Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Optimal and robust a posteriori error estimates in $L^\infty(L^2)$ for the approximation of Allen-Cahn equations past singularities

Sören Bartels[1], Rüdiger Müller[2]

submitted:

[1] Institut für Numerische Simulation
Rheinische Friedrichs-Wilhelms-Universität Bonn
Wegelerstr. 6
53115 Bonn
E-Mail: bartels@ins.uni-bonn.de

[2] Weierstraß-Institut
für Angewandte Analysis
und Stochastik
Mohrenstr. 39
10117 Berlin
E-Mail: mueller@wias-berlin.de

No. 1416
Berlin 2009

2000 *Mathematics Subject Classification.* 65M60,65M15,35K55.

Key words and phrases. Allen-Cahn equation, mean curvature flow, finite element method, error analysis, adaptive methods.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract. Optimal a posteriori error estimates in $L^\infty(0,T;L^2(\Omega))$ are derived for the finite element approximation of Allen-Cahn equations. The estimates depend on the inverse of a small parameter only in a low order polynomial and are valid past topological changes of the evolving interface. The error analysis employs an elliptic reconstruction of the approximate solution and applies to a large class of conforming, nonconforming, mixed, and discontinuous Galerkin methods. Numerical experiments illustrate the theoretical results.

1. Introduction

In this paper, we derive optimal a posteriori error estimates in $L^\infty(0,T;L^2(\Omega))$ for the finite element approximation of the Allen-Cahn problem

$$\begin{aligned} \partial_t u - \Delta u + \varepsilon^{-2} f(u) &= 0 \quad \text{in } (0,T)\times\Omega, \\ \partial_n u &= 0 \quad \text{on } (0,T)\times\partial\Omega, \\ u(0,\cdot) &= u_0, \end{aligned} \tag{1}$$

with $T>0$, $\Omega\subseteq\mathbb{R}^d$, $d=2,3$, $u_0\in L^2(\Omega)$, $f(u)=u^3-u$, and $0<\varepsilon\ll 1$. Our ultimate goal is to prove estimates that are robust in the small parameter ε past generic singularities in the evolution described by (1).

The mathematical model (1) is the simplest version of a phase field model and was introduced in [AC79] to model the motion of phase boundaries by surface tension. The interface $\Gamma_t := \{x\in\Omega : u(x,t)=0\}$ separates regions in which $u(t,\cdot)\approx+1$ from those in which $u(t,\cdot)\approx-1$. As $\varepsilon\to 0$, the evolution of the interface approaches the motion of a hypersurface governed by Brakke's mean curvature flow [Bra78, Ilm93]. An important feature of the diffuse interface model (1) is that topological changes in Γ_t are captured whereas sharp interface models typically require artificial adaptations to model such effects.

A straightforward error analysis for the numerical approximation of (1) leads to an exponential dependence of error estimates on ε^{-1}. The first successful attempt to establish robust a priori error estimates, i.e., error estimates that depend on ε^{-1} only in a polynomial, for the approximation of Allen-Cahn equations is due to [FP03]. Those results are based on uniform bounds for the principal eigenvalue of the linearized Allen-Cahn operator about the exact solution, i.e., for the quantity

$$-\lambda_{AC}(t) := \inf_{v\in H^1(\Omega)\setminus\{0\}} \frac{\|\nabla v\|^2 + (f'(u(t))v,v)}{\|v\|^2},$$

where $(\cdot,\cdot)$ and $\|\cdot\|$ denote the inner product and the norm in $L^2(\Omega)$, respectively. Such bounds are available as long as (1) describes the smooth evolution of a developed interface Γ_t, cf. [AF93, Che94, dMS95]. The ideas of [FP03] have been carried over to an a posteriori error analysis in [KNS04, FW05] employing a continuation argument. Instead of using a priori bounds for $\lambda_{AC}(t)$ to derive a posteriori error estimates, it has been proposed in [Bar05] to extract the relevant information about the stability of the evolution from the approximate solution U by considering the principal eigenvalue of the linearized Allen-Cahn operator about $U(t)$, i.e.,

$$-\Lambda_{AC}(t) := \inf_{v\in H^1(\Omega)\setminus\{0\}} \frac{\|\nabla v\|^2 + (f'(U(t))v,v)}{\|v\|^2}.$$

This still allows to rigorously derive a posteriori error estimates and establishes a mechanism to detect critical times at which uniform bounds for λ_{AC} and its approximation Λ_{AC} break down. In the recent paper [BMO09b] it has been shown that the weaker bound

$$\int_0^T \Lambda_{AC}^+(t)\,\mathrm{d}t \le C_0 + \log(\varepsilon^{-\kappa})$$

is sufficient for a robust a posteriori error analysis and that this bound is realistic for generic topological changes of Allen-Cahn evolutions. Specifically, the computable left-hand side of the estimate enters the error estimates of [BMO09b] exponentially and hence no bounds are required a priori.

The estimates of [BMO09b] hold provided that the computable upper bound $\eta_{L^2(H^1)}$ for the error in $L^2(0,T;H^1(\Omega))$ satisfies

$$\eta_{L^2(H^1)} \leq C\varepsilon^{4+3\kappa}$$

which imposes restrictive conditions on discretizations since we only expect $\eta_{L^2(H^1)} \sim \varepsilon^{-5/2}(\tau + h)$ for an implicit scheme with temporal and spatial step sizes τ and h, respectively. The quantity $\eta_{L^2(H^1)}$ also controls the error in the weaker norm of $L^\infty(0,T;L^2(\Omega))$ but this bound is suboptimal since the optimal convergence rate is $\tau + h^2$ for the error measured in this norm. By establishing optimal estimates for the error in $L^\infty(0,T;L^2(\Omega))$ we expect to obtain a posteriori error estimates that are valid under less restrictive conditions on the corresponding computable estimator $\eta_{L^\infty(L^2)}$.

