A POSTERIORI ERROR ANALYSIS FOR THE MEAN CURVATURE FLOW OF GRAPHS*

OMAR LAKKIS† AND RICARDO H. NOCHETTO‡

Abstract. We study the equation describing the motion of a nonparametric surface according to its mean curvature flow. This is a nonlinear nonuniformly parabolic PDE that can be discretized in space via a finite element method. We conduct an a posteriori error analysis of the spatial discretization and derive upper bounds of the error in terms of computable estimators based on local residual indicators. The reliability of the estimators is illustrated with two numerical simulations, one of which treats the case of a singular solution.

AMS subject classifications. Primary: 65N30, 65G20, 35K60; Secondary: 57R99, 40A30.

Key words. Finite element, mean curvature, error analysis, a posteriori, nonlinear PDE, parabolic equation, geometric motion, convergence, reliability, effectivity index.

1. Introduction. The objective of this article is the derivation of reliable a posteriori error estimates for the *mean curvature flow (MCF)* of a d-dimensional, time-dependent submanifold $\Gamma(t)$ of the Euclidean space $\mathbb{R}^{d+1}$. We pay special attention to the physically relevant cases ($d = 1, 2, 3$) and we refer to $\Gamma(t)$ simply as *moving surface*. A geometric definition of the MCF, whose details can be found in Huisken [15] and the references therein, is given by

$$\boldsymbol{V}(\boldsymbol{x}, t) = -\boldsymbol{\kappa}(\boldsymbol{x}, t), \text{ for } \boldsymbol{x} \in \Gamma(t), t \in \mathbb{R}, \tag{1.1}$$

where $\boldsymbol{V}$ and $\boldsymbol{\kappa}$ are respectively the velocity and the vector mean curvature of Γ. More general definitions of MCF are found in the literature, but will not be used [5, 11, 4].

In this paper we are interested in the *graph* (also called *nonparametric*) *description* in which the moving surface is described as the graph of a function u defined on a space-time domain $\Omega \times [0, T] \subset \mathbb{R}^d \times \mathbb{R}$. This description leads to the following PDE, referred to as the *Mean Curvature Flow of Graphs (MCFG)*,

$$\frac{\partial_t u(\boldsymbol{x}, t)}{Qu(\boldsymbol{x}, t)} - \frac{1}{d} \operatorname{div} \frac{\nabla u(\boldsymbol{x}, t)}{Qu(\boldsymbol{x}, t)} = 0, \qquad \text{for } \boldsymbol{x} \in \Omega, t \in [0, T] \tag{1.2}$$

where ∇ denotes the derivative with respect to $\boldsymbol{x}$ and Q the *elementary area* operator defined by

$$Qw := (1 + |\nabla w|^2)^{1/2}. \tag{1.3}$$

We drop the factor $1/d$, through a time rescaling by d, and we study the following initial-boundary value problem associated with (1.2) .

1.1 Problem (Cauchy-Dirichlet problem for the MCFG) *Given functions* $f : \Omega \times (0, T] \to \mathbb{R}$, *and* $g : \partial_{\mathrm{p}} (\Omega \times (0, T)) \to \mathbb{R}$, *find* $u : \overline{\Omega} \times [0, T] \to \mathbb{R}$ *such that*

$$\frac{\partial_t u(\boldsymbol{x}, t)}{Qu(\boldsymbol{x}, t)} - \operatorname{div} \frac{\nabla u(\boldsymbol{x}, t)}{Qu(\boldsymbol{x}, t)} = f(\boldsymbol{x}, t), \text{ for } (\boldsymbol{x}, t) \in \Omega \times (0, T], \tag{1.4}$$

* This work was partially supported by the NSF grants DMS-9971450, DMS-0204670, INT-9910086; Omar Lakkis was also supported by *Progetto MURST Cofin 2000* at *Università di Milano.*

†Institute of Applied and Computational Mathematics – FORTH PO Box 1527 – GR-71110 Iraklio, Greece. http://www.iacm.forth.gr/~omar – Email: omar@iacm.forth.gr

‡University of Maryland – Department of Mathematics and Institute for Physical Science and Technology – College Park, Maryland 20742. Email: rhn@math.umd.edu

$$u(\boldsymbol{x},t) = g(\boldsymbol{x},t), \text{ for } (\boldsymbol{x},t) \in \partial_{\mathrm{p}}\left(\Omega \times (0,T)\right), \tag{1.5}$$

where $\partial_{\mathrm{p}}\left(\Omega \times (0,T)\right)$ *is the* parabolic boundary *defined as* $\Omega \times \{0\} \cup \partial\Omega \times [0,T]$.

Arguably the MCF plays the role of model geometric motion, in the same way as the heat equation plays the role of model diffusion equation. Since more than two decades the MCF has been the object of mathematical analysis [5, 11, 15, 1, 16] as well as computer simulations [5, 22, 8] and numerical analysis [26, 7, 6]. It has also attracted the interests of practitioners, especially in the fields of material sciences and phase transition where the MCF, or some closely related geometric motion, models often the motion of a free boundary [3, 13].

A straightforward way to approximate numerically the solution of Problem 1.1 is firstly to discretize the spatial variable through a finite element method—which comes natural as (1.4) is written in "divergence form"—and secondly to discretize the time variable with a finite difference scheme known as *semi-implicit* in which the nonlinearity is treated explicitly and the linear part implicitly [8]. The first stage of this process, discussed in §2.5–2.6, is referred to as the *spatial (semi-) discretization.* Deckelnick & Dziuk [7] have derived a priori error estimates for both the spatially discrete and the semi-implicit fully discrete scheme.

The study of *a posteriori error estimates* for evolution equations, which has developed in the last 15 years, is mainly motivated by their successful use in deriving adaptive mesh refinement algorithms. The lack of such estimates in the case of the MCFG and the interest of adaptive methods for this problem are the driving motives behind this article. Our main results, discussed in §3, are a posteriori upper bounds of the error for the spatially discrete approximation. A posteriori error estimates have been established for linear parabolic problems [9, 19] and used to derive adaptive mesh refinement algorithms. Analogous results have been also derived for certain nonlinear elliptic [24, 12] and parabolic [18, 10, 17] equations, but these cannot be applied to the MCFG.

As observed since the early days of adaptive FEM [2], an adaptive mesh refinement algorithm must satisfy two fundamental properties: namely *reliability* and *efficiency.* These two algorithmic concepts are closely related to the nature of the error bounds. An algorithm for numerical approximation is called *reliable* if the error between its output and the exact solution is bounded by a given tolerance from above. In terms of estimators, reliability is achieved if the error/estimator ratio is bounded by a positive constant that is independent of the meshsize; this ratio is known as *effectivity index* in the literature. Notice that for the effectivity index to be positive and independent of the meshsize *it is necessary for the estimator to have the same order of convergence as the error*; this allows the upper bound to be used as a stopping criterion in adaptive algorithms. In this paper, our main concern, besides proving the error estimates, will be to understand whether these are reliable. For this, the numerical examples we shall present in §7 are geared toward comparing the numerical asymptotic decay rate of the error with that of the estimators.

The MCFG is an example of evolution equation that is not covered by any of the general techniques developed so far for the derivation of a posteriori error estimates for nonlinear equations [10, 17, 25]. The reason for this is mainly due to the nonuniformly parabolic nature of the equation and, more philosophically, to the fact that general non-linear theories end up being less reliable and harder to apply. In this paper we employ an ad-hoc energy technique to derive the estimates. This is the only practical way up to our knowledge. A distinctive feature of this paper is the use of special quantities to quantify the error. As for most nonuniformly parabolic equations, for

the MCFG the Sobolev norms are extremely hard to handle and we are naturally led to use the *geometric errors* introduced next. These are not Sobolev norms of the error $u-u_h$ where u and u_h are respectively the exact and approximate solutions (see §2.6). The geometric errors are not even symmetric in u and u_h, yet they quantify satisfactorily the error.

1.2 Definition (Geometric error) Let u be the solution of Problem 1.1 and u_h be the finite element solution given by Problem 2.6. For each $t \in [0,T]$, define

$$A(t) := \int_\Omega |\boldsymbol{N}u_h(\boldsymbol{x},t) - \boldsymbol{N}u(\boldsymbol{x},t)|^2 \, Qu(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x} \tag{1.6}$$

$$B(t) := \int_0^t \int_\Omega \left(Vu_h(\boldsymbol{x},s) - Vu(\boldsymbol{x},s)\right)^2 Qu(\boldsymbol{x},s)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}s, \tag{1.7}$$

where

$$\begin{gathered} \mathrm{W}^1(\Omega) \ni w \mapsto \boldsymbol{N}w := (\nabla w; -1)/Qw \in \mathrm{L}_\infty(\Omega)^{d+1} \\ \text{and } \mathrm{W}_1^1(\Omega\times(0,T)) \ni w \mapsto Vw := \partial_t w/Qw \in \mathrm{L}_1^{\mathrm{loc}}(\Omega\times(0,T)) \end{gathered} \tag{1.8}$$

are respectively the *normal vector* and the *normal velocity* operators. We will denote by $\mathrm{C}^k(\Omega)$ (resp. $\mathrm{W}^k(\Omega)$) the space of k times continuously (resp. weakly) differentiable functions, by $\mathrm{W}_p^k(\Omega)$ the usual Sobolev space of functions in $\mathrm{W}^k(\Omega)$ with derivatives in $\mathrm{L}_p(\Omega)$ and by $\mathring{\mathrm{W}}_p^k(\Omega)$ the subspace of functions with vanishing trace. The functions of time A and B are the building blocks of the *total geometric error* E defined by

$$E(t)^2 := B(t) + \sup_{[0,t]} A(s) = \int_0^t \int_\Omega \left(Vu_h - Vu\right)^2 Qu + \sup_{(0,t)} \int_\Omega |\boldsymbol{N}u_h - \boldsymbol{N}u|^2 \, Qu. \tag{1.9}$$

We refer to $\sup_{[0,t]} A^{1/2}$ and $B^{1/2}$ as *geometric energy error* and *normal velocity error*, respectively.

The integrals of the form $\int_\Omega \cdot\, Qu(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}$ in (1.9) can be interpreted as integrals over the moving surface $\Gamma(t)$ which give us the $\mathrm{L}_2(\Gamma)$ norm of the *difference of normals* and the *difference of normal velocities*. A comparison with the integrals appearing in the left-hand side of (2.6) explains in part why they "fit" the problem.

We point out that, despite the natural relation between our notion of error and MCFG, no related concept of error has yet been used in the context of a posteriori error control for parabolic equations. In fact, the geometric nature contrasts sharply with the pure analytic setting found for instance in Verfürth's monograph [24]. A related, symmetric, geometric error is employed by Fierro & Veeser for the stationary case [12].

It is important to observe that the sharpest estimate in this article, given by Theorem 3.6 is a *conditional estimate*. By conditional we mean that the estimate is valid only if a certain condition on how close the approximate solution is to the exact solution is satisfied. A relevant feature of our result in this respect is that the condition can be machine checked since it entails computable quantities. This is of paramount importance for a result to be fully "a posteriori" (see Remark 3.7). In this sense, and up to our knowledge, our result is the first conditional a posteriori estimate for nonlinear parabolic equations. Conditional results have been derived also for the prescribed mean curvature (elliptic) equation by Fierro & Veeser [12]. We notice that Verfürth has established also conditional results, but the conditions are not fully a posteriori and cannot be machine-checked [24]. In order to appreciate the

sharpness of the conditional result of Theorem 3.6, an *unconditional estimate* is given in Theorem 3.4 for sake of comparison. Our numerical results provide a practical comparison between the two theoretical bounds and show that the reliability is very good for the conditional estimate while it is poor for the unconditional estimate.

Deckelnick & Dziuk showed an a priori error bound of rate $\mathrm{O}(h)$ for the geometric error in the spatially discrete case [7]. The geometric error introduced in 1.2 is similar to that of Deckelnick & Dziuk, but in their case the integrals are evaluated on the discrete surface while we compute them on the exact surface. In this respect our a posteriori view-point approach can be seen, roughly speaking, as dual to their a priori approach. We notice however, that our results are valid under weaker regularity assumptions on the exact solution u (see 7.6). Our analysis also includes *time-dependent boundary value* g and non-homogeneous right-hand side f, while their analysis is limited to the homogeneous and time-independent boundary value case.

The rest of this paper is organized as follows. In §2 we discuss some properties of Problem 1.1 and introduce the associated spatial finite element method. In §3 we state the main results and make some observations. Next, in §§4–6 we prove these results. Finally, numerical tests are discussed in §7.

