_n^2\right)^{1/2} + \tilde{\mathcal{E}}_{m,1},$$

where $\tilde{\mathcal{E}}_{m,1}$ is defined in Theorem 5.4.

Combining the above estimates we complete the proof of Theorem 5.4.

6. Behavior of schemes and estimators. In this section we discuss in more detail the computational example presented in the introduction, and we include for completeness the main conclusions of [4] regarding the qualitative and computational behavior of the estimators derived in this paper.

6.1. Standard and modified schemes. We present further computational examples regarding the behavior of the standard and modified CN schemes. We consider the following test case: We solve the homogeneous heat equation with $u^0 = \sin(2\pi x)$, $x \in [0, 1]$. Oscillations for the standard scheme are noticeable at reasonable small time-step sizes. We have observed that they grow as time increases. On the other hand, oscillations tend to diminish as the time-step size decreases; see Figure 2. The appearance of the oscillations is a phenomenon due to the mesh change and not to the subtle behavior of the CN scheme per se. To verify this we plot the corresponding solutions of the CN scheme without mesh change for various values of the discretization parameters. In all cases the approximations are free from oscillations. A mathematical explanation of the appearance of oscillations during mesh change follows. In the standard CN scheme, the discrete Laplacian corresponding to the new grid (in the case of refinement, the fine grid) is applied to a function on the old grid (in our case, the coarse grid). It seems that this action produces oscillations. To see why, consider in one dimension a function $W \in \mathbb{V}_h$, where $\mathbb{V}_h$ is the standard finite element space of piecewise linear functions on a uniform partition of $[a, b]$ with mesh size h. Then

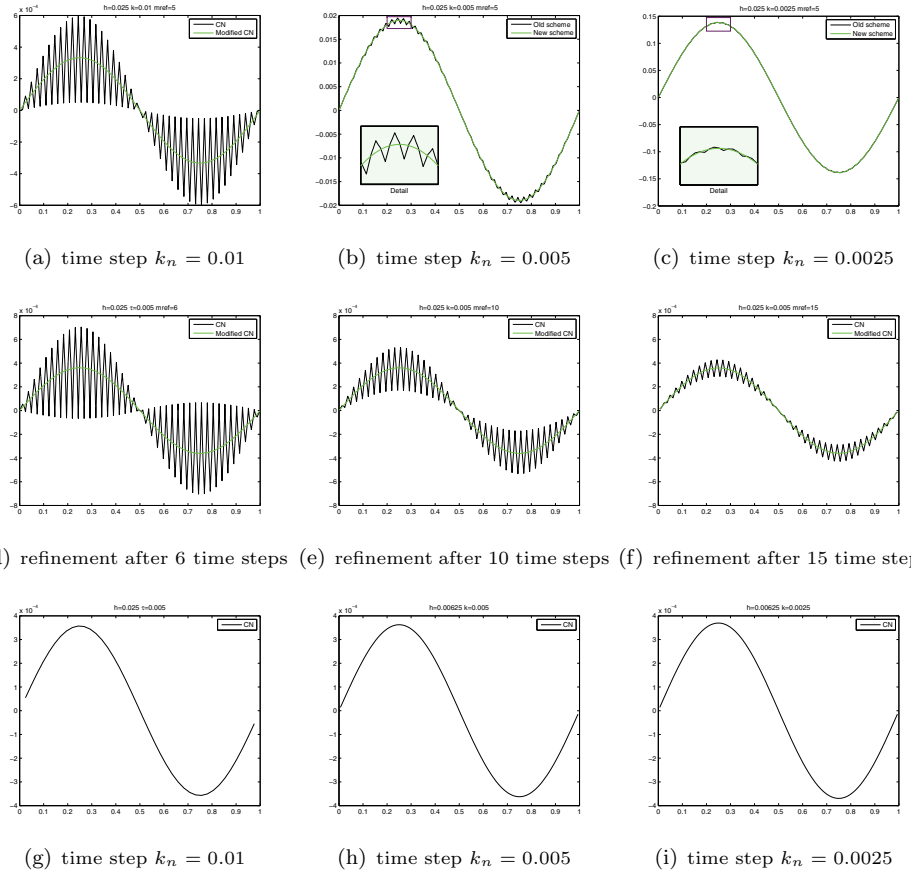
$$\int_a^b W'(x)\phi'(x)\,dx = \sum_j h\frac{1}{h^2}(W_{j+1} - W_j)(\phi_{j+1} - \phi_j) = \sum_j h\, g_j\, \phi_j.$$