Optimal a posteriori error estimates in $L^\infty(0,T;L^2(\Omega))$ for parabolic problems have been derived under certain conditions on triangulations in [EJ95a, EJ95b] using duality arguments. A different approach to the derivation of such estimates by energy techniques has been proposed and analyzed for semidiscrete schemes in [MN03] and investigated for fully discrete schemes in [LM06, GL08]. The approach consists in constructing at each time step t_j a function w^j such that the approximate solution U^j of the linear parabolic problem at time t_j is the Galerkin approximation to an elliptic problem whose exact solution is w^j. This concept is called *elliptic reconstruction* and allows to derive a posteriori error estimates for parabolic problems by reducing a large part of the analysis to known a posteriori error estimates for elliptic problems. Elliptic reconstruction may be regarded as the a posteriori analogue of elliptic projection which has been used to derive optimal a priori error estimates in $L^\infty(0,T;L^2(\Omega))$ for parabolic equations in [Whe73].

We combine the method of elliptic reconstruction of [MN03, LM06] with techniques recently developed in [BMO09b] to derive robust and optimal a posteriori error estimates in $L^\infty(0,T;L^2(\Omega))$ for the numerical approximation of the nonlinear parabolic partial differential equation (1). Let $(U^j)_{j=0,1,\dots,J} \subset L^2(\Omega)$ denote a sequence of approximations to the exact solution of (1) obtained with the implicit Euler scheme in time and some finite element method in space, i.e., for given $U^{j-1} \in \mathbb{V}_h^{j-1}$ the function $U^j \in \mathbb{V}_h^j$ satisfies

$$\tau_j^{-1}(U^j - U^{j-1}, V) + a_h^j(U^j, V) = -\varepsilon^{-2}(f(U^j), V) \tag{2}$$

for all $V \in \mathbb{V}_h^j$. Here, τ_j is a time-step size, $\mathbb{V}_h^j$ an approximation space, and a_h^j a bilinear form on $\mathbb{V}_h^j$ that approximates the Laplace operator. We let $U \in H^1(0,T;L^2(\Omega))$ denote the function that is obtained by piecewise affine interpolation of the approximations $(U^j)_{j=0,1,\dots,J}$ subordinate to the partition of the time interval $(0,T)$ defined by the time steps $(\tau_j)_{j=1,2,\dots,J}$. Under moderate consistency and compatibility conditions on the bilinear forms a_h^j, cf. (7) and Assumption (COMP) below, that allow conforming, nonconforming, mixed, and discontinuous

Galerkin methods, we establish the computable error bound

$$\sup_{s\in(0,T)}\|(u-U)(s)\| \le \max_{j=0,1,\dots,J}\mathcal{E}_{L^2}(U^j;\mathbb{V}_h^j)$$
$$+8\Big\{\sum_{j=0}^{J}\tau_j\Big(\tau_j\|d_t\Delta_h^jU^j\|+\tau_j\varepsilon^{-2}C_{f',I_j}\|d_tU^j\|+\mathcal{E}_{L^2}(d_tU^j;\widehat{\mathbb{V}}_h^j)+\varepsilon^{-2}C_{f',I_j}\max_{k=j-1,j}\mathcal{E}_{L^2}(U^k;\mathbb{V}_h^k)\Big)$$
$$+\varepsilon^{-1}\Big(\sum_{j=0}^{J}\tau_jC_{C\ell}\big(\tau_j^{-2}\|h_j(U^{j-1}-P_h^jU^{j-1})\|^2+\varepsilon^{-4}\|h_j(f(U^j)-P_h^jf(U^j))\|^2\big)\Big)^{1/2}$$
$$+\|u_0-U^0\|+\mathcal{E}_{L^2}(U^0;\mathbb{V}_h^0)\Big\}\exp\Big(4\sum_{j=0}^{J}\tau_j\big((1-\varepsilon^2)\overline{\Lambda}_{AC}^j+1+\varepsilon^{-2}\eta_{f'}^j\big)^+\Big),$$

which holds provided that the terms inside the curly brackets, denoted $\eta_{L^\infty(L^2)}$, and the exponential factor, denoted E, satisfy

$$\eta_{L^\infty(L^2)} \le \frac{\varepsilon^4}{(4\mu_gC_S(1+T))^2}(4E)^{-3} \le C\varepsilon^{4+3\kappa} \tag{3}$$

The symbol d_t denotes the backward difference operator, $-\Delta_h^j$ is a discrete version of the Laplace operator defined by a_h^j, h_j is a positive meshsize function, P_h^j is the L^2 projection onto $\mathbb{V}_h^j$, C_{f',I_j}, $\eta_{f'}^j$, and μ_g are computable quantities related to the nonlinearity f, and $\overline{\Lambda}_{AC}^j$ stands for a computable upper bound for $\Lambda_{AC}(t_j)$. We refer the reader to the subsequent sections for further details. It is important to notice the linear accumulation of error estimators for space and time discretization residuals in the first sum inside the curly brackets on the right-hand side of our error estimate. To evaluate the upper bound it is not necessary to compute $\Delta_h^jU^j$ explicitly since this term is known from (2). Moreover, the computation of the nonlocal operator P_h^j can be avoided if the scheme (2) and the error estimate are slightly modified by incorporating appropriate local mesh transfer operators, cf. Remark 4.4 below. For lowest order conforming methods based on regular triangulations $\mathcal{T}_h^j$ of Ω that define the spaces $\mathbb{V}_h^j$ the estimator $\mathcal{E}_{L^2}(U^j;\mathbb{V}_h^j)$ is, up to generic constants, given by

$$\mathcal{E}_{L^2}(U^j;\mathbb{V}_h^j)=\big\|h_j^2\big(\tau_j^{-1}(U^j-P_h^jU^{j-1})+\varepsilon^{-2}P_h^jf(U^j)\big)\big\|+\big\|h_j^{3/2}\big[\nabla U^j\cdot n_{\mathcal{F}_h^j}\big]\big\|_{L^2(\cup\mathcal{F}_h^j)},$$