2. The Cauchy-Dirichlet problem and its spatial discretization.

2.1 Hypothesis (Solvability and regularity) *Unless otherwise stated, the following assumptions will hold:*

(a) Classical solvability: *Problem 1.1 admits a unique classical solution* u *in* $\mathrm{C}^{2,1}(\Omega\times(0,T])\cap\mathrm{C}^0(\overline{\Omega}\times[0,T])$, *for some* $T>0$.

(b) Boundary regularity of contact angle:[1]

$$\frac{\nabla u(t)}{Qu(t)}\in \mathrm{W}^1_d(\Omega),\ \forall t\in[0,T]. \tag{2.1}$$

(c) Regularity of normal velocity

$$Vu(t)=\frac{\partial_t u(t)}{Qu(t)}\in \mathrm{L}_d(\Omega),\ \forall t\in[0,T]. \tag{2.2}$$

(d) Regularity of vertical velocity:

$$\partial_t u(t)\in \mathrm{W}^1_1(\Omega)\cap\mathrm{L}_2(\Omega),\ \forall t\in[0,T]. \tag{2.3}$$

2.2 Remark (About the regularity assumptions) Assumption 2.1 (a) is backed by the fact that Problem (1.4) admits indeed classical solutions under certain sufficient conditions relating the mean-convexity of $\partial\Omega$ and the function $|f|$ [16, § 12.8]. Solutions, which are classical up to blow-up, can exist also in more general situations where the domain is non-mean-convex or compatibility conditions are violated [23]. There are two implicit assumptions that are immediate consequences of Hypothesis 2.1 (a): we necessarily have $f\in\mathrm{C}^0(\Omega\times(0,T])$ and $g\in\mathrm{C}^0(\partial_{\mathrm{p}}(\Omega\times(0,T]))$. Although a "weak form" of Problem 1.1 will be derived in §2.3, we do not know of any satisfactory concept of "weak solution" for it.

[1] We use the following convention throughout this article: Whenever a space time function $w:\Omega\times[0,T]\to\mathbb{R}^N$ $(N=1,d)$ is written with only one argument it means that the argument is a time variable, and that value—e.g., $w(t)$, or $w(1/2)$—is a function with domain Ω.

The reason we assume (2.3) is technical: this assumption will be needed to test (1.4) by $\partial_t u$ (see §2.3, 2.4 and 4.3). Notice that for $d \le 2$, in view of the Sobolev embedding, this assumption can be simplified to $\partial_t u \in \mathrm{W}^1_1(\Omega)$ and implies (2.2). Notice also that, for $d \ge 1$, the Sobolev embedding and (2.3) imply that $\partial_t u \in \mathrm{L}_{d'}(\Omega)$ for $d' = d/(d-1)$.

2.3 Proposition (Weak form) *Let $u \in \mathrm{C}^{2,1}(\Omega \times (0,T]) \cap \mathrm{C}^0(\overline{\Omega} \times [0,T])$ be a given function that satisfies* (2.1) *and* (2.2)*. The function u is a classical solution of Problem 1.1 if and only if*

$$\left\langle \frac{\partial_t u(t)}{Qu(t)}, \phi \right\rangle + \left\langle \frac{\nabla u(t)}{Qu(t)}, \nabla \phi \right\rangle = \langle f(t), \phi \rangle \,, \ \forall \phi \in \mathring{\mathrm{W}}^1_1(\Omega), t \in (0,T], \tag{2.4}$$

$$u(t) - \tilde{g}(t) \in \mathring{\mathrm{W}}^1_1(\Omega) \, \forall t \in (0,T], \quad \text{and} \quad u(0) = g(0) \tag{2.5}$$

where $\tilde{g}(t)$ is an extension of $g(t)$ to all of Ω.

We use the notation $\langle v, w \rangle_D := \int_D v(\boldsymbol{x}) w(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x})$, for functions v and w such that $vw \in \mathrm{L}_1(D,\mu)$, $D \subset \mathbb{R}^d$ and $\mathrm{d}\mu$ is the Lebesgue measure "$\mathrm{d}\cdot$" or the $(d-1)$-dimensional Hausdorff measure, depending on the Hausdorff dimension of D. If $D = \Omega$ we omit the subscript in the brackets.

The proof of Proposition 2.3 follows basic PDE techniques and is omitted. We observe that the existence of $\tilde{g}$, for $g(t) \in \mathrm{L}_1(\Omega)$, is guaranteed in view of [21, Eq. (5.5)].

2.4 Lemma (Stability estimate) *If $f \in \mathrm{L}_2(0,T;\mathrm{L}_\infty(\Omega))$ and $g \in \mathrm{W}^1_1(\partial_{\mathrm{p}}(\Omega \times (0,T)))$ then*

$$\begin{aligned} &\frac{1}{2} \int_0^t \int_\Omega |Vu|^2 \, Qu + \int_\Omega Qu(t) \\ &\qquad \le \exp\left(\frac{1}{2} \int_0^t \|f\|^2_{\mathrm{L}_\infty(\Omega)} \right) \left(\|Qg(0)\|_{\mathrm{L}_1(\Omega)} + \|\partial_t g\|_{\mathrm{L}_1(\partial\Omega \times (0,t))} \right). \end{aligned} \tag{2.6}$$

Proof Test (1.4) by $\partial_t u \in \mathrm{L}_{d'}(\Omega)$ and, owing to (2.1) and (2.3), apply the integration by parts formula on Ω:

$$0 = \int_\Omega |Vu|^2 \, Qu + \int_\Omega \frac{\nabla u}{Qu} \cdot \nabla \partial_t u - \int_{\partial\Omega} \frac{\nabla u \cdot \boldsymbol{\nu}}{Qu} \partial_t u - \int_\Omega f \partial_t u. \tag{2.7}$$

The first term, which is equal to $\int_\Omega Vu \, \partial_t u$, is well defined thanks to (2.3) and (2.2). The third and fourth terms are bounded as follows:

$$\int_{\partial\Omega} \frac{\nabla u \cdot \boldsymbol{\nu}}{Qu} \partial_t u = \int_{\partial\Omega} \left(\frac{\nabla u}{Qu} \cdot \boldsymbol{\nu} \right) \partial_t g \le \|\partial_t g\|_{\mathrm{L}_1(\partial\Omega)} \,, \tag{2.8}$$

$$\int_\Omega f \partial_t u = \int_\Omega f \sqrt{Qu} \frac{\partial_t u}{\sqrt{Qu}} \le \frac{1}{2} \int_\Omega |Vu|^2 \, Qu + \frac{1}{2} \|f\|^2_{\mathrm{L}_\infty(\Omega)} \int_\Omega Qu. \tag{2.9}$$

Next we observe that the basic identity

$$\partial_t Qu(\boldsymbol{x},t) = \partial_t \sqrt{1 + |\nabla u|^2} = \frac{\nabla u \cdot \partial_t \nabla u}{Qu} \tag{2.10}$$

implies

$$\frac{1}{2} \int_\Omega |Vu|^2 \, Qu + \mathrm{d}_t \int_\Omega Qu \le \|\partial_t g\|_{\mathrm{L}_1(\partial\Omega)} + \frac{1}{2} \|f\|^2_{\mathrm{L}_\infty(\Omega)} \int_\Omega Qu. \tag{2.11}$$

The result is obtained by integrating on $[0,t]$ and applying the Gronwall lemma. ∎

Inequality (2.6) acquires a geometric meaning upon observing that $\int_\Omega Qu$ is the area of graph(u). This gives us a control on the growth of the area in time in terms of data. In particular, if the forcing term $f = 0$ and the boundary conditions are time-independent, then (2.6) quantifies the decrease in area of the graph that tends toward a nonparametric minimal surface as time grows. The MCF, with $f = 0$, is thus interpreted as the gradient descent method for the area functional, with respect to the $\mathrm{L}_2(\Gamma(t))$ norm.

2.5. Finite element discretization. We start by introducing $\{\mathscr{T}_h\}_h$, a shape-regular family of triangulations (simplicial partitions) of the domain Ω. This means that there exists a constant $\sigma_0 \in \mathbb{R}^+$, independent of the particular triangulation $\mathscr{T}_h$, such that

$$\frac{\sup\{\rho \in \mathbb{R}^+ :\ \mathrm{B}_\rho(\boldsymbol{x}) \subset K\}}{\operatorname{diam}(K)} \geq \sigma_0,\ \forall K \in \mathscr{T}_h. \tag{2.12}$$

We will refer to σ_0 as the *shape-regularity* of the family $\{\mathscr{T}_h\}_h$. We assume that the *approximate domain* $\Omega_h = \operatorname{int}\left(\bigcup_{K\in\mathscr{T}_h} \overline{K}\right)$, coincides with Ω; this is a simplifying assumption that could be removed at the cost of seriously complicating the analysis, without adding much content to the results we intend to present. The symbol h stands for both the *local meshsize function* and the global meshsize of $\mathscr{T}_h$; this abuse of notation should not cause confusion.

Given a simplex $K \in \mathscr{T}_h$ and $\psi : \Omega \to \mathbb{R}$, we denote by ψ_K the restriction $\psi|_K$—e.g., if $\psi = h$ we have $h_K = \operatorname{diam}(K)$—and by $\mathscr{U}_K^h$ the $\mathscr{T}_h$*-neighborhood* of K

$$\mathscr{U}_K^h := \operatorname{int}\left(\bigcup\left\{\overline{K'} \in \mathscr{T}_h :\ \overline{K'} \cap \overline{K} \neq \emptyset\right\}\right). \tag{2.13}$$

We also associate with $\mathscr{T}_h$ its *internal mesh* $\Sigma_h := \bigcup_{S\in\mathscr{S}_h^\circ} S$, where $\mathscr{S}_h^\circ$ is the set of internal edges (or faces) of the simplexes in $\mathscr{T}_h$. The *finite element spaces* constructed on $\mathscr{T}_h$ that will be employed are

$$\mathbb{V}_h := \left\{\phi \in \mathrm{W}_1^1(\Omega) :\ \phi_K \in \mathbb{P}^\ell,\ \forall K \in \mathscr{T}_h\right\} \text{ and } \mathring{\mathbb{V}}_h := \mathbb{V}_h \cap \mathring{\mathrm{W}}_1^1(\Omega), \tag{2.14}$$

where $\ell \in \mathbb{Z}^+$ and $\mathbb{P}^\ell$ is the space of polynomials of degree at most ℓ. A spatial finite element discretization of Problem 1.1 can be now derived from (2.4).

2.6 Problem (Spatially discrete scheme for the MCFG) *Let $\tilde{g}_h(t) \in \mathbb{V}_h$ be an* interpolation *of $\tilde{g}(t)$. Find $u_h \in \mathrm{C}^1([0,T];\mathbb{V}_h)$ such that for each $t \in [0,T]$*

$$u_h(t) - \tilde{g}_h(t) \in \mathring{\mathbb{V}}_h, \tag{2.15}$$

$$\left\langle \frac{\partial_t u_h(t)}{Qu_h(t)}, \phi_h \right\rangle + \left\langle \frac{\nabla u_h(t)}{Qu_h(t)}, \nabla\phi_h \right\rangle = \langle f(t), \phi_h\rangle,\ \forall \phi_h \in \mathring{\mathbb{V}}_h. \tag{2.16}$$

Solvability of Problem 2.6 and a priori error estimates are studied by Deckelnick & Dziuk [7]. Throughout the paper u_h will denote the solution of Problem 2.6.

3. A posteriori error estimates. In this section we state our main results. We start by introducing some definitions.

3.1 Definition (Residual functions) For each $t \in [0,T]$, let $r(t)$ be the *internal residual* and $j(t)$ be the *jump residual* associated with u_h. These two functions are defined on $\Omega \setminus \Sigma_h$ and Σ_h respectively, and are given by

$$r(\boldsymbol{x},t) := \frac{\partial_t u_h(\boldsymbol{x},t)}{Qu_h(\boldsymbol{x},t)} - f(\boldsymbol{x},t) - \operatorname{div}\left(\frac{\nabla u_h(\boldsymbol{x},t)}{Qu_h(\boldsymbol{x},t)}\right),\ \text{for } \boldsymbol{x} \in \Omega \setminus \Sigma_h, \tag{3.1}$$

$$j(\boldsymbol{x},t) := \left[\!\!\left[\frac{\nabla u_h(\boldsymbol{x},t)}{Qu_h(\boldsymbol{x},t)} \right]\!\!\right]_S, \text{ for } \boldsymbol{x} \in S \in \mathscr{S}_h^\circ, \tag{3.2}$$

where the *jump of a vector field* $\boldsymbol{\psi}$ across an edge S is defined as

$$[\![\boldsymbol{\psi}]\!]_S(\boldsymbol{x}) := \lim_{\varepsilon \to 0} \left(\boldsymbol{\psi}(\boldsymbol{x} + \varepsilon \boldsymbol{\nu}_S) - \boldsymbol{\psi}(\boldsymbol{x} - \varepsilon \boldsymbol{\nu}_S)\right) \cdot \boldsymbol{\nu}_S \tag{3.3}$$

with $\boldsymbol{x} \in S$ and $\boldsymbol{\nu}_S$ denoting one of the two normals to S (the choice is arbitrary and does not affect the definition).