Thus, modulo higher order quadrature errors, the discrete Laplacian of a piecewise linear function $W \in \mathbb{V}_h$ with nodal values W_j equals at the nodes

$$g_j = -\frac{1}{h^2}(W_{j+1} - 2W_j + W_{j-1}).$$

Now, let $v \in \mathbb{V}_{2h}$; then $v \in \mathbb{V}_h$. Let W_j be the nodal values of v in $\mathbb{V}_h$. Then for any other point $g_j = 0$, and obviously the other values of g_j are not zero in general. In particular, v can be chosen so that the discrete Laplacian produces noticeable oscillations. Similarly, we can see why the modified scheme does not produce spurious oscillations: first the discrete Laplacian of v in the $\mathbb{V}_{2h}$ space is calculated, a normal procedure since originally $v \in \mathbb{V}_{2h}$; then the resulting element of $\mathbb{V}_{2h}$ and thus of $\mathbb{V}_h$ is used in the modified CN scheme.

6.2. General remarks. We have verified that CN is a sensitive scheme for diffusion problems when it is combined with self-adjusted meshes. This sensitivity is



(a) time step $k_n = 0.01$ (b) time step $k_n = 0.005$ (c) time step $k_n = 0.0025$

(d) refinement after 6 time steps (e) refinement after 10 time steps (f) refinement after 15 time steps

(g) time step $k_n = 0.01$ (h) time step $k_n = 0.005$ (i) time step $k_n = 0.0025$

FIG. 2. *Behavior of the standard and modified CN schemes. First row: comparison of standard and modified schemes as time-step size decreases* ($k_n = 0.01, 0.005, 0.0025$). *Initial number of nodes* 40*; refinement after* 6 *time steps. Graph shown after* 20 *time steps. Second row:* $k_n = 0.01$, *initial number of nodes* 40. *Graphs at the same time but with decreasing number of refinements (after* 6*,* 10*, and* 15 *time steps). Third row: behavior of the CN scheme without mesh refinement at different time-step sizes and spatial nodes.*

related to mesh change with time, the behavior of the a posteriori error estimators, and, as expected, its dependence on the smoothness of the data.

First we have found that *refinement can spoil CN schemes*, and we have suggested a modified scheme which is natural and more robust with respect to mesh change. This finding is of particular interest since up to now refinement during mesh change was considered an error free procedure; see, e.g., [12]. In this paper we provide a complete a posteriori analysis for the modified scheme and also discuss the analysis of the standard CN discretization.

In [4] we have studied the behavior of the estimators from different perspectives. First, we have verified by presenting detailed computational experiments that the a posteriori estimators analyzed in this paper are of optimal order and their terms capture the spatial and the temporal errors separately; in Figure 3 we present here just an example showing this behavior. Here, for quantities of interest we look at their experimental order of convergence (EOC). The EOC is defined as follows: for a