where we use standard notation for the jumps across element sides contained in $\mathcal{F}_h^j$, cf. Remark 4.1 below. Analogously, the estimator $\mathcal{E}_{L^2}(d_tU^j;\mathbb{V}_h^{j-1/2})$ is given by

$$\mathcal{E}_{L^2}(d_tU^j;\widehat{\mathbb{V}}_h^j)=\big\|\hat{h}_j^2d_t\big(\tau_j^{-1}(U^j-P_h^jU^{j-1})+\varepsilon^{-2}P_h^jf(U^j)\big)\big\|+\big\|\hat{h}_j^{3/2}\big[\nabla d_tU^j\cdot n_{\widehat{\mathcal{F}}_h^j}\big]\big\|_{L^2(\cup\widehat{\mathcal{F}}_h^j)},$$

where the triangulation $\widehat{\mathcal{T}}_h^j$ defines $\widehat{\mathbb{V}}_h^j$ and is the finest common coarsening of $\mathcal{T}_h^j$ and $\mathcal{T}_h^{j-1}$. A similar estimator is needed to obtain pointwise control over certain residuals related to the nonlinearities in the error equation.

We expect that $\eta_{L^\infty(L^2)}\sim\varepsilon^{-7/2}(\tau+h^2)$ and therefore, we obtain a significantly weaker condition for the validity of the error estimate than the one in [BMO09b]. For smooth evolutions of developed interfaces we deduce $E\sim 1$ from [Che94, dMS95] while for evolutions that undergo topological changes we observe $E\sim\varepsilon^{-\kappa}$ with a small number κ, cf. [BMO09b]. In particular, E does not grow exponentially in ε^{-1}.

As a byproduct we obtain an error estimate in the seminorm of $L^2(0,T;H^1(\Omega))$ that holds under a significantly weaker condition than the one stated in [BMO09b], namely, if (3) holds then we

have for a lowest order conforming method that

$$\int_0^T |||(u-U)(s)|||^2 \,\mathrm{d}s \le \sum_{j=0}^{J} \frac{\tau_j}{2}\big(\mathcal{E}_{H^1}^2(U^{j-1},\mathbb{V}_h^{j-1}) + \mathcal{E}_{H^1}^2(U^j,\mathbb{V}_h^j)\big) + 2\varepsilon^{-2}\eta_{L^\infty(L^2)}^2 E^2,$$

with $|||\cdot||| = \|\nabla\cdot\|$ and

$$\mathcal{E}_{H^1}(U^j;\mathbb{V}_h^j) = \big\|h_j\big(\tau_j^{-1}(U^j - P_h^j U^{j-1}) + \varepsilon^{-2}P_h^j f(U^j)\big)\big\| + \big\|h_j^{1/2}\big[\nabla U^j\cdot n_{\mathcal{F}_h^j}\big]\big\|_{L^2(\cup\mathcal{F}_h^j)}.$$

In contrast to the result of [BMO09b] we assume H^2 regularity of the Laplace operator in Ω and we require one additional order of differentiability of the potential function f here. A similar result can be derived for nonconforming and discontinuous Galerkin finite element methods by choosing an appropriate extension $|||\cdot|||$ of the seminorm in $H^1(\Omega)$. For ease of presentation we do not aim at stating the most general conditions on discretizations that lead to such estimates and instead refer the reader to [GL08] for a related, more detailed discussion in case of the linear heat equation.

Our estimates naturally lead to adaptive algorithms for the efficient approximation of (1) by local mesh refinement. The contributions to the right-hand sides of our estimates can be categorized into localizable estimators related to spatial and temporal discretization errors as well as mesh-change and oscillation residuals which allows an individual local adjustment of time-step and mesh-sizes. Owing to the strongly localized features of solutions to (1), adaptivity is of fundamental importance for the development of efficient approximation schemes and the techniques discussed in this paper directly transfer to other, more sophisticated phase field models such as Ginzburg-Landau, Cahn-Hilliard, and Cahn-Larché equations, cf. [BM08, BMO09a]. In particular, the estimates presented in this paper do not rely on the validity of a maximum principle.

The outline of this paper is as follows. We state some preliminaries in Section 2, derive an abstract a posteriori error estimate in Section 3, and discuss the application to various finite element methods in Section 4. Numerical experiments that illustrate the reliability of our method are reported in Section 5.

2. Preliminaries

Let $\Omega\subset\mathbb{R}^d$, $d=2,3$ be a bounded, polygonal or polyhedral Lipschitz domain. The outer unit normal on $\partial\Omega$ is denoted by n and $\partial_n v$ is the normal derivative of a function v on $\partial\Omega$. For a real number $r\ge 0$ we set $B_r := \{x\in\mathbb{R}^\ell : |x|<r\}$; the positive part of a real number is denoted by s^+, i.e., $s^+ = \max\{s,0\}$ for all $s\in\mathbb{R}$. Standard notation is used for Sobolev and Lebesgue spaces and we write $\|\cdot\|$ whenever $\|\cdot\|_{L^2(\Omega)}$ is meant; $(\cdot,\cdot)$ is the inner product in $L^2(\Omega;\mathbb{R}^\ell)$, $\ell\in\mathbb{N}$. For a Banach space X its dual is denoted X^* and $\langle\cdot;\cdot\rangle$ is the corresponding duality pairing. We define

$$\mathbb{V} := H^1(\Omega)$$

and write $\|\cdot\|_*$ for the induced norm on $\mathbb{V}^*$. The bilinear form $a:\mathbb{V}\times\mathbb{V}\to\mathbb{R}$ is for $v,w\in\mathbb{V}$ defined through

$$a(v,w) := (\nabla v,\nabla w). \tag{4}$$

We assume that $0<\varepsilon\le 1$ and that the potential function f has the following properties.