3.2 Definition (Local indicators and weights) Denote by C_1 and C_2 the Scott-Zhang interpolation inequality constants, which depend only on the shape-regularity σ_0 of $\mathscr{T}_h$ and which we introduce later in inequalities (6.2) and (6.3), respectively. With each $K \in \mathscr{T}_h$ we associate the *local*

$$\textit{elliptic indicator} \qquad \eta_0^K(t) := h_K^{d/2} \left(C_1 \left\| r(t) \right\|_{\mathrm{L}_d(K)} + C_2 \left\| j(t) \right\|_{\mathrm{L}_\infty(\partial K)} \right), \tag{3.4}$$

$$\textit{parabolic indicator} \qquad \eta_1^K(t) := h_K^{d/2} \left(C_1 \left\| \partial_t r(t) \right\|_{\mathrm{L}_d(K)} + C_2 \left\| \partial_t j(t) \right\|_{\mathrm{L}_\infty(\partial K)} \right) \tag{3.5}$$

and the *local weights*

$$\omega^K(t) := \sup_{\boldsymbol{x} \in \mathscr{U}_K^h} Qu_h(\boldsymbol{x},t)^2, \qquad \alpha^K(t) := \omega^K(t)^2 \sup_{\boldsymbol{x} \in \mathscr{U}_K^h} \frac{1}{Qu(\boldsymbol{x},t)}. \tag{3.6}$$

3.3 Definition (A posteriori error estimators) Denote by M and γ two positive constants, depending only on the shape-regularity σ_0, which we will introduce in detail in the proof of Lemma 6.4. We define the *elliptic part of the proper estimator*

$$\mathscr{E}_{2,0}(t) := \sup_{s \in [0,t]} \hat{\mathscr{E}}_{2,0}(s) \qquad \text{where} \quad \hat{\mathscr{E}}_{2,0}(t)^2 := \gamma^2 \sum_{K \in \mathscr{T}_h} \alpha^K(t) \eta_0^K(t)^2, \tag{3.7}$$

the *parabolic part of the proper estimator*

$$\mathscr{E}_{2,1}(t) := \int_0^t \dot{\mathscr{E}}_{2,1}(s)\,\mathrm{d}s \qquad \text{where} \quad \dot{\mathscr{E}}_{2,1}(t)^2 := \gamma^2 \sum_{K \in \mathscr{T}_h} \alpha^K(t) \eta_1^K(t)^2, \tag{3.8}$$

the *elliptic part of the vicinity estimator*

$$\mathscr{E}_{\infty,0}(t) := \sup_{s \in [0,t]} \hat{\mathscr{E}}_{\infty,0}(s) \quad \text{where} \quad \hat{\mathscr{E}}_{\infty,0}(t) := M \max_{K \in \mathscr{T}_h} \left(h_K^{-d/2} \omega^K(t) \eta_0^K(t) \right) \tag{3.9}$$

and the *parabolic part of the vicinity estimator*

$$\mathscr{E}_{\infty,1}(t) := \int_0^t \dot{\mathscr{E}}_{\infty,1}(s)\,\mathrm{d}s \quad \text{where} \quad \dot{\mathscr{E}}_{\infty,1}(t) := M \max_{K \in \mathscr{T}_h} \left(h_K^{-d/2} \omega^K(t) \eta_1^K(t) \right). \tag{3.10}$$

These definitions allow us to introduce the *proper estimator* and the *vicinity estimator* respectively as

$$\mathscr{E}_2(t) := \left(\mathscr{E}_{2,0}(t)^2 + \mathscr{E}_{2,1}(t)^2 \right)^{1/2} \quad \text{and} \quad \mathscr{E}_\infty(t) := \mathscr{E}_{\infty,0}(t) + \mathscr{E}_{\infty,1}(t). \tag{3.11}$$

We finally introduce the *initial estimator* and *total estimator* respectively as

$$\mathscr{E}_0 := \left(\left(1 + 2\mathscr{E}_{\infty,0}(0)\right) A(0) + 2\mathscr{E}_{2,0}(0) \sqrt{A(0)} \right)^{1/2} \tag{3.12}$$

and

$$\mathscr{E}(t) := \left(\mathscr{E}_0^2 + \mathscr{E}_2(t)^2 + \mathscr{E}_\infty(t)\right)^{1/2}. \tag{3.13}$$

The motivation for our terminology will be clear in Theorem 3.6 below: there the vicinity estimator $\mathscr{E}_\infty$ does not enter directly in the conditional estimate, but dictates a "closeness condition" that must be satisfied for the estimate to hold. This conditional estimate then involves the initial and proper estimators $\mathscr{E}_0$ and $\mathscr{E}_2$.

We are now ready to state the main results whose proof is spread through §§4–6.

3.4 Theorem (Unconditional a posteriori estimate) *Let u be the solution of Problem 1.1 and u_h the finite element solution of Problem 2.6. For all $t \in [0,T]$, there exist $C = C[u_h, f, t]$ and $C' = C'[f, g, t]$ such that*

$$C \le \exp \int_0^t \left(2 \left\|\partial_t u_h(s)\right\|^2_{\mathrm{L}_\infty(\Omega)} + 4 \left\|\nabla \partial_t u_h(s)\right\|_{\mathrm{L}_\infty(\Omega)}\right) \mathrm{d}s, \tag{3.14}$$

$$C' \le \exp\left(\frac{1}{2}\left\|f(s)\right\|^2_{\mathrm{L}_\infty(\Omega)}\right)\left(\left\|Qg(0)\right\|_{\mathrm{L}_1(\Omega)} + \left\|\partial_t g\right\|_{\mathrm{L}_1(\partial\Omega\times(0,t))}\right), \tag{3.15}$$

$$\begin{aligned}\int_0^t \int_\Omega (Vu_h - Vu)^2 Qu + \frac{1}{2}\sup_{[0,t]} \int_\Omega \left|\boldsymbol{N}u_h - \boldsymbol{N}u\right|^2 Qu \\ \le C\left(\mathscr{E}_0{}^2 + 4\mathscr{E}_2(t)^2 + 8C'\mathscr{E}_\infty(t)\right).\end{aligned} \tag{3.16}$$

3.5 Remark (Poor reliability of the unconditional estimate) The estimate (3.16) holds regardless whether the approximate solution u_h is close or far from exact solution u. The presence of the vicinity estimator $\mathscr{E}_\infty$ on the right-hand side is undesirable because, even under the most optimistic assumptions of regularity on u, there is no indication that this estimator would have the same order of convergence, as h goes to zero, as the square of the geometric error on the left-hand side. In fact, the numerical tests described in §7 bear strong evidence that $\mathscr{E}_\infty$ does not decay with a sufficient high power of h. This means that the above estimate cannot be relied on absolutely as a stopping criterion in an adaptive scheme. A crucial point of this paper is that *this estimate can be improved, provided u_h is sufficiently close to u,* as stated in the next theorem.

3.6 Theorem (Conditional a posteriori estimate) *Let u be the solution of Problem 1.1 and u_h the finite element solution of Problem 2.6. For each $t \in [0,T]$, if*

$$\mathscr{E}_\infty(t) \le \frac{1}{8} \tag{3.17}$$

then there exists a constant $C = C[u_h, t]$ such that

$$C \le \exp \int_0^t \left(2 \left\|\partial_t u_h(s)\right\|^2_{\mathrm{L}_\infty(\Omega)} + 4 \left\|\nabla \partial_t u_h(s)\right\|_{\mathrm{L}_\infty(\Omega)}\right) \mathrm{d}s, \tag{3.18}$$

$$\int_0^t \int_\Omega (Vu_h - Vu)^2 Qu + \frac{1}{2}\sup_{[0,t]} \int_\Omega \left|\boldsymbol{N}u_h - \boldsymbol{N}u\right|^2 Qu \le C\left(\mathscr{E}_0{}^2 + 8\mathscr{E}_2(t)^2\right). \tag{3.19}$$

3.7 Remark (A posteriori nature of condition (3.17)**)** Theorem 3.6 is a conditional result, typical in nonlinear analysis. The condition (3.17) can be interpreted as follows: the approximate solution u_h needs to be sufficiently close to the exact solution u for the estimate to hold. The technique we use can be thought of as a linearization of the equation about u_h, instead of a linearization about u which would

be natural in an a priori setting. This leads to the important fact that *condition* (3.17) *can be effectively verified* since it involves exclusively a posteriori, therefore computable, quantities. Thus, in a practical adaptive method where a stopping criterion is needed, Theorem 3.4 would be used in the early pre-asymptotic stages in order to get close enough to the exact solution, the estimate of Theorem 3.6 would then provide a sharper criterion once the algorithm enters a second stage in which the condition (3.17) is satisfied.

4. The error equation. We divide the proof of Theorems 3.4 and 3.6 into several steps that will spread over the next two sections. Here we introduce the residual based energy technique and set-up the error equation.

4.1. The residual. The *residual* is defined as the difference between the exact operator acting on the approximate solution and the exact operator acting on the exact solution. In our setting, the result has to be understood in the following weak sense

$$\langle \mathscr{R} \,|\, \phi \rangle := \left\langle \frac{\partial_t u_h}{Q u_h} - \frac{\partial_t u}{Qu}, \phi \right\rangle + \left\langle \frac{\nabla u_h}{Q u_h} - \frac{\nabla u}{Qu}, \nabla \phi \right\rangle, \ \forall \phi \in \mathring{\mathrm{W}}{}^1_1(\Omega). \tag{4.1}$$

Here $\langle \cdot \,|\, \cdot \rangle$ stands for the duality pairing. The distribution $\mathscr{R}$ is time dependent and, owing to Hypothesis 2.1, $\mathscr{R}(t)$ is a bounded linear functional on $\mathring{\mathrm{W}}{}^1_1(\Omega)$ for all $t \in [0,T]$. We will refer to $\mathscr{R}$ as the *residual functional*. The use of (2.4) and an integration by parts in the space variable lead to the residual functional *representation*

$$\begin{aligned} \langle \mathscr{R} \,|\, \phi \rangle &= \left\langle \frac{\partial_t u_h}{Q u_h} - f - \operatorname{div}\left(\frac{\nabla u_h}{Q u_h} \right), \phi \right\rangle + \left\langle \left[\!\!\left[\frac{\nabla u_h}{Q u_h} \right]\!\!\right], \phi \right\rangle_{\Sigma_h} \\ &= \langle r, \phi \rangle + \langle j, \phi \rangle_{\Sigma_h}, \ \forall \phi \in \mathring{\mathrm{W}}{}^1_1(\Omega), \end{aligned} \tag{4.2}$$

where the residual functions r and j are those introduced in §3.1.

4.2. Galerkin orthogonality and the error equation. The starting point of our residual based a posteriori estimation, is to exploit the property that $\mathscr{R}$ vanishes on $\mathring{\mathbb{V}}_h$. This is the so called *Galerkin orthogonality* property, which yields the following *error equation* for all $\phi \in \mathring{\mathrm{W}}{}^1_1(\Omega)$, $\phi_h \in \mathring{\mathbb{V}}_h$

$$\left\langle \frac{\partial_t u_h}{Q u_h} - \frac{\partial_t u}{Qu}, \phi \right\rangle + \left\langle \frac{\nabla u_h}{Q u_h} - \frac{\nabla u}{Qu}, \nabla \phi \right\rangle = \langle \mathscr{R} \,|\, \phi - \phi_h \rangle \,. \tag{4.3}$$

4.3. Choice of the test function. The energy technique relies on an appropriate choice of test functions ϕ and ϕ_h in (4.3). Let us denote by e the error

$$e(\boldsymbol{x}, t) := u_h(\boldsymbol{x}, t) - u(\boldsymbol{x}, t) \tag{4.4}$$

and make the following choices for the test functions

$$\phi(\boldsymbol{x}, t) := \partial_t e(\boldsymbol{x}, t), \tag{4.5}$$

$$\phi_h(\boldsymbol{x}, t) := I_h \phi(\boldsymbol{x}, t) \tag{4.6}$$

where I_h is the Scott-Zhang interpolation operator which will be briefly discussed in §6.2 and §6.3. For $\partial_t e$ to be admissible as a test function ϕ in (4.3) it must vanish on $\partial\Omega$ which is not necessarily true. This motivates the following temporary assumption, which will be removed in §6.8 where we deal with general boundary data.