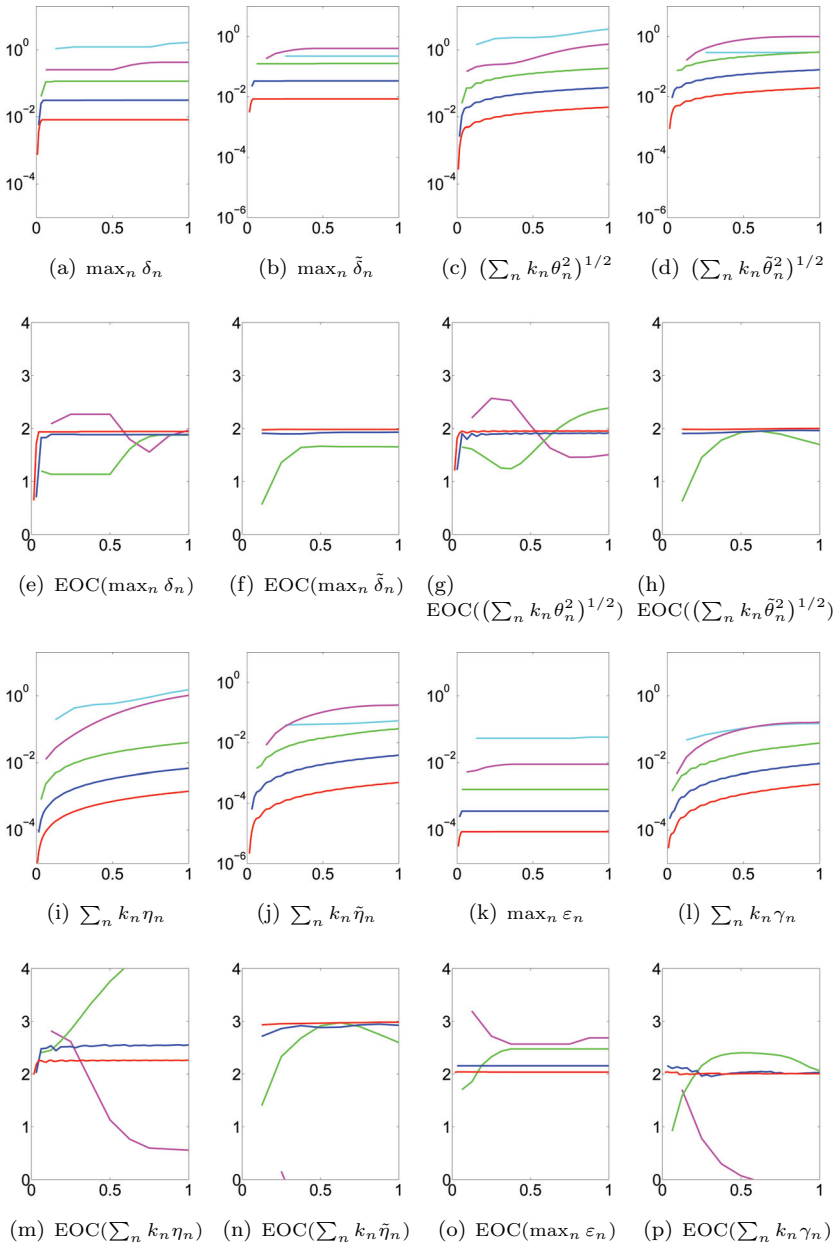


FIG. 3. *Numerical results for the problem* $u(x, y, t) = \sin(15\pi t)\sin(\pi x)\sin(\pi y)$. *In the first and third rows we plot the logs of estimators and below them the corresponding EOCs. We observe that all the estimators decrease with at least second order with respect to time and spatial mesh size. Moreover, in both cases the time error estimator and the time reconstruction dominate all other estimators. The space estimator* $\sum_n k_n \tilde{\eta}_n$ *superconverges.*

given finite sequence of successive runs (indexed by i), the EOC of the corresponding sequence of quantities of interest $E(i)$ (estimator or part of an estimator) itself is a sequence defined by $EOC(E(i)) = \frac{\log(E(i+1)/E(i))}{\log(h(i+1)/h(i))}$, where $h(i)$ denotes the mesh size of the run i. Each curve of the plots in Figure 3 corresponds to a given run (or an EOC of a quantity of interest) performed with equal time and spatial mesh sizes for $t \in [0, 1]$ (x-axis). The most coarse grid corresponds to $k = h = 0.125$ (curve with the largest error), and the finest grid corresponds to $k = h = 0.0078125$ (curve with the smallest error). The time and spatial mesh sizes are divided by two while moving from the highest to the lowest curve. On odd rows of each figure we plot the logs of the errors and the estimators and below them the corresponding EOC. The values of the EOC of a given estimator indicates its order. From the computational experiments it is clear that the order of the estimators is optimal; see [4] for details.

Next we studied the behavior of the estimators with respect to the discrete smoothness of the data. It is well known that the CN method requires further regularity assumptions on the data in order to be second order accurate [14, 19]. Our main conclusion using spectral arguments as well as computations is that the smoothness of the data can have a significant affect on the behavior of the error estimators. Here we refer to the smoothness of the discrete approximations of u^0 and not on the smoothness of u^0 per se. We also compared the two- and three-point estimators derived herein with respect to this criterion. It follows that the three-point estimator is less sensitive to the discrete smoothness of the data. Additionally, in [4], we provide an explicit computation which yields the leading terms of the two-point estimator w_n and those of the three-point estimator $\tilde{w}_n$ in a comparable form. It turns out that $\tilde{w}_n$ contains an additional "backward Euler" smoothing step compared to w_n, a fact that explains this difference in the behavior. It is interesting to note that we have found that the only difference between these terms is the fact that w_n involves the difference $U^n - U^{n-1}$ while $\tilde{w}_n$ involves the difference $U^n - U^{n-2}$. Hence, one may conclude that in a posteriori estimators equivalent terms in order corresponding to discrete derivatives are independent objects with possibly different behavior; see [4] for details.