Assumption (POT). *(i) There exists a nonnegative function $F\in C^3(\mathbb{R})$ such that $f=F'$.*
(ii) There exists $C_f\ge 0$ such that $f'(u)\ge -C_f$ for all $u\in\mathbb{R}$.
(iii) There exist $\delta>0$ with $\delta<2$ if $d=2$ and $\delta\le 1$ if $d=3$ and a nonnegative function $g\in C(\mathbb{R})$ such that for all $a,b\in\mathbb{R}$ we have

$$\big(f(a)-f(b)-f'(b)(a-b)\big)(a-b)\ge -g(b)|a-b|^{2+\delta}.$$

For $F(u) = (u^2-1)^2/4$, $u \in \mathbb{R}$, and $f = F'$ the estimate $f'(u) = 3u^2 - 1 \geq -1$, $u \in \mathbb{R}$, and the Taylor expansion

$$f(a) - f(b) - f'(b)(a-b) = 3b(a-b)^2 + (a-b)^3,$$

valid for all $a, b \in \mathbb{R}$, imply that (POT) holds with $C_f = 1$, $\delta = 1$, and $g(b) = 3|b|$, $b \in \mathbb{R}$.

Assumption (POT) implies that there exists a unique function

$$u \in X_{AC} := H^1(0,T;\mathbb{V}^*) \cap L^\infty(0,T;\mathbb{V})$$

satisfying $u(0) = u_0$ continuously in $L^2(\Omega)$ and

$$\langle \partial_t u(t), v\rangle + a(u(t), v) = -\varepsilon^{-2}(f(u(t)), v) \tag{5}$$

for almost every $t \in (0,T)$ and every $v \in \mathbb{V}$. The function u is called weak solution of the Allen-Cahn equation. We suppress the dependence of u upon ε but stress that all appearing constants do not depend on ε^{-1}. Notice that (5) is the L^2 gradient flow of the energy functional

$$E_\varepsilon(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 \,\mathrm{d}x + \varepsilon^{-2}\int_\Omega F(u)\,\mathrm{d}x.$$

The following generalization of Gronwall's lemma, which allows an additional superlinear term that can be controlled as long as the function remains sufficiently small, is an essential tool for our error analysis. Its proof is adapted from [KNS04, BMO09b].

Lemma 2.1 (Generalized Gronwall lemma)**.** *Suppose that the nonnegative functions $y_1 \in C([0,T])$, $y_2, y_3 \in L^1(0,T)$, $\alpha \in L^\infty(0,T)$, and the real number $A \geq 0$ are such that y_1 is monotonically increasing and that*

$$y_1(t) + \int_0^t y_2(s)\,\mathrm{d}s \leq A + \int_0^t \alpha(s) y_1(s)\,\mathrm{d}s + \int_0^t y_3(s)\,\mathrm{d}s$$

for all $t \in [0,T]$. Assume that for $B \geq 0$, $\beta > 0$, and every $t \in [0,T]$ we have

$$\int_0^t y_3(s)\,\mathrm{d}s \leq B y_1^\beta(t) \int_0^t (y_1(s) + y_2(s))\,\mathrm{d}s.$$

Set $E := \exp\big(\int_0^T \alpha(s)\,\mathrm{d}s\big)$ and assume that $4AE \leq (4B(1+T)E)^{-1/\beta}$. We then have

$$y_1(T) + \int_0^T y_2(s)\,\mathrm{d}s \leq 4A \exp\Big(\int_0^T \alpha(s)\,\mathrm{d}s\Big).$$

Proof. Set $\theta := 4AE$ if $A > 0$ and let $\theta > 0$ such that $2B(1+T)\theta^\beta E \leq 1$ otherwise. Define

$$I_\theta := \big\{t' \in [0,T] : \Upsilon(t') := y_1(t') + \int_0^{t'} y_2(s)\,\mathrm{d}s \leq \theta\big\}.$$

Since $y_1(0) \leq A < \theta$ and since Υ is continuous and monotonically increasing we have $I_\theta = [0, t_m]$ for some $0 < t_m \leq T$. For every $t \in [0, t_m]$ we have

$$\begin{aligned} y_1(t) + \int_0^t y_2(s)\,\mathrm{d}s &\leq A + \int_0^t \alpha(s) y_1(s)\,\mathrm{d}s + B y_1^\beta(t)\int_0^t (y_1(s) + y_2(s))\,\mathrm{d}s \\ &\leq A + \int_0^t \alpha(s) y_1(s)\,\mathrm{d}s + B(1+T)\theta^{1+\beta}. \end{aligned}$$

An application of Gronwall's lemma, cf., e.g., [IT79], the condition on A, and the choice of θ yield that for all $t \in [0, t_m]$ we have

$$y_1(t) + \int_0^t y_2(s)\,\mathrm{d}s \leq (A + B(1+T)\theta^{1+\beta})E \leq \frac{\theta}{2}.$$

This implies $\Upsilon(t_m) < \theta$, hence $t_m = T$, and thus proves the lemma if $A > 0$. If $A = 0$ we may choose θ arbitrarily small to deduce the assertion. □

Remark 2.2. *The factor 4 on the right-hand side of the estimate of the lemma can be replaced by any number bigger than 2 or by 2 if $\alpha \not\equiv 0$.*

3. Abstract a posteriori error analysis

Given a sequence of positive time steps $(\tau_j)_{j=0,1,...,J}$ that defines the partition $0 = t_0 < t_1 < ... < t_J = T$ of $(0,T)$ and subspaces $(\mathbb{V}_h^j)_{j=0,1,...,J}$ of $L^2(\Omega)$, we assume that $(U^j)_{j=0,1,...,J} \subset L^2(\Omega)$ is such that for $j = 1, 2, ..., J$ we have $U^j \in \mathbb{V}_h^j$ and

$$\tau_j^{-1}(U^j - U^{j-1}, V) + a_h^j(U^j, V) = -\varepsilon^{-2}(f(U^j), V)$$

for all $V \in \mathbb{V}_h^j$. Here, $a_h^j : \mathbb{V}_h^j \times \mathbb{V}_h^j \to \mathbb{R}$ is a bilinear form that approximates the bilinear form a from (4). Equivalently, we have for $j = 1, 2, ..., J$ that

$$\tau_j^{-1}(U^j - P_h^j U^{j-1}) - \Delta_h^j U^j = -\varepsilon^{-2} P_h^j f(U^j), \tag{6}$$