4.4 Hypothesis (Exact boundary data resolution) *Until §6.8, let either*
(a) the boundary value g be approximated exactly by g_h, or
(b) g be time independent.

5. Coercivity. Our objective in this section is to derive a lower bound of the left-hand side of (4.3) with the choice made in (4.5). To achieve this objective we exhibit as much coercivity as the nonlinearity allows; we will make a liberal use of the word "coercivity" in this sense. The geometric error functions of time A and B, introduced in §1.2, will be used extensively in this section and in the next one. We begin by stating some simple yet fundamental geometric relations observed by Dziuk.
5.1 Lemma (Basic geometry [8]) *Given $\boldsymbol{p}_1, \boldsymbol{p}_2 \in \mathbb{R}^d$, if $q_i := (1 + |\boldsymbol{p}_i|^2)^{1/2}$ and $\boldsymbol{n}_i := (\boldsymbol{p}_i; -1)/q_i \in \mathbb{R}^d$, for $i = 1, 2$, then the following geometric relations hold:*

$$1 - \frac{1 + \boldsymbol{p}_1 \cdot \boldsymbol{p}_2}{q_1 q_2} = \frac{1}{2} |\boldsymbol{n}_1 - \boldsymbol{n}_2|^2 , \tag{5.1}$$

$$\left| \left(\frac{1}{q_1} - \frac{1}{q_2} \right) \left(\frac{\boldsymbol{p}_1}{q_1} - \frac{\boldsymbol{p}_2}{q_2} \right) \right| \le \frac{1}{2} |\boldsymbol{n}_1 - \boldsymbol{n}_2|^2 , \tag{5.2}$$

$$\frac{|\boldsymbol{p}_1 - \boldsymbol{p}_2|}{q_1} \le (1 + |\boldsymbol{p}_2|) |\boldsymbol{n}_1 - \boldsymbol{n}_2| . \tag{5.3}$$

5.2 Lemma (Dziuk identity [8]) *If v and w are sufficiently differentiable functions on $\Omega \times [0, T]$, then*

$$\begin{aligned} \frac{1}{2} \partial_t \left(|\boldsymbol{N} v - \boldsymbol{N} w|^2 Q w \right) = & \left(\frac{\nabla v}{Q v} - \frac{\nabla w}{Q w} \right) \cdot \nabla \left(\partial_t v - \partial_t w \right) \\ & - \nabla \partial_t v \cdot \left(\frac{\nabla w}{Q v} - \frac{\nabla w}{Q w} + \frac{\nabla v}{Q v} - \frac{1 + \nabla w \cdot \nabla v}{(Q v)^2} \frac{\nabla v}{Q v} \right) \end{aligned} \tag{5.4}$$

The first term in the left-hand side of (4.3) is handled through the following inequality.
5.3 Lemma (Coercivity of the velocity term) *With the notation*

$$\varrho_1(t) := \frac{1}{2} \|\partial_t u_h(t)\|^2_{\mathrm{L}_\infty(\Omega)} , \tag{5.5}$$

we have that, for all $t \in [0, T]$,

$$\left\langle \frac{\partial_t u_h}{Q u_h} - \frac{\partial_t u}{Q u}, \partial_t u_h - \partial_t u \right\rangle \ge \frac{1}{2} \mathrm{d}_t B(t) - \varrho_1 A(t). \tag{5.6}$$

Proof Basic manipulations imply

$$\begin{aligned} & \left\langle \frac{\partial_t u_h}{Q u_h} - \frac{\partial_t u}{Q u}, \partial_t u_h - \partial_t u \right\rangle \\ & = \int_\Omega (V u_h - V u)^2 Q u + \int_\Omega \partial_t u_h \left(\frac{1}{Q u} - \frac{1}{Q u_h} \right) (V u_h - V u) Q u \\ & \ge \mathrm{d}_t B(t) - \|\partial_t u_h\|_{\mathrm{L}_\infty(\Omega)} \int_\Omega \left| \frac{1}{Q u} - \frac{1}{Q u_h} \right| \sqrt{Q u} \, |V u_h - V u| \sqrt{Q u} \\ & \ge \mathrm{d}_t B(t) - \|\partial_t u_h\|_{\mathrm{L}_\infty(\Omega)} \left(\int_\Omega |\boldsymbol{N} u_h - \boldsymbol{N} u|^2 Q u \right)^{1/2} \left(\int_\Omega (V u_h - V u)^2 Q u \right)^{1/2} . \end{aligned}$$

Consequently

$$\left\langle \frac{\partial_t u_h}{Qu_h} - \frac{\partial_t u}{Qu}, \partial_t u_h - \partial_t u \right\rangle \geq \mathrm{d}_t B(t) - \frac{1}{2}\,\mathrm{d}_t B(t) - \varrho_1(t)A(t) = \frac{1}{2}\,\mathrm{d}_t B(t) - \varrho_1 A(t),$$

as asserted. ∎

5.4 Lemma (Coercivity for normals and gradients) *With the notation*

$$\varrho_2(t) := \|\nabla \partial_t u_h(t)\|_{\mathrm{L}_\infty(\Omega)}\,, \tag{5.7}$$

we have that, for all $t \in [0,T]$,

$$\left\langle \frac{\nabla u_h}{Qu_h} - \frac{\nabla u}{Qu}, \nabla(\partial_t u_h - \partial_t u) \right\rangle \geq \frac{1}{2}\,\mathrm{d}_t A(t) - \varrho_2(t)A(t). \tag{5.8}$$

Proof Integrating in space both sides of (5.4) and rearranging terms yield

$$\begin{aligned}\left\langle \frac{\nabla u_h}{Qu_h} - \frac{\nabla u}{Qu}, \nabla(\partial_t u_h - \partial_t u) \right\rangle &= \frac{1}{2}\,\mathrm{d}_t A(t) \\ &\quad + \int_\Omega \nabla \partial_t u_h \cdot \left(\frac{\nabla u}{Qu_h} - \frac{\nabla u}{Qu} - \frac{\nabla u_h}{Qu_h} - \frac{1 + \nabla u \cdot \nabla u_h}{(Qu_h)^2} \frac{\nabla u_h}{Qu_h} \right).\end{aligned}$$

To show the result it is sufficient to show that the last integral above is bounded from below by $-\varrho_2(t)A(t)$. To do this we add and subtract $-(Qu\nabla u_h)/(Qu_h)^2$ and rewrite this term as the sum of two integrals

$$\begin{aligned} I_1 + I_2 := &\int_\Omega \nabla \partial_t u_h \cdot \left(\frac{\nabla u}{Qu_h} - \frac{\nabla u}{Qu} + \frac{\nabla u_h}{Qu_h} - \frac{Qu\nabla u_h}{(Qu_h)^2} \right) \\ &+ \int_\Omega \nabla \partial_t u_h \cdot \left(\frac{Qu\nabla u_h}{(Qu_h)^2} - \frac{1 + \nabla u \cdot \nabla u_h}{(Qu_h)^2} \frac{\nabla u_h}{Qu_h} \right).\end{aligned} \tag{5.9}$$

The integrals in (5.9) are bounded, by using (5.2) for the first one

$$I_1 \geq -\,\|\nabla \partial_t u_h\|_{\mathrm{L}_\infty(\Omega)} \int_\Omega \left| \left(\frac{1}{Qu} - \frac{1}{Qu_h} \right) \left(\frac{\nabla u_h}{Qu_h} - \frac{\nabla u}{Qu} \right) Qu \right| \geq -\frac{\varrho_2(t)}{2} A(t),$$

and with the help of (5.1) for the second one

$$I_2 \geq -\,\|\nabla \partial_t u_h\|_{\mathrm{L}_\infty(\Omega)} \int_\Omega \left| \left(1 - \frac{1 + \nabla u_h \cdot \nabla u}{Qu_h Qu} \right) Qu \frac{\nabla u_h}{(Qu_h)^2} \right| \geq -\frac{\varrho_2(t)}{2} A(t).$$

This proves the assertion. ∎

5.5 Lemma (Estimate of the geometric terms) *With the notation*

$$\varrho(t) := \varrho_1(t) + \varrho_2(t), \tag{5.10}$$

we have that, for all $t \in [0,T]$,

$$A(t) + B(t) \leq A(0) + 2\int_0^t \varrho(s)A(s)\,\mathrm{d}s + 2\int_0^t \langle \mathscr{R}(s) \,|\, \partial_t \left(e(s) - I_h e(s) \right) \rangle\,\mathrm{d}s. \tag{5.11}$$

Proof Using (4.5), (4.6), (5.6) and (5.8) in (4.3), we obtain

$$\frac{1}{2}\left(\mathrm{d}_t A(t) + \mathrm{d}_t B(t) \right) \leq \langle \mathscr{R} \,|\, \partial_t e(t) - I_h \partial_t e(t) \rangle + \varrho(t)A(t), \tag{5.12}$$

for all $t \in [0,T]$. An integration in time over the interval $[0,t]$ yields the result. ∎

6. Bounding the residual by the estimators. We prove in this section Theorems 3.4 and 3.6 by estimating $\int_0^t \langle \mathscr{R} \,|\, \partial_t (e - I_h e) \rangle$ appearing in (5.11). We will denote by $d' = d/(d-1)$ the conjugate exponent of d, the latter being the surface's dimension. We start by stating two lemmas bearing a fundamental geometric relationship and an interpolation theory result respectively.

6.1 Lemma (Fierro-Veeser inequality [12]) *Adopting the same notation as in Lemma 5.1, the following inequality holds:*

$$|\boldsymbol{p}_1 - \boldsymbol{p}_2| \, \frac{1}{q_1^2} \leq 2 \, |\boldsymbol{n}_1 - \boldsymbol{n}_2| + |\boldsymbol{n}_1 - \boldsymbol{n}_2|^2 \, q_2. \tag{6.1}$$

6.2 Lemma (The Scott-Zhang interpolation [21]) *If I_h denotes the averaging interpolation operator introduced by Scott & Zhang, then the following interpolation inequalities hold:*

$$\|\psi - I_h \psi\|_{\mathrm{L}_{d'}(K)} \leq C_1 \, |\psi|_{\mathrm{W}^1_1(\mathscr{U}_K^h)} \,, \tag{6.2}$$

$$\|\psi - I_h \psi\|_{\mathrm{L}_1(\partial K)} \leq 2 C_2 \, |\psi|_{\mathrm{W}^1_1(\mathscr{U}_K^h)} \,, \tag{6.3}$$

where $\mathscr{U}_K^h$ is the $\mathscr{T}_h$-neighborhood of K defined in (2.13).

6.3 Remark The particular choice of the norms in Lemma 6.2 is motivated by our wish for $\sqrt{A(t)}$ to appear in an upper bound of the right-hand side of (5.11).