Finally, in [4] we present a computational case study where refinement occurs at a given time level. This corresponds to the most "harmless" mesh modification, namely refinement. We observe that, when the exact solution of the heat equation is "fast" in the space variable, both the two- and the three-point estimators jump under the selected refinement procedure but the global three point estimator is only marginally affected. In case that the exact solution of the problem changes faster in time, the influence of the given refinement on the behavior of the estimators is very small or nonexistent.

We claim that the mesh change procedure is in some way related to data effects also discussed in [4]. In this light, nonstandard projections passing information from the previous finite element space to that corresponding to the next time step might be desirable since they may provide additional smoothness. This subtle issue requires further investigation. The analysis of the present paper allows such a possibility.

REFERENCES

[1] M. Ainsworth and J. T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Wiley-Interscience, New York, 2000.

[2] G. Akrivis, C. Makridakis, and R. H. Nochetto, *A posteriori error estimates for the Crank-Nicolson method for parabolic equations*, Math. Comp., 75 (2006), pp. 511–531.

[3] G. Akrivis, C. Makridakis, and R. H. Nochetto, *Optimal order a posteriori error estimates for a class of Runge-Kutta and Galerkin methods*, Numer. Math., 114 (2009), pp. 133–160.
[4] E. Bänsch, F. Karakatsani, and C. Makridakis, *The effect of mesh modification in time on the error control of fully discrete approximations for parabolic equations*, Appl. Numer. Math., to appear.
[5] W. Dörfler, *A time- and space-adaptive algorithm for the linear time-dependent Schrödinger equation*, Numer. Math., 73 (1996), pp. 419–448.
[6] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems* I*: A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.
[7] K. Eriksson, C. Johnson, and A. Logg, *Adaptive Computational Methods for Parabolic Problems*, in Encyclopedia of Computational Mechanics, John Wiley and Sons, New York, 2004.
[8] D. J. Estep, M. G. Larson, and R. D. Williams, *Estimating the error of numerical solutions of systems of reaction-diffusion equations*, Mem. Amer. Math. Soc., 146 (2000), no. 696.
[9] C. Johnson, Y.-Y. Nie, and V. Thomée, *An a posteriori error estimate and adaptive timestep control for a backward Euler discretization of a parabolic problem*, SIAM J. Numer. Anal., 27 (1990), pp. 277–291.
[10] O. Karakashian and C. Makridakis, *Error control for DG approximations of generalized KdV equations*, to appear.
[11] T. Katsaounis and I. Kyza, *A posteriori error estimates in the $l^\infty(l^2)$ norm for a fully discrete scheme of linear Schrödinger-type equations*, to appear.
[12] 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.
[13] A. Lozinski, M. Picasso, and V. Prachittham, *An anisotropic error estimator for the Crank–Nicolson method: Application to a parabolic problem*, SIAM J. Sci. Comput., 31 (2009), pp. 2757–2783.
[14] M. Luskin and R. Rannacher, *On the smoothing property of the Crank-Nicolson scheme*, Applicable Anal., 14 (1982/83), pp. 117–135.
[15] C. Makridakis, *Space and time reconstructions in a posteriori analysis of evolution problems*, in ESAIM Proceedings. Vol. 21 (2007) [Journées d'Analyse Fonctionnelle et Numérique en l'honneur de Michel Crouzeix], ESAIM Proc. 21, EDP Sciences, Les Ulis, France, 2007, pp. 31–44.
[16] 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.
[17] V. Prachittham, *Space-Time Adaptive Algorithms for Parabolic problems: A Posteriori Error Estimates and Application to Microfluidics*, Ph.D. thesis, EPFL Lausanne, Lausanne, Switzerland, 2009.
[18] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[19] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, Berlin, 1997.
[20] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Appl. Numer. Meth., John Wiley and Sons, Chichester, B. G. Teubner, Stuttgart, 1996.
[21] R. Verfürth, *A posteriori error estimates for finite element discretizations of the heat equation*, Calcolo, 40 (2003), pp. 195–212.