where $P_h^j : L^2(\Omega) \to \mathbb{V}_h^j$ denotes the L^2 projection onto $\mathbb{V}_h^j$ and $-\Delta_h^j : \mathbb{V}_h^j \to \mathbb{V}_h^j$ is for $V \in \mathbb{V}_h^j$ defined through the identity

$$(-\Delta_h^j V, W) = a_h^j(V, W)$$

for all $W \in \mathbb{V}_h^j$. We assume that for $j = 0, 1, ..., J$ constant functions are included in $\mathbb{V}_h^j$ and a_h^j vanishes for constant functions, i.e.,

$$1 \in \mathbb{V}_h^j \qquad \text{and} \qquad a_h^j(V, 1) = 0 \tag{7}$$

for all $V \in \mathbb{V}_h^j$. This ensures that the elliptic reconstruction of a function $-\Delta_h^j V$ for $V \in \mathbb{V}_h^j$ is well defined.

Definition 3.1 (Elliptic reconstruction)**.** *For $j = 0, 1, ..., J$ define*

$$\xi_h^j := -\Delta_h^j U^j$$

and let $w^j \in \mathbb{V}$ be such that

$$(\nabla w^j, \nabla v) = (\xi_h^j, v) \qquad \text{and} \qquad \int_\Omega w^j \,\mathrm{d}x = \int_\Omega U^j \,\mathrm{d}x$$

for all $v \in \mathbb{V}$. Let $w, U \in H^1(0,T; L^2(\Omega))$ be for $j = 1, 2, ..., J$ and $t \in [t_{j-1}, t_j]$ defined through

$$w(t) := \ell_{j-1}(t) w^{j-1} + \ell_j(t) w^j,$$
$$U(t) := \ell_{j-1}(t) U^{j-1} + \ell_j(t) U^j,$$

where $\ell_j(t) = (t - t_{j-1})/\tau_j$ and $\ell_{j-1}(t) = 1 - \ell_j(t)$ for $t \in [t_{j-1}, t_j]$.

Notice that for $j = 0, 1, ..., J$ we have

$$-\Delta w^j = \xi_h^j \quad \text{in } \Omega, \qquad \partial_n w^j = 0 \quad \text{on } \partial\Omega.$$

Moreover, owing to the definition of $-\Delta_h^j$, we have that $U^j \in \mathbb{V}_h^j$ is for $j = 0, 1, ..., J$ the Galerkin approximation of the Poisson problem with homogeneous Neumann boundary conditions and right-hand side $\xi_h^j = -\Delta_h^j U^j$, i.e., we have

$$a(w^j, v) = (\xi_h^j, v), \qquad a_h^j(U^j, V) = (\xi_h^j, V)$$

for all $v \in \mathbb{V}$ and all $V \in \mathbb{V}_h^j$.

Lemma 3.2 (Perturbed parabolic evolution). *For $j = 1, 2, ..., J$ and $t \in (t_{j-1}, t_j)$ define $\Gamma(w, U; t) \in \mathbb{V}^*$ through*

$$\Gamma(w, U; t) := \partial_t(w - U) - \Delta(w - w^j) - \tau_j^{-1}(U^{j-1} - P_h^j U^{j-1}) + \varepsilon^{-2}(f(w) - P_h^j f(U^j)).$$

Then we have for almost every $t \in (0, T)$ that

$$\partial_t w - \Delta w = -\varepsilon^{-2} f(w) + \Gamma(w, U; t). \tag{8}$$

Proof. The identity follows from (6) upon noting that

$$\partial_t U - \tau_j^{-1}(U^j - P_h^j U^{j-1}) = -\tau_j^{-1}(U^{j-1} - P_h^j U^{j-1})$$

and $\langle \Delta w^j - \Delta_h^j U^j, v\rangle = 0$ for all $v \in \mathbb{V}$. □

The motivation for the following theorem is that the quantity $\exp\big(\int_0^T \lambda_{AC}^+(s)\,\mathrm{d}s\big)$ is bounded by some power of ε^{-1} and that computable bounds are available for the difference $w - U$ in various norms which can be made arbitrarily small by local mesh refinement.

Theorem 3.3 (General a posteriori estimate). *Let δ, C_f, and g be as in (POT). Suppose that $\overline{\Lambda}_{AC} \in L^1(0, T)$ is such that for almost every $t \in (0, T)$ we have*

$$-\overline{\Lambda}_{AC}(t) \le -\Lambda_{AC}(t) := \inf_{v \in \mathbb{V}\setminus\{0\}} \frac{\|\nabla v\|^2 + \varepsilon^{-2}(f'(U(t))v, v)}{\|v\|^2}$$

and assume that $\eta_{\Gamma,0}, \eta_{\Gamma,1}, \eta_{f'} : (0, T) \to \mathbb{R}$ and $\mu_g \in \mathbb{R}$ are such that

$$\begin{aligned} \langle \Gamma(w, U; t), v\rangle &\le \eta_{\Gamma,0}(t)\|v\| + \eta_{\Gamma,1}(t)\|\nabla v\| \\ \|f'(w(t)) - f'(U(t))\|_{L^\infty(\Omega)} &\le \eta_{f'}(t), \\ \sup_{s\in(0,T)} \|g(w(s))\|_{L^\infty(\Omega)} &\le \mu_g \end{aligned}$$

for almost every $t \in (0, T)$ and all $v \in \mathbb{V}$ and set $\mu_\Lambda(t) := 8\big((1 - \varepsilon^2)\overline{\Lambda}_{AC}(t) + C_f + \varepsilon^{-2}\eta_{f'}(t)\big)^+$. If

$$\eta^2 := 16\Big(\int_0^T \eta_{\Gamma,0}\,\mathrm{d}s\Big)^2 + 4\varepsilon^{-2}\int_0^T \eta_{\Gamma,1}^2\,\mathrm{d}s + 4\|u_0 - w^0\|^2 \\ \le \frac{\varepsilon^{8/\delta}}{(8\mu_g C_S(1 + T))^{2/\delta}}\Big(4\exp\big(\int_0^T \mu_\Lambda\,\mathrm{d}s\big)\Big)^{-1-2/\delta}$$