Indeed, estimating the residual $\mathscr{R}$ in energy norms would typically lead to dealing with $|\nabla u_h - \nabla u|$. In light of the geometric errors A and B in the left-hand side of (5.11), a straightforward idea would be to bound its L_2-norm, that is $|\nabla u_h - \nabla u|^2$, from above by $C \, |\boldsymbol{N} u_h - \boldsymbol{N} u|^2 \, Q u$, with the constant $C = C[u_h]$ independent of u (think of u_h being unrelated to u in this paragraph). The only practical way to derive such a bound would be a pointwise geometric relation like

$$\frac{|\boldsymbol{p}_1 - \boldsymbol{p}_2|^2}{\kappa(\boldsymbol{p}_1) \, |\boldsymbol{n}_1 - \boldsymbol{n}_2|^2 \, q_2} \leq 1, \tag{6.4}$$

where $p_1 = \nabla u_h$, $\boldsymbol{n}_1 = N u_h$, $q_1 = Q u_h$, the quantities with subscript 2 refer to u and κ is some function of $\boldsymbol{p}_1$ only. Unfortunately this is not possible because (6.4) is false. To see this, fix $\boldsymbol{p}_1$ and observe that $\boldsymbol{n}_1 - \boldsymbol{n}_2$ is bounded; by letting $|\boldsymbol{p}_2| \to \infty$ we obtain, in contrast with (6.4),

$$\frac{|\boldsymbol{p}_1 - \boldsymbol{p}_2|^2}{\kappa(\boldsymbol{p}_1) \, |\boldsymbol{n}_1 - \boldsymbol{n}_2|^2 \, q_2} \geq C \frac{|\boldsymbol{p}_1 - \boldsymbol{p}_2|^2}{q_2} = \mathrm{O}(|\boldsymbol{p}_2|) \to \infty. \tag{6.5}$$

This difficulty can be circumvented by using the L_1-norm of $|\nabla u - \nabla u_h|$, instead of the L_2-norm, and the Fierro-Veeser inequality (6.1) which reads

$$\begin{aligned} |\nabla u_h - \nabla u| &= (Q u_h)^2 \, \frac{|\nabla u_h - \nabla u|}{(Q u_h)^2} \\ &\leq (Q u_h)^2 \left(2 \, |\boldsymbol{N} u_h - \boldsymbol{N} u| + |\boldsymbol{N} u_h - \boldsymbol{N} u|^2 \, Q u \right). \end{aligned} \tag{6.6}$$

Notice that the last term is cumbersome because its power is too high—it is the "price to pay". This term will yield a term of the form $\varrho(t) A(t)$ on the right-hand side which has to be handled carefully in order to close the estimate.

Recalling first the notation in §1.2 and §3.3, we now state and prove the central result of this paper.

6.4 Lemma (Residual estimate) *The following inequality holds for all* $t \in [0,T]$

$$\begin{aligned} A(t) + B(t) &\le \mathscr{E}_0{}^2 + 2\hat{\mathscr{E}}_{2,0}(t)A(t)^{1/2} + 2\hat{\mathscr{E}}_{\infty,0}(t)A(t) \\ &\quad + 2\int_0^t \dot{\mathscr{E}}_{2,1}(s)A(s)^{1/2}\,\mathrm{d}s + 2\int_0^t \dot{\mathscr{E}}_{\infty,1}(s)A(s)\,\mathrm{d}s + 2\int_0^t \varrho(s)A(s)\,\mathrm{d}s. \end{aligned} \tag{6.7}$$

Proof Apply the representation formula (4.2) with $\phi = \partial_t \delta_h e$, where $\delta_h e(t) := e(t) - I_h e(t)$, integrate by parts in time, and use the commutativity property $\partial_t I_h = I_h \partial_t$, to obtain

$$\begin{aligned} \int_0^t \langle \mathscr{R}(s) \,|\, \partial_t \delta_h e(s)\rangle \,\mathrm{d}s &= \int_0^t \langle r(s), \partial_t \delta_h e(s)\rangle + \langle j(s), \partial_t \delta_h e(s)\rangle_{\Sigma_h} \,\mathrm{d}s \\ &= \big[\langle r, \delta_h e\rangle + \langle j, \delta_h e\rangle_{\Sigma_h}\big]_0^t - \int_0^t \langle \partial_t r(s), \delta_h e(s)\rangle + \langle \partial_t j(s), \delta_h e(s)\rangle_{\Sigma_h} \,\mathrm{d}s. \end{aligned}$$

Hence

$$\begin{aligned} &\int_0^t \langle \mathscr{R}(s) \,|\, \partial_t \delta_h e(s)\rangle \,\mathrm{d}s \le \\ &\quad \sum_{K\in\mathscr{T}_h} \Bigg(\sum_{s\in\{0,t\}} \Big[\|r(s)\|_{\mathrm{L}_d(K)} \|\delta_h e(s)\|_{\mathrm{L}_{d'}(K)} + \frac{1}{2} \|j(s)\|_{\mathrm{L}_\infty(\partial K)} \|\delta_h e(s)\|_{\mathrm{L}_1(\partial K)} \Big] \\ &\qquad + \int_0^t \|\partial_t r(s)\|_{\mathrm{L}_d(K)} \|\delta_h e(s)\|_{\mathrm{L}_{d'}(K)} + \frac{1}{2} \|\partial_t j(s)\|_{\mathrm{L}_\infty(\partial K)} \|\delta_h e(s)\|_{\mathrm{L}_1(\partial K)} \,\mathrm{d}s \Bigg). \end{aligned}$$

Owing to the approximation properties of the Scott-Zhang interpolator of Lemma 6.2, and using the local indicators η_i^K introduced in Definition 3.2, we may write

$$\begin{aligned} \int_0^t \langle \mathscr{R}(s) \,|\, \partial_t \delta_h e(s)\rangle \,\mathrm{d}s &\le \sum_{K\in\mathscr{T}_h} h_K^{-d/2} \Big(\eta_0^K(0) \|\nabla e(0)\|_{\mathrm{L}_1(\mathscr{U}_K^h)} \\ &\qquad + \eta_0^K(t) \|\nabla e(t)\|_{\mathrm{L}_1(\mathscr{U}_K^h)} + \int_0^t \eta_1^K(s) \|\nabla e(s)\|_{\mathrm{L}_1(\mathscr{U}_K^h)} \,\mathrm{d}s \Big). \end{aligned} \tag{6.8}$$

We proceed by observing that inequality (6.6) implies

$$\|\nabla e(t)\|_{\mathrm{L}_1(\mathscr{U}_K^h)} \le \sup_{\mathscr{U}_K^h} (Qu_h)^2 \int_{\mathscr{U}_K^h} \left(\frac{2\mathscr{N}(t)}{\sqrt{Qu(t)}} + \mathscr{N}(t)^2 \right), \tag{6.9}$$

where, in order to simplify notation, we introduce the shorthand

$$\mathscr{N} := |\boldsymbol{N}u_h - \boldsymbol{N}u| \sqrt{Qu}. \tag{6.10}$$

We follow up the bound in (6.8) by using (6.9) as follows:

$$\begin{aligned} &\int_0^t \langle \mathscr{R}(s) \,|\, \partial_t e(s) - I_h \partial_t e(s)\rangle \,\mathrm{d}s \\ &\quad \le \sum_{K\in\mathscr{T}_h} \eta_0^K(0) h_K^{-d/2} \omega^K(0) \int_{\mathscr{U}_K^h} \left(\frac{2\mathscr{N}(0)}{\sqrt{Qu(0)}} + \mathscr{N}(0)^2 \right) \\ &\quad + \sum_{K\in\mathscr{T}_h} \eta_0^K(t) h_K^{-d/2} \omega^K(t) \int_{\mathscr{U}_K^h} \left(\frac{2\mathscr{N}(t)}{\sqrt{Qu(t)}} + \mathscr{N}(t)^2 \right) \\ &\quad + \sum_{K\in\mathscr{T}_h} \int_0^t \eta_1^K(s) h_K^{-d/2} \omega^K(s) \int_{\mathscr{U}_K^h} \left(\frac{2\mathscr{N}(s)}{\sqrt{Qu(s)}} + \mathscr{N}(s)^2 \right) \mathrm{d}s. \end{aligned} \tag{6.11}$$

The first two terms in (6.11) can be bounded at once through the following inequality (where we simply take $t = 0$ for the first term):

$$\begin{aligned}
&\sum_{K\in\mathscr{T}_h} \eta_0^K(t) h_K^{-d/2}\omega^K(t) \int_{\mathscr{U}_K^h} \left(\frac{2\mathscr{N}(t)}{\sqrt{Qu(t)}} + \mathscr{N}(t)^2 \right) \\
&\quad \le 2 \left(\sum_{K\in\mathscr{T}_h} \eta_0^K(t)^2 h_K^{-d} \omega^K(t)^2 \left|\mathscr{U}_K^h\right| \sup_{\mathscr{U}_K^h} \frac{1}{Qu(t)} \right)^{1/2} \left(\sum_{K\in\mathscr{T}_h} \int_{\mathscr{U}_K^h} \mathscr{N}(t)^2 \right)^{1/2} \\
&\qquad + \max_{K\in\mathscr{T}_h} \left(\eta_0^K(t) h_K^{-d/2} \omega^K(t) \right) \left(\sum_{K\in\mathscr{T}_h} \int_{\mathscr{U}_K^h} \mathscr{N}(t)^2 \right).
\end{aligned}$$

Likewise, the last term in (6.11) is bounded by

$$\begin{aligned}
\int_0^t \Bigg[&2 \left(\sum_{K\in\mathscr{T}_h} \eta_1^K(s)^2 h_K^{-d} \omega^K(s)^2 \left|\mathscr{U}_K^h\right| \sup_{\mathscr{U}_K^h} \frac{1}{Qu(s)} \right)^{1/2} \left(\sum_{K\in\mathscr{T}_h} \int_{\mathscr{U}_K^h} \mathscr{N}(s)^2 \right)^{1/2} \\
&+ \max_{K\in\mathscr{T}_h} \left(\eta_1^K(s) h_K^{-d/2} \omega^K(s) \right) \left(\sum_{K\in\mathscr{T}_h} \int_{\mathscr{U}_K^h} \mathscr{N}(s)^2 \right) \Bigg] \mathrm{d}s.
\end{aligned}$$

To conclude the proof, we observe that the shape regularity of $\mathscr{T}_h$ (2.12) implies the existence of two constants $\gamma \in \mathbb{R}^+$ and $M \in \mathbb{Z}^+$, depending only on σ_0, such that $\left|\mathscr{U}_K^h\right| \le (\gamma^2/4) h_K^d$ and the number of simplexes of $\mathscr{T}_h$ contained in $\mathscr{U}_K^h$ does not exceed M. It follows that

$$\begin{aligned}
&\int_0^t \langle \mathscr{R}(s) \,|\, \partial_t \delta_h e(s) \rangle \, \mathrm{d}s \\
&\le \gamma \left(\sum_{K\in\mathscr{T}_h} \alpha^K(0) \eta_0^K(0)^2 \right)^{1/2} A(0)^{1/2} + M \max_{K\in\mathscr{T}_h} \left(h_K^{-d/2} \omega^K(0) \eta_0^K(0) \right) A(0) \\
&+ \gamma \left(\sum_{K\in\mathscr{T}_h} \alpha^K(t) \eta_0^K(t)^2 \right)^{1/2} A(t)^{1/2} + M \max_{K\in\mathscr{T}_h} \left(h_K^{-d/2} \omega^K(t) \eta_0^K(t) \right) A(t) \\
&+ \gamma \int_0^t \left(\sum_{K\in\mathscr{T}_h} \alpha^K(s) \eta_1^K(s)^2 \right)^{1/2} A(s)^{1/2} \, \mathrm{d}s + M \int_0^t \max_{K\in\mathscr{T}_h} \left(h_K^{-d/2} \omega^K(s) \eta_1^K(s) \right) A(s) \, \mathrm{d}s.
\end{aligned}$$

Recalling Definition 3.3, we combine the last inequality with (5.11) and obtain (6.7) as asserted. ∎

Next we prove the theorems stated in §3 with the aid of Lemma 6.4. For (6.7) to be useful we must control the terms containing $A(t)$ on the right-hand side by those on the left-hand side. We distinguish two main ways of doing this. The first way, which is direct and somewhat naive, uses the stability Lemma 2.4 and leads to the unconditional a posteriori estimate in Theorem 3.4. The second, more careful, way results in the conditional but sharper estimate in Theorem 3.6. To shorten the discussion, we show first the latter and then the former which is simpler.

6.5. Proof of Theorem 3.6. Our starting point is inequality (6.7). Introduce the notation $A^*(t) := \sup_{[0,t]} A$, apply Hölder inequality, and Young inequality with pa-

rameter μ at our disposal, to obtain

$$\begin{aligned} A(t) + B(t) \leq & \mathscr{E}_0{}^2 + \mu A(t) + \frac{1}{\mu}\hat{\mathscr{E}}_{2,0}(t)^2 + \mu A^*(t) + \frac{1}{\mu}\mathscr{E}_{2,1}(t)^2 \\ & + 2\hat{\mathscr{E}}_{\infty,0}(t)A(t) + 2A^*(t)\mathscr{E}_{\infty,1}(t) + 2\int_0^t \varrho(s)A(s)\,\mathrm{d}s. \end{aligned} \tag{6.12}$$

Choosing $\mu = 1/8$, taking the supremum over $[0,t]$ on both sides, recalling that B, $\mathscr{E}_{2,1}$ and $\mathscr{E}_{\infty,1}$ are nondecreasing, and using Definition 3.3, we can write

$$A^*(t) + B(t) \leq \mathscr{E}_0^2 + \frac{1}{4}A^*(t) + 8\mathscr{E}_2(t)^2 + 2\mathscr{E}_\infty A^*(t) + 2\int_0^t \varrho(s)A^*(s)\,\mathrm{d}s. \tag{6.13}$$

The condition (3.17), i.e., $\mathscr{E}_\infty \leq 1/8$, and the last inequality imply

$$\frac{1}{2}A^*(t) + B(t) \leq \mathscr{E}_0{}^2 + 8\mathscr{E}_2(t)^2 + 2\int_0^t \varrho(s)A^*(s)\,\mathrm{d}s. \tag{6.14}$$

To conclude the proof, it suffices now to apply Gronwall lemma in the above inequality, and to recall (5.10), (5.5) and (5.7), in order to derive (3.18) and (3.19). ∎

6.6. Proof of Theorem 3.4. The proof is a direct combination of Lemma 6.4 and the elementary fact that

$$A(t) = \int_\Omega |\boldsymbol{N}u_h(t) - \boldsymbol{N}u(t)|^2\, Qu(t) \leq 4\int_\Omega Qu(t). \tag{6.15}$$

a The stability Theorem 2.4 provides us with an upper bound for the last integral in terms of the data f and g. To conclude, it is enough to proceed along the lines of §6.5 with $\mu = 1/4$ and apply Gronwall lemma. ∎

6.7 Remark (Slowly varying solutions) Notice that if $\int_0^t \varrho$ is small enough (for which it is necessary for $\|\partial_t u_h\|_{\mathrm{L}_1(\mathrm{W}^1_\infty)}$ to be small), the Gronwall inequality argument is not needed and the exponential bound on C can be dropped. This is in particular true for solutions that are close to stationary points, i.e., if $\partial_t f$ and $\partial_t g$ are very small. We will not pursue this issue further in this paper, but we remark that this condition is also a posteriori and could be checked automatically if needed.

6.8. Time dependent Dirichlet boundary data. As promised earlier, we remove now Hypothesis 4.4; that is, we allow

$$\partial_t(u_h - u)|_{\partial\Omega} = \partial_t(g_h - g) \neq 0. \tag{6.16}$$

We study the case where the boundary value g is discretized as follows

$$\tilde{g}_h := I_h\tilde{g} \text{ and } g_h := \tilde{g}_h|_{\partial\Omega} \tag{6.17}$$

where I_h it the Scott-Zhang interpolator of Lemma 6.2 and $\tilde{g}$ denotes the extension of g to the whole domain Ω [21, (5.5)]. The error $e = u_h - u$ can thus be decomposed as follows

$$e = e_0 + \epsilon := (u_h - \tilde{g}_h - u + \tilde{g}) + (\tilde{g}_h - \tilde{g}). \tag{6.18}$$

The residual $\mathscr{R}$, as defined in (4.1), can be naturally extended to be a functional on $\mathrm{W}^1_1(\Omega)$. It follows that if we take $\phi = \partial_t e$ in (4.3) we have

$$\langle \mathscr{R} \,|\, \partial_t e\rangle = \langle \mathscr{R} \,|\, \partial_t e_0\rangle + \langle \mathscr{R} \,|\, \partial_t \epsilon\rangle = \langle \mathscr{R} \,|\, \partial_t e_0 - I_h\partial_t e_0\rangle + \langle \mathscr{R} \,|\, \partial_t \epsilon\rangle\,. \tag{6.19}$$

Notice that a Galerkin orthogonality argument can be applied directly to the part with the admissible error $e_0 \in \mathring{\mathrm{W}}^1_1$. As for the last term in (6.19), we use the $\mathbb{V}_h$-invariance property the Scott-Zhang interpolator I_h, namely $I_h I_h \psi = I_h \psi$ for all $\psi \in \mathrm{W}^1_1(\Omega)$, and (6.17) to conclude that

$$I_h \epsilon = I_h \tilde{g}_h - I_h \tilde{g} = I_h I_h \tilde{g} - I_h \tilde{g} = 0.$$

This implies that $\partial_t I_h \epsilon = 0$, and thus $\langle \mathscr{R} \,|\, \partial_t \epsilon \rangle = \langle \mathscr{R} \,|\, \partial_t (\epsilon - I_h \epsilon) \rangle$, whence the following representation formula follows from (6.19) and element-wise integration by parts:

$$\langle \mathscr{R} \,|\, \partial_t e \rangle = \langle r, \partial_t e - I_h \partial_t e \rangle + \langle j, \partial_t e - I_h \partial_t e \rangle + \langle \beta - \beta_h, \partial_t \epsilon \rangle_{\partial\Omega} \,. \tag{6.20}$$

where $\beta := (\nabla u \cdot \boldsymbol{\nu})/Qu$ and $\beta_h = (\nabla u_h \cdot \boldsymbol{\nu})/Qu_h$.

In order to obtain a lower bound of the left-hand side of equation (6.20), which is equal to the left-hand side of (4.3) with $\phi = \partial_t e$,we proceed in the same fashion as in §5 and thereby we derive again (5.11). The first two terms on the right-hand side of (6.20) can be dealt with exactly as in §6; while the fact that $\beta, \beta_h \leq 1$ implies the following bound for the last term

$$\langle \beta - \beta_h, \partial_t \epsilon \rangle_{\partial\Omega} \leq 2 \left\| \partial_t \epsilon \right\|_{\mathrm{L}_1(\partial\Omega)} = 2 \left\| \partial_t g - \partial_t g_h \right\|_{\mathrm{L}_1(\partial\Omega)} . \tag{6.21}$$

This proves the following generalization of Lemma (6.4).

6.9 Lemma (Residual estimate with boundary values) *With the notation* $\mathscr{E}_\partial(t) := \int_0^t \|\partial_t (g - g_h)\|_{\mathrm{L}_1(\partial\Omega)}$, *we have that, for all* $t \in [0, T]$,

$$\begin{aligned} A(t) + B(t) \leq {\mathscr{E}_0}^2 + 2\mathscr{E}_\partial(t) + 2\hat{\mathscr{E}}_{2,0}(t) A(t)^{1/2} + 2\hat{\mathscr{E}}_{\infty,0} A(t) \\ + 2\int_0^t \dot{\mathscr{E}}_{2,1}(s) A(s)^{1/2} \,\mathrm{d}s + 2\int_0^t \dot{\mathscr{E}}_{\infty,1}(s) A(s) \,\mathrm{d}s + 2\int_0^t \varrho(s) A(s) \,\mathrm{d}s. \end{aligned} \tag{6.22}$$

This lemma enables us to obtain extended versions of Theorems 3.6 and 3.4 by just adding $\mathscr{E}_\partial$ to the estimators therein. We omit the statement of these results as they can be written in a straightforward manner.

7. Numerical results and reliability tests. We present now some numerical computations that we have performed in order to test the reliability of the error estimates derived in Theorem 3.6 and Theorem 3.4.

7.1 Definition (Fully discrete semi-implicit scheme [7, 8]) Let $N \in \mathbb{Z}^+$ and $0 = t_0 < t_1 < \cdots < t_N = T$ be a partition of the time interval $[0, T]$. For each $n \in [1 : N]$, denote by $\tau_n := t_n - t_{n-1}$ the n-th step size. Given $g_h(0) \in \mathbb{V}_h$ (an interpolant of $g(0)$), and $\tilde{g}^n_h$ (the extension to Ω of an interpolant of $g(t_n)$), find a sequence of functions $U^n_h \in \mathbb{V}_h$ such that, for each $n \in [1 : N]$,

$$\left\langle \frac{\nabla U^n_h}{Q U^{n-1}_h}, \nabla \phi_h \right\rangle + \left\langle \frac{U^n_h}{\tau_n Q U^{n-1}_h}, \phi_h \right\rangle = \left\langle \frac{U^{n-1}_h}{\tau_n Q U^{n-1}_h} + f^n, \phi_h \right\rangle, \ \forall \phi_h \in \mathring{\mathbb{V}}_h; \tag{7.1}$$

$$U^n_h - \tilde{g}^n_h \in \mathring{\mathbb{V}}_h. \tag{7.2}$$

We implemented this scheme, which is due to Dziuk [8], with the help of the C finite element toolbox ALBERT of Schmidt & Siebert [20]. All the computations are based on piecewise linear ($\mathbb{P}^1$) finite elements.

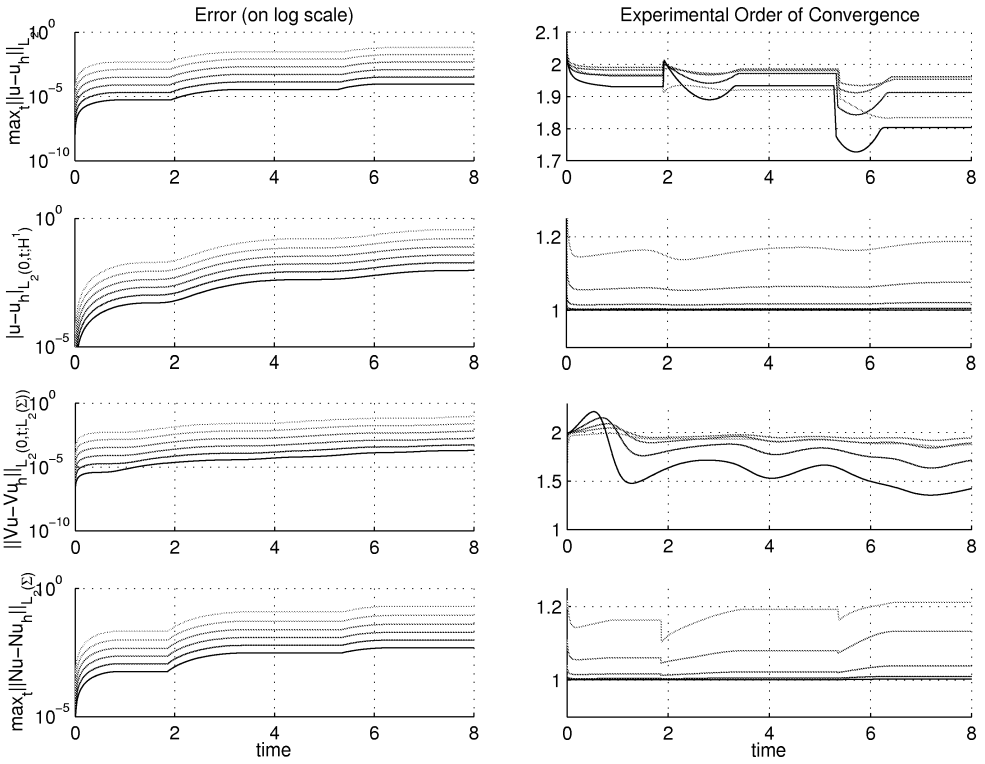

FIG. 7.1. *Errors and experimental orders of convergence (EOC) vs. time of Example 7.5. On the left column, we plot the errors in the customary Sobolev norms, related to the heat equation, and the geometric energy error and normal velocity error introduced in §1.2. On the right column, we plot the corresponding EOC. The different gray shades (or colors), from light to dark, correspond to the decreasing meshsizes h. Notice that the behavior of the Sobolev energy norm error $|e|_{\mathrm{L}_2(\mathrm{H}^1_0)}$ and the geometric energy error is similar and that both have EOC close to* 1.