then

$$\sup_{s\in(0,T)} \|(u - U)(s)\| \le \sup_{s\in(0,T)} \|(U - w)(s)\| + 2\eta\exp\Big(\frac{1}{2}\int_0^T \mu_\Lambda\,\mathrm{d}s\Big)$$

and, for any seminorm $|||\cdot|||$ defined on the span of $\mathbb{V} \cup \bigcup_{j=0}^J \mathbb{V}_h^j$ such that $|||v||| = \|\nabla v\|$ for all $v \in \mathbb{V}$,

$$\Big(\int_0^T |||(u - U)(s)|||^2\,\mathrm{d}s\Big)^{1/2} \le \Big(\int_0^T |||(U - w)(s)|||^2\,\mathrm{d}s\Big)^{1/2} + \varepsilon^{-1}\sqrt{2}\eta\exp\Big(\frac{1}{2}\int_0^T \mu_\Lambda\,\mathrm{d}s\Big).$$

Proof. We abbreviate $\varrho := u - w$ and omit the argument t in the following. Subtracting (8) from (5) and testing the resulting equation by ϱ we have, incorporating (iii) of (POT),

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\varrho\|^2 + \|\nabla\varrho\|^2 &= -\varepsilon^{-2}(f(u) - f(w), \varrho) - \langle \Gamma(u, w), \varrho\rangle \\ &\le -\varepsilon^{-2}(f'(w)\varrho, \varrho) + \varepsilon^{-2}\|g(w)\|_{L^\infty(\Omega)}\|\varrho\|_{L^{2+\delta}(\Omega)}^{2+\delta} + \eta_{\Gamma,0}\|\varrho\| + \eta_{\Gamma,1}\|\nabla\varrho\| \\ &\le -\varepsilon^{-2}(f'(U)\varrho, \varrho) + \varepsilon^{-2}\eta_{f'}\|\varrho\|^2 + \varepsilon^{-2}\mu_g\|\varrho\|_{L^{2+\delta}(\Omega)}^{2+\delta} + \eta_{\Gamma,0}\|\varrho\| + \eta_{\Gamma,1}\|\nabla\varrho\|. \end{aligned}$$

Hölder's and Young's inequality, item (ii) of (POT), and straightforward manipulations lead to

$$\frac{1}{2}\frac{d}{dt}\|\varrho\|^2 + \|\nabla\varrho\|^2 \leq -(1-\varepsilon^2)\varepsilon^{-2}(f'(U)\varrho,\varrho) + C_f\|\varrho\|^2 + \varepsilon^{-2}\eta_{f'}\|\varrho\|^2 \\ + \varepsilon^{-2}\mu_g\|\varrho\|^{2+\delta}_{L^{2+\delta}(\Omega)} + \eta_{\Gamma,0}\|\varrho\| + \frac{1}{2\varepsilon^2}\eta_{\Gamma,1}^2 + \frac{\varepsilon^2}{2}\|\nabla\varrho\|^2.$$

The assumed property of $\overline{\Lambda}_{AC}$ implies that we have

$$-\varepsilon^{-2}(f'(U)\varrho,\varrho) \leq \overline{\Lambda}_{AC}\|\varrho\|^2 + \|\nabla\varrho\|^2.$$

This yields that

$$\frac{d}{dt}\|\varrho\|^2 + \varepsilon^2\|\nabla\varrho\|^2 \leq 2\eta_{\Gamma,0}\|\varrho\| + \varepsilon^{-2}\eta_{\Gamma,1}^2 + 2\big((1-\varepsilon^2)\overline{\Lambda}_{AC} + C_f + \varepsilon^{-2}\eta_{f'}\big)^+\|\varrho\|^2 + 2\varepsilon^{-2}\mu_g\|\varrho\|^{2+\delta}_{L^{2+\delta}(\Omega)}.$$

We integrate this estimate over $(0,t)$ and employ Hölder's and Young's inequality to verify that

$$\frac{1}{2}\sup_{s\in(0,t)}\|\varrho\|^2 + \frac{\varepsilon^2}{2}\int_0^t\|\nabla\varrho\|^2\,\mathrm{d}s \leq \|\varrho(0)\|^2 + \frac{1}{4}\sup_{s\in(0,t)}\|\varrho\|^2 + 4\Big(\int_0^t\eta_{\Gamma,0}\,\mathrm{d}s\Big)^2 \\ + \varepsilon^{-2}\int_0^t\eta_{\Gamma,1}^2\,\mathrm{d}s + \frac{1}{4}\int_0^t\mu_\Lambda\sup_{r\in(0,s)}\|\varrho\|^2\,\mathrm{d}s + 2\varepsilon^{-2}\mu_g\int_0^t\|\varrho\|^{2+\delta}_{L^{2+\delta}(\Omega)}\,\mathrm{d}s, \tag{9}$$

where we used that $\sup_{s\in(0,t)}a(s)+b(t) \leq 2c(t)$ if $a(t)+b(t)\leq c(t)$ for all $t\in(0,t)$. The conditions on δ in (iii) of (POT) together with Hölder's inequality and a Sobolev estimate permit us to derive the bound

$$\int_0^t\|\varrho\|^{2+\delta}_{L^{2+\delta}(\Omega)}\,\mathrm{d}s \leq \int_0^t\|\varrho\|^\delta\|\varrho\|^2_{L^{4/(2-\delta)}(\Omega)}\,\mathrm{d}s \leq C_S\big(\sup_{s\in(0,t)}\|\varrho\|^2\big)^{\delta/2}\int_0^t(\|\varrho\|^2+\|\nabla\varrho\|^2)\,\mathrm{d}s. \tag{10}$$

Setting

$$y_1(t) := \sup_{s\in(0,t)}\|\varrho(s)\|^2, \qquad y_2(t) := 2\varepsilon^2\|\nabla\varrho(t)\|^2, \qquad y_3(t) := 8\varepsilon^{-2}\mu_g\|\varrho(t)\|^{2+\delta}_{L^{2+\delta}(\Omega)}$$

for almost every $t\in(0,T)$, the estimates (9) and (10) show that we are in the situation of Lemma 2.1 with $A=\eta^2$, $B=8\varepsilon^{-4}\mu_g C_S$, $E=\exp\big(\int_0^T\mu_\Lambda(s)\,\mathrm{d}s\big)$, and $\beta=\delta/2$. Hence, the assumption on η implies that

$$\sup_{s\in(0,T)}\|\varrho(s)\|^2 + 2\varepsilon^2\int_0^T\|\nabla\varrho(s)\|^2\,\mathrm{d}s \leq 4\eta^2\exp\Big(\int_0^T\mu_\Lambda\,\mathrm{d}s\Big).$$

Applications of the triangle inequality yield the asserted estimates. □

4. Application to finite element methods

We next discuss how Theorem 3.3 can be specified for various spatial discretizations of (1). Owing to the employed elliptic reconstruction, this reduces to a posteriori error estimates for elliptic equations and we assume that we are given a posteriori error estimators for the approximation error of the Poisson problem in various norms. For the discussion of the construction of a computable function $\overline{\Lambda}_{AC}$ that fulfills the requirements of Theorem 3.3 we refer the reader to [BMO09b].

Assumption ($\mathbf{EST}_{L^p}$). *The subspace $\mathbb{V}_h$ and the bilinear form $a_h:\mathbb{V}_h\times\mathbb{V}_h\to\mathbb{R}$ satisfy* assumption (EST_{L^p}) *if for all $\xi\in L^2(\Omega)$ with $\int_\Omega\xi\,\mathrm{d}x=0$ the following holds: If $w\in\mathbb{V}$ and $W\in\mathbb{V}_h$ are such that $\int_\Omega W\,\mathrm{d}x=\int_\Omega w\,\mathrm{d}x$ and*

$$a(w,v)=(\xi,v) \qquad \textit{and} \qquad a_h(W,V)=(\xi,V)$$

for all $v \in \mathbb{V}$ and all $V \in \widehat{\mathbb{V}}_h$ for some nontrivial subspace $\widehat{\mathbb{V}}_h \subseteq \mathbb{V}_h$ then for $p = 2$ and $p = \infty$ we have

$$\|w - W\|_{L^p(\Omega)} \leq \mathcal{E}_{L^p}(W, \xi; \widehat{\mathbb{V}}_h)$$

for a computable quantity $\mathcal{E}_{L^p}(W, \xi; \widehat{\mathbb{V}}_h)$.

Remark 4.1. *For lowest order conforming methods assumption (EST_{L^p}) is well established provided that the Laplace operator is H^2 regular in Ω, cf., e.g.,* [Noc95, Ver96, DDP00, NSSV06]. *In particular, we may choose*

$$\mathcal{E}_{L^2}(W, \xi; \widehat{\mathbb{V}}_h) = C_2\big(\|h^2_{\widehat{\mathcal{T}}_h}(\Delta_{\widehat{\mathcal{T}}_h} W + \xi)\| + \|h^{3/2}_{\widehat{\mathcal{T}}_h}[\nabla W \cdot n_{\widehat{\mathcal{F}}_h}]\|_{L^2(\cup\widehat{\mathcal{F}}_h)}\big),$$
$$\mathcal{E}_{L^\infty}(W, \xi; \widehat{\mathbb{V}}_h) = C_\infty \log(\underline{h}_{\widehat{\mathcal{T}}_h})^{4/3}\big(\|h_{\widehat{\mathcal{T}}_h}(\Delta_{\widehat{\mathcal{T}}_h} W + \xi)\|_{L^\infty(\Omega)} + \|h^{1/2}_{\widehat{\mathcal{T}}_h}[\nabla W \cdot n_{\widehat{\mathcal{F}}_h}]\|_{L^\infty(\cup\widehat{\mathcal{F}}_h)}\big),$$

if $\widehat{\mathbb{V}}_h$ is the lowest order conforming finite element space related to the regular triangulation $\widehat{\mathcal{T}}_h$ with meshsize function $h_{\widehat{\mathcal{T}}_h}$ whose minimum is $\underline{h}_{\widehat{\mathcal{T}}_h}$ and with interelement sides contained in $\widehat{\mathcal{F}}_h$; $\Delta_{\widehat{\mathcal{T}}_h}$ denotes the elementwise application of the Laplace operator on $\widehat{\mathcal{T}}_h$.

Another assumption is needed that guarantees compatibility of successive discretizations of the Laplace operator.

Assumption (COMP). *For $j = 1, 2, ..., J$ there exists a subspace $\mathbb{V}_h^{j-1/2} \subseteq \mathbb{V}_h^{j-1} \cap \mathbb{V}_h^j$ and a bilinear form*

$$a_h^{j-1/2} : \big(\mathbb{V}_h^{j-1} + \mathbb{V}_h^j\big) \times \mathbb{V}_h^{j-1/2} \to \mathbb{R}$$

such that the pair $(\mathbb{V}_h^{j-1/2}, a_h^{j-1/2})$ satisfies assumption (EST_{L^p}) and

$$a_h^{j-1/2}(W_1 + W_2, V) = a_h^{j-1}(W_1, V) + a_h^j(W_2, V)$$

for all $W_1 \in \mathbb{V}_h^{j-1}$, $W_2 \in \mathbb{V}_h^j$, and $V \in \mathbb{V}_h^{j-1/2}$.