7.2. Main goal of the numerical results. With reference to Definition 3.3, we introduce the *full proper estimator* defined as $\widetilde{\mathscr{E}} := \left(\mathscr{E}_0{}^2 + \mathscr{E}_2{}^2\right)^{1/2}$, and we recall that we denote by $\mathscr{E}_\infty$ the *vicinity estimator*, by $\mathscr{E}$ the *total estimator* and by E the *geometric error*, introduced in (1.9). With this notation the *unconditional estimate* of Theorem 3.4 can be written as

$$E \leq C\mathscr{E} = C\left(\widetilde{\mathscr{E}}^2 + C'\mathscr{E}_\infty\right)^{1/2}, \tag{7.3}$$

while the *conditional estimate* provided by Theorem 3.6 can be summarized as follows

$$\mathscr{E}_\infty \leq c \quad \Rightarrow \quad E \leq \widetilde{C}\widetilde{\mathscr{E}}. \tag{7.4}$$

The main goal of our numerical experiments is to see that the error bound of (7.4) is reliable whereas that of (7.3) is not. This will be illustrated by comparing the *experimental order of convergence (EOC for short)* of E, $\widetilde{\mathscr{E}}$, and $\mathscr{E}_\infty^{1/2}$. The EOC is defined as follows: for a given finite sequence of uniform triangulations $\{\mathscr{T}_{h_i}\}_{i=1,\dots,I}$ of meshsize h_i, the EOC of a corresponding sequence of some triangulation-dependent

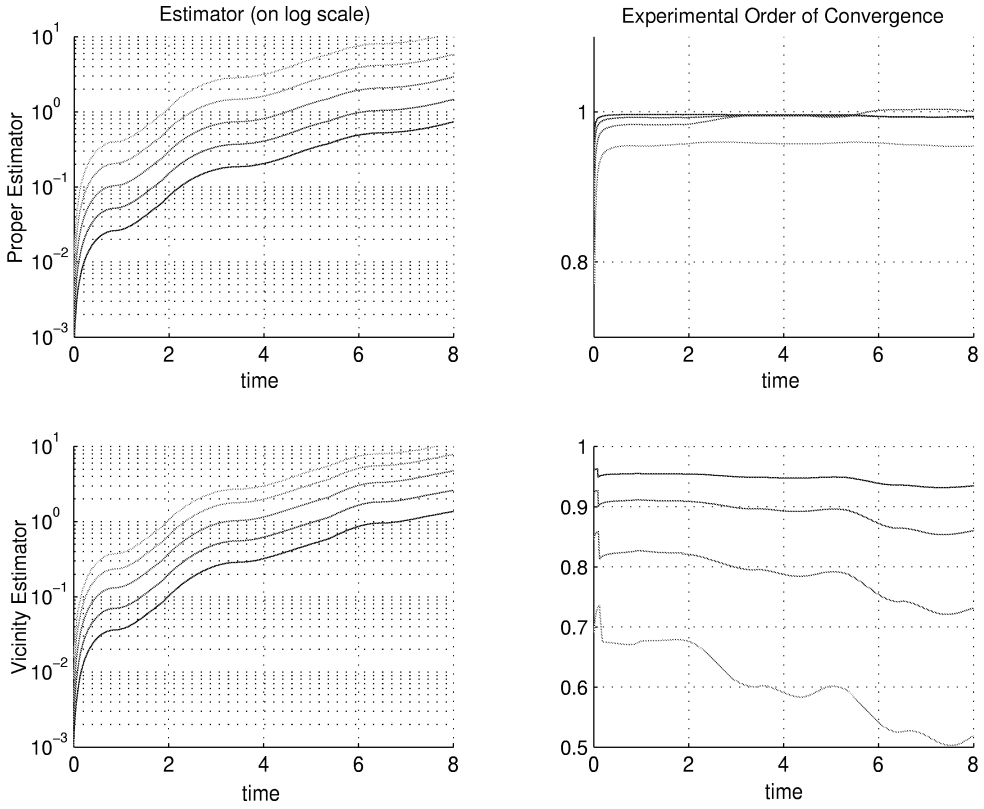

FIG. 7.2. *Estimators and experimental orders of convergence (EOC) vs. time of Example 7.5, for $i \geq 2$. Above we show the proper estimator $\widetilde{\mathscr{E}}$ and the corresponding EOC (notice that in this particular figure we do not plot the proper estimator for the first meshsize $h = 0.5$ for clarity reasons). Below we exhibit the vicinity estimator behavior which is seen to converge to zero. According to Figure 7.1, we have* $\mathrm{EOC}\, E \approx 1$ *and, according to this figure, we have* $\mathrm{EOC}\, \widetilde{\mathscr{E}} \approx 1$ *which means that the proper estimator is reliable. Notice also that, for the vicinity estimator, we have* $\mathrm{EOC}\, \mathscr{E}_\infty \approx 0.95 \leq 1$*, which implies that* $\mathrm{EOC}\, \mathscr{E} \leq 1/2$*: this is a strong numerical evidence that the unconditional estimate* (7.3) *is not reliable, in that the estimators decay with a much lower order than the errors, and justifies the need for the sharper conditional result of Theorem 3.6.*

quantity $e(i)$ (like an error or an estimator), is itself a sequence defined as

$$\mathrm{EOC}\, e(i) = \frac{\log(e(i+1)/e(i))}{\log(h_{i+1}/h_i)}. \tag{7.5}$$

Notice that for (7.4) to be reliable, it is sufficient to have $\mathrm{EOC}\, E \approx \mathrm{EOC}\, \widetilde{\mathscr{E}}$ and $\mathscr{E}_\infty = \mathrm{o}(1)$, as i increases—this will be satisfied in our numerical tests—whereas for (7.3) to be reliable it is necessary to have the stronger requirement that $\mathrm{EOC}\, E \approx \mathrm{EOC}\, \widetilde{\mathscr{E}}$ and $\mathrm{EOC}\, E \approx \mathrm{EOC}\, \mathscr{E}_\infty^{1/2}$—this will fail in our numerical tests. We will focus also on understanding when $\mathscr{E}_\infty = \mathrm{o}(1)$ might fail and on computing the *effectivity index*, which is a practical bound of the constant $\widetilde{C}$, and is defined as $E/\widetilde{\mathscr{E}}$, at the finest level I.

Notice that since the errors and the estimators are functions of time, also the EOC and the effectivity index are functions of time. Most of the comments in this

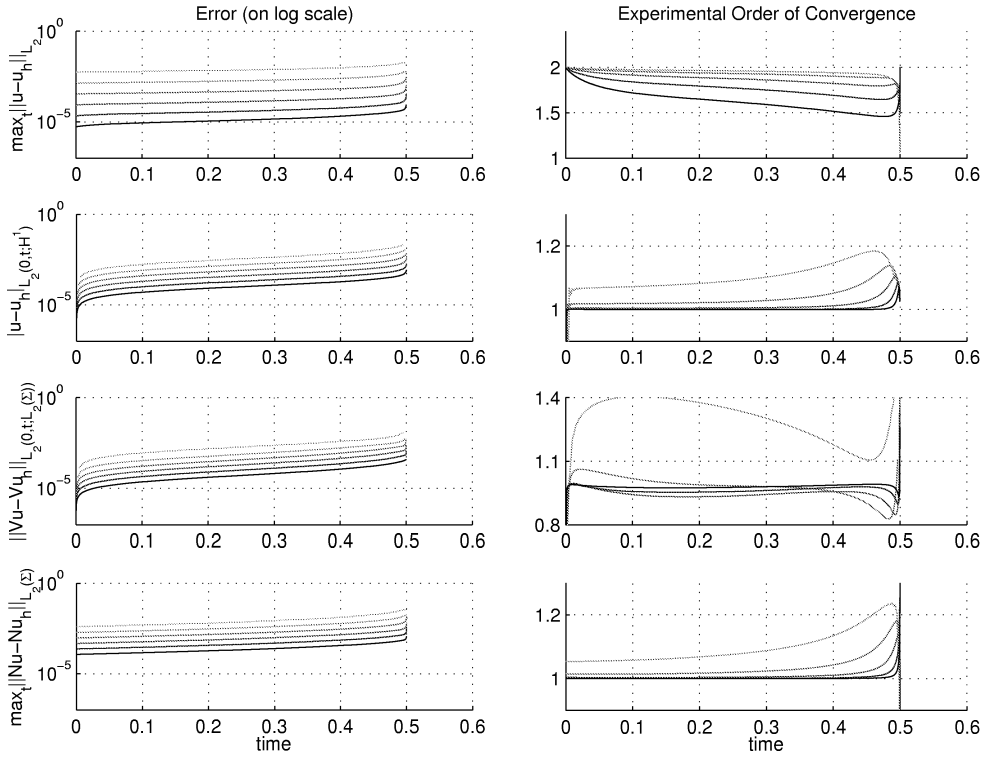

FIG. 7.3. *Sobolev norm errors and geometric errors for the shrinking sphere of Example 7.6. The different shades of gray (or colors) correspond to decreasing meshsize h. In this example a blow-up in the gradient occurs at the boundary at time $t = 0.5$.*

section are given as figure captions in order to make the reading easier.

7.3 Remark (Practical version of the error estimators) To test the reliability of the upper bound given by the estimators, we compute a fully discrete version of the spatially discrete global estimators introduced in Definition 3.3. These estimators are sums of the local indicators

$$\eta_i^K(t) := h_K^{d/2} \left(C_1 \left\| (\partial_t)^i r(t) \right\|_{\mathrm{L}_d(K)} + C_2 \left\| (\partial_t)^i j(t) \right\|_{\mathrm{L}_\infty(\partial K)} \right), \quad i = 0, 1, \tag{7.6}$$

which involve the L_∞ norm that is not so practical. Since we use piecewise linear elements, the jump residuals are constant functions on each edge, and thus the L_∞ norm can be replaced by the L_2 norm using the inverse estimate

$$\|v\|_{\mathrm{L}_\infty(\partial K)} \leq C h_K^{(1-d)/2} \|v\|_{\mathrm{L}_2(\partial K)}, \tag{7.7}$$

for all v that are constants on each edge of ∂K. It is hence legitimate to use instead of η_K^i, the handier local indicators

$$\bar{\eta}_i^K(t) := h_K^{d/2} C_1 \left\| (\partial_t)^i r(t) \right\|_{\mathrm{L}_d(K)} + h_K^{1/2} C_2 \left\| (\partial_t)^i j(t) \right\|_{\mathrm{L}_2(\partial K)}, \quad i = 0, 1. \tag{7.8}$$

All the integrals are in fact quadratures: while ALBERT's built-in Gaussian quadrature is used to approximate the space integrals, a simple midpoint rule is used for the time integrals. Time derivatives are replaced by backward finite differences.

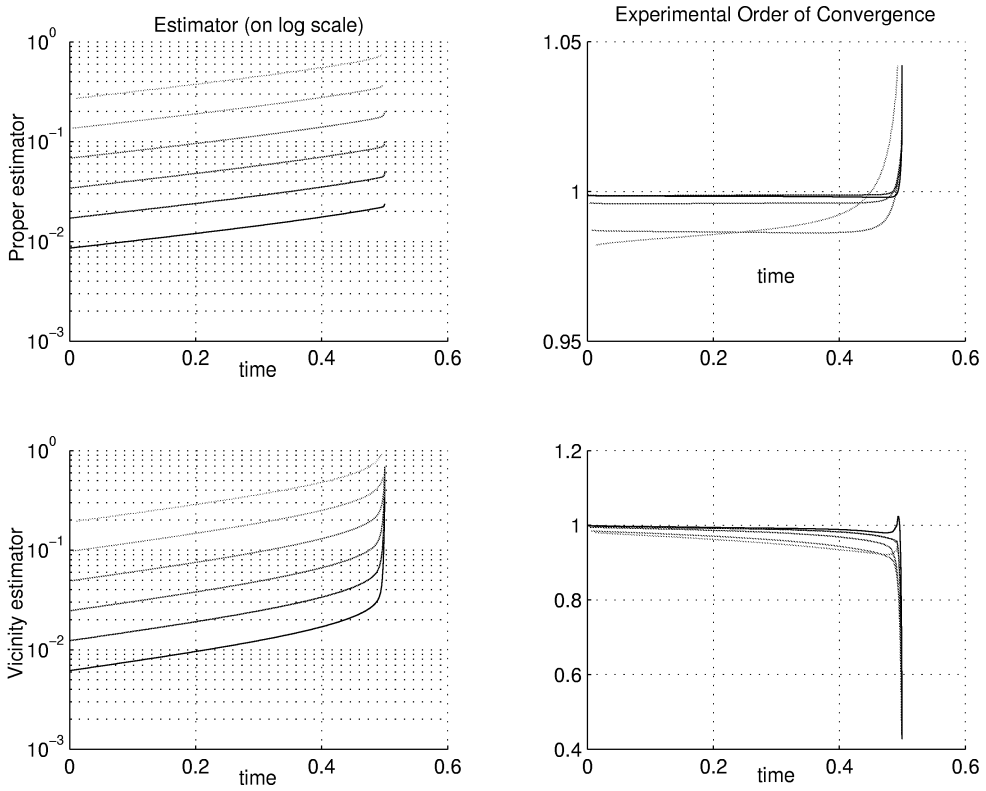

FIG. 7.4. *Estimators and experimental orders of convergence (EOC) vs. time for the shrinking sphere segment of Example 7.6. We exhibit the behavior of the proper estimator in the first row and that of the vicinity estimator in the second row. The different shades of gray (or colors) correspond to decreasing meshsizes h. We can observe two stages as time approaches the blow-up $t = 0.5$. In the first stage, the same observations made for Example 7.5 are valid in that* $\mathrm{EOC}\,\widetilde{\mathscr{E}} \approx \mathrm{EOC}\,E$ *and* $\mathscr{E}_\infty \to 0$. *In the second stage, the vicinity estimator* $\mathscr{E}_\infty$ *exhibits a blow-up, which means that the condition* (3.17) *of Theorem 3.6 is violated and that we can no longer rely on the proper estimator* $\widetilde{\mathscr{E}}$. *The vicinity estimator blow-up can be interpreted as numerical evidence of the boundary gradient blow-up occurring at* $t = 0.5$.

7.4 Remark (The discrete initial condition) In our computations, we take the *minimal surface projection* for the discrete initial values, i.e., $U_h^0 := u_h(0) = M_h u(0)$, where $M_h v$ is defined, for each $v \in \mathrm{W}_1^1(\Omega)$, as the unique function in $\mathbb{V}_h$ such that

$$\left\langle \frac{\nabla M_h v}{Q M_h v}, \nabla \phi_h \right\rangle = \left\langle \frac{\nabla v}{Qv}, \nabla \phi_h \right\rangle, \ \forall \phi_h \in \mathring{\mathbb{V}}_h, \tag{7.9}$$

and that interpolates v on the boundary.

This choice of the discrete initial value, reduces the initial transients that can occur with other choices for the discrete initial values such as Lagrange interpolation.

7.5 Example (Smooth exact solution on a square) Our first series of tests use the following exact solution as a benchmark

$$u(x,y;t) = t\left(\sin(t) - \sin(t - x(1-x)y(1-y))\right), \quad (x,y,t) \in [0,1]^2 \times [0,8]. \tag{7.10}$$

The function u is smooth and has zero initial and boundary values which allows us to focus on the effect of the estimators only. It is the solution of Problem 1.1 where

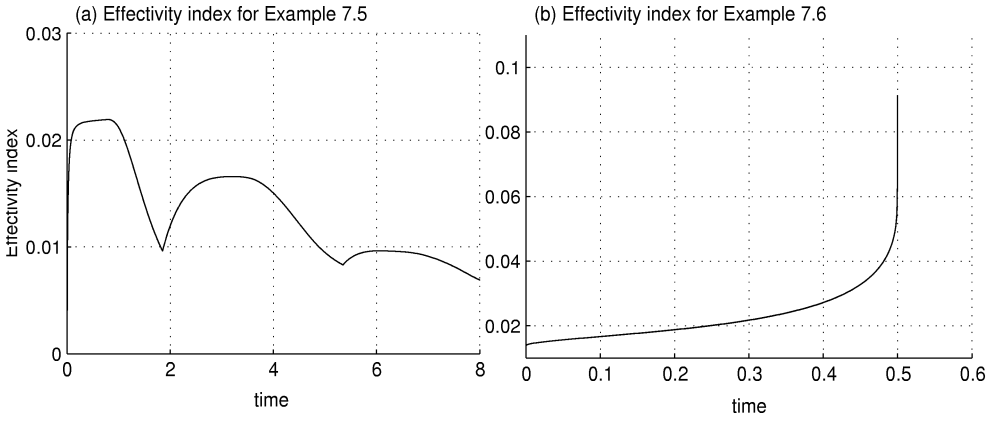

FIG. 7.5. *Effectivity indexes for Examples 7.5 and 7.6 These indexes, defined for the finest mesh, are numerical realizations of the constants on the right-hand side of* (3.16) *or* (3.18). *Figure (a) refers to the smooth exact solution of Example 7.5; the effectivity index behaves well in time. Figure (b) shows the effectivity index for the proper estimator 7.6 which has a blow-up at time* $t = 0.5$. *Consequently, the exponential behavior predicted for the factor* C *in Theorem 3.6 might be sharp. The graph's behavior in (b) close to* $t = 0.5$ *is questionable though, as the vicinity estimator blows up there according to Figure 7.4, and thus the conditional estimate is not guaranteed to hold anymore.*

the right-hand side f is obtained by applying the differential operator of (1.2) on u. We performed a series of computations, on uniform meshes, with the meshsizes $h_i = (0.5)^i$ for $1 \leq i \leq 6$. The results are reported in form of graphs where the abscissa always denotes the time variable; this allows us to track the behavior of the errors and estimators in time. Figure 7.1 shows the behavior of the exact spatial errors, namely the geometric errors and those in the customary Sobolev norms for evolution equations. Figure 7.2, shows the proper and vicinity estimator's behavior in time.

As shown by the right-hand side in Figure 7.1, the $\mathrm{EOC}\, E \approx 1$—this is to be expected by the a priori results, derived in the case of smooth solutions [8]. Although the normal velocity error tends to decrease faster, the geometric error decreases like the geometric energy error which has order 1.

Reliability of the estimate (7.4) can be seen by the fact that $\mathrm{EOC}\, \widetilde{\mathscr{E}} \approx 1$ and that $\mathscr{E}_\infty \to 0$. On the other hand, we notice that $\mathrm{EOC}\, \mathscr{E}_\infty \leq 1$, which implies that $\mathrm{EOC}(\mathscr{E}_\infty^{1/2}) \prec \mathrm{EOC}\, E$ and that the unconditional estimate (7.3) is not reliable.

The effectivity index $\widetilde{C}$, relative to the estimate (7.4) is plotted in Figure 7.5 (a) as a function of time. In this example, the effectivity index is bounded in time and we do not detect the exponential behavior predicted by the worst-case-scenario bound in (3.19).

7.6 Example (Shrinking spherical segment) This second numerical example is inspired from a simple geometric situation. A sphere that moves by mean curvature flow shrinks to a point in finite time [14]. If we assume that the initial radius of the sphere equals 2 and that the center is fixed at $(0, 0, 0)$, then the segment of the surface that lies above the square $[0, 1] \times [0, 1] \times \{0\} \in \mathbb{R}^3$ is the graph of the function

$$u(\boldsymbol{x}, t) = \sqrt{4 - 4t - |\boldsymbol{x}|^2}, \quad (\boldsymbol{x}, t) \in [0, 1]^2 \times [0, 0.5]. \tag{7.11}$$

The function u constitutes thus a solution of Problem 1.1 with zero right-hand side f and time-dependent Dirichlet boundary value g. This is an interesting example

because of a blow-up of gradient occurs at the space-time boundary point $(1, 1; 1/2)$. Despite this singular behavior, the function u still satisfies Hypothesis 2.1 and our a posteriori error analysis applies. Notice that the a priori error analysis of Deckelnick & Dziuk does not apply [7, Prop. 3] in this case because of too stringent regularity assumptions. This example allows us to appreciate the exponential worst-case-scenario bound on $\widetilde{C}$ (factor C in Theorem 3.6), as that bound is expected to behave like $\exp(1/\sqrt{0.5-t})$ as $t \to 0.5$. Numerical solutions have been computed on uniform triangulations with meshsizes $h_i = (0.5)^i, \quad i = 2, \ldots, 7$. The type of data we report is similar to that in §7.5: the errors and their asymptotic behavior are reported in Figure 7.3, while Figure 7.4 shows the behavior of the estimators. We refer to the caption for a comment on the blow-up at $t = 0.5$ and its effect on the estimators and estimate validity. In Figure 7.5 (b) we report the effectivity index of the proper estimator, that justifies in part the exponential behavior predicted by the theory. Notice that because of the blow-up behavior the effectivity index is not so meaningful in the last part of the graph, close to $t = 0.5$ where the vicinity estimator is too big.

Acknowledgments. We express our thanks to Gerhard Dziuk for fruitful discussions, Kunibert Siebert and Alfred Schmidt for introducing us to ALBERT [20], and Claudio Verdi for his advice and support during the stay of O. Lakkis at the Dipartimento di Matematica, Università di Milano, Italia. We are also grateful to the Isaac Newton Institute, Cambridge UK, where this work was finished. The programme *Computational Challenges in PDEs 2003* at the Isaac Newton Institute provided an excellent support as well as a stimulating and friendly scientific environment.

REFERENCES

[1] L. Ambrosio, *Lecture notes on geometric evolution problems, distance function and viscosity solutions*, Pubblicazioni 1029, Istituto di Analisi Numerica del CNR, Pavia, Italy, 1997.

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

[3] E. Bänsch, *Finite element discretization of the Navier-Stokes equations with a free capillary surface*, Numer. Math., 88 (2001), pp. 203–235.

[4] G. Barles and P. E. Souganidis, *A new approach to front propagation problems: theory and applications*, Arch. Rational Mech. Anal., 141 (1998), pp. 237–296.

[5] K. A. Brakke, *The motion of a surface by its mean curvature*, Princeton University Press, Princeton, N.J., 1978.

[6] K. Deckelnick, *Error bounds for a difference scheme approximating viscosity solutions of mean curvature flow*, Interfaces Free Bound., 2 (2000), pp. 117–142.

[7] K. Deckelnick and G. Dziuk, *Error estimates for a semi-implicit fully discrete finite element scheme for the mean curvature flow of graphs*, Interfaces Free Bound., 2 (2000), pp. 341–359.

[8] G. Dziuk, *Numerical schemes for the mean curvature flow of graphs*, in Variations of Domain and Free-Boundary Problems in Solid Mechanics (Paris, 1997), Kluwer Acad. Publ., Dordrecht, 1999, pp. 63–70.

[9] 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.

[10] ———, *Adaptive finite element methods for parabolic problems. IV. Nonlinear problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1729–1749.

[11] L. C. Evans and J. Spruck, *Motion of level sets by mean curvature. I*, J. Differential Geom., 33 (1991), pp. 635–681.

[12] F. Fierro and A. Veeser, *On the a posteriori error analysis for equations of prescribed mean curvature*, Math. Comp., (2003). Posted March 26.

[13] M. Fried and A. Veeser, *Simulation and numerical analysis of dendritic growth*, in Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems, Springer, Berlin, 2001, pp. 225–252, 812–813.

[14] G. Huisken, *Asymptotic behavior for singularities of the mean curvature flow*, J. Differential Geom., 31 (1990), pp. 285–299.

[15] ———, *Local and global behaviour of hypersurfaces moving by mean curvature*, in Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990), Amer. Math. Soc., Providence, RI, 1993, pp. 175–191.

[16] G. M. Lieberman, *Second Order Parabolic Differential Equations*, World Scientific Publishing Co. Inc., River Edge, NJ, 1996.

[17] R. H. Nochetto, G. Savaré, and C. Verdi, *A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations*, Comm. Pure Appl. Math., 53 (2000), pp. 525–589.

[18] R. H. Nochetto, A. Schmidt, and C. Verdi, *A posteriori error estimation and adaptivity for degenerate parabolic problems*, Math. Comp., 69 (2000), pp. 1–24.

[19] M. Picasso, *Adaptive finite elements for a linear parabolic problem*, Comput. Methods Appl. Mech. Engrg., 167 (1998), pp. 223–237.

[20] A. Schmidt and K. G. Siebert, *ALBERT—software for scientific computations and applications*, Acta Math. Univ. Comenian. (N.S.), 70 (2000), pp. 105–122.

[21] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[22] J. Sethian and S. J. Osher, *The design of algorithms for hypersurfaces moving with curvature-dependent speed*, in Nonlinear hyperbolic equations—theory, computation methods, and applications (Aachen, 1988), Vieweg, Braunschweig, 1989, pp. 544–551.

[23] N. N. Uraltseva, *Boundary regularity for flows of nonparametric surfaces driven by mean curvature*, in Motion by Mean Curvature and Related Topics (Trento, 1992), G. Buttazzo and A. Visintin, eds., de Gruyter, Berlin, 1994, pp. 198–209.

[24] R. Verfürth, *A Review of* A Posteriori *Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, Chichester-Stuttgart, 1996.

[25] ———, *A posteriori error estimates for nonlinear problems: $L^r(0,T;W^{1,\rho}(\omega))$-error estimates for finite element discretizations of parabolic equations*, Numer. Methods Partial Differential Equations, 14 (1998), pp. 487–518.

[26] N. J. Walkington, *Algorithms for computing motion by mean curvature*, SIAM J. Numer. Anal., 33 (1996), pp. 2215–2238.