Remarks 4.2. *(i) Requiring that the pair $(\mathbb{V}_h^{j-1/2}, a_h^{j-1/2})$ satisfies assumption (EST_{L^p}) avoids that assumption (COMP) is trivially satisfied with the choice $\mathbb{V}_h^{j-1/2} = \{0\}$.*
(ii) Assumption (COMP) is trivially satisfied if the same spatial discretization that fulfills (EST_{L^p}) is used in each time step.
(iii) For lowest order conforming methods assumption (COMP) is satisfied provided that there exists a common coarsening $\widehat{\mathcal{T}}_h$ of the triangulations $\mathcal{T}_h^{j-1}$ and $\mathcal{T}_h^j$ that define the spaces $\mathbb{V}_h^{j-1}$ and $\mathbb{V}_h^j$, respectively. In this case an efficient choice for $\mathbb{V}_h^{j-1/2}$ is the finite element space defined through the finest common coarsening of $\mathcal{T}_h^j$ and $\mathcal{T}_h^{j-1}$, cf. [LM06].

In the following, $h_j \in L^\infty(\Omega)$ denotes for $j = 1, 2, ..., J$ a positive meshsize function related to the space $\mathbb{V}_h^j$. In particular, we assume that there exists a constant $C_{C\ell} > 0$ such that for every $v \in \mathbb{V}$ and $j = 1, 2, ..., J$ we have the Clément type quasi-interpolation estimate

$$\inf_{V \in \mathbb{V}_h^j} \|h_j^{-1}(v - V)\| \leq C_{C\ell}\|\nabla v\|.$$

Given any sequence $(a^j)_{j=0,1,...,J}$ we set

$$d_t a^j := \tau_j^{-1}(a^j - a^{j-1})$$

for $j = 1, 2, ..., J$. If (COMP) is satisfied and all involved bilinear forms fulfill (EST_{L^p}) then we immediately obtain bounds for the functional Γ. Recall that $\xi_h^j = -\Delta_h^j U^j$ and $-\Delta w^j = \xi_h^j$ for $j = 0, 1, ..., J$.

Proposition 4.3 (Computable bounds). *Suppose that the pairs $(\mathbb{V}_h^j, a_h^j)$, $j = 0, 1, ..., J$, satisfy (EST_{L^p}) and that assumption (COMP) holds. Then,*
(a) we have

$$\langle \Gamma(w, U; t), v\rangle \le \big(\eta_{\Gamma,0}^{\mathrm{t}}(t) + \eta_{\Gamma,0}^{\mathrm{s}}(t)\big)\|v\| + \eta_{\Gamma,1}^{\mathrm{c}}(t)\|\nabla v\|$$

with $\eta_{\Gamma,0}^{\mathrm{t}}$, $\eta_{\Gamma,0}^{\mathrm{s}}$ and $\eta_{\Gamma,1}^{\mathrm{c}}$ defined for $t \in (t_{j-1}, t_j]$, $j = 1, 2, ..., J$ by

$$\begin{aligned}
\eta_{\Gamma,0}^{\mathrm{t}}(t) &:= \|\Delta_h^{j-1}U^{j-1} - \Delta_h^j U^j\| + \varepsilon^{-2}C_{f',I_j}\|U^{j-1} - U^j\|,\\
\eta_{\Gamma,0}^{\mathrm{s}}(t) &:= \mathcal{E}_{L^2}(d_t U^j, d_t\xi_h^j; \mathbb{V}_h^{j-1/2}) + \varepsilon^{-2}C_{f',I_j}\max_{k=j-1,j}\mathcal{E}_{L^2}(U^k, \xi_h^k; \mathbb{V}_h^k),\\
\eta_{\Gamma,1}^{\mathrm{c}}(t) &:= C_{C\ell}\tau_j^{-1}\|h_j(U^{j-1} - P_h^j U^{j-1})\| + C_{C\ell}\varepsilon^{-2}\|h_j(f(U^j) - P_h^j f(U^j))\|
\end{aligned}$$

where for $\ell = 1, 2$ we set

$$C_{f^{(\ell)},I_j} := \|f^{(\ell)}\|_{L^\infty(B_{d_j})}, \quad d_j := \max_{k=j-1,j}\big(\|U^k\|_{L^\infty(\Omega)} + \mathcal{E}_{L^\infty}(U^k, \xi_h^k; \mathbb{V}_h^k)\big); \tag{11}$$

(b) we have

$$\sup_{s\in(0,T)}\|(U - w)(s)\| \le \max_{j=0,1,...,J}\mathcal{E}_{L^2}(U^j, \xi_h^j; \mathbb{V}_h^j)$$

and

$$\|u_0 - w^0\| \le \|u_0 - U^0\| + \mathcal{E}_{L^2}(U^0, \xi_h^0; \mathbb{V}_h^0);$$

(c) with C_{f'',I_j} from (11) we have for $t \in [t_{j-1}, t_j]$, $j = 1, 2, ..., J$ that

$$\|f'(w(t)) - f'(U(t))\|_{L^\infty(\Omega)} \le C_{f'',I_j}\max_{k=j-1,j}\mathcal{E}_{L^\infty}(U^k, \xi_h^k; \mathbb{V}_h^k);$$

(d) we have

$$\begin{aligned}
&\sup_{s\in(0,T)}\|g(w(s))\|_{L^\infty(\Omega)}\\
&\quad\le \max_{j=0,1,...,J}\big(\|g(U^j)\|_{L^\infty(\Omega)} + C_{g',I_j}\|U^j - U^{j-1}\|_{L^\infty(\Omega)} + C_{g',I_j}\max_{k=j-1,j}\mathcal{E}_{L^\infty}(U^k, \xi_h^k; \mathbb{V}_h^k)\big),
\end{aligned}$$

where C_{g',I_j} is defined as in (11) with $f^{(\ell)}$ replaced by g'.

Proof. (a) Given $t \in (t_{j-1}, t_j)$, $j = 1, 2, ..., J$, we recast the functional Γ as

$$\begin{aligned}
\langle \Gamma(w, U; t), v\rangle = \langle \partial_t(w - U), v\rangle + (\nabla(w - w^j), \nabla v) + \tau_j^{-1}(U^{j-1} - P_h^j U^{j-1}, v) + \varepsilon^{-2}(f(w) - f(U), v)\\
+ \varepsilon^{-2}(f(U) - f(U^j), w) + \varepsilon^{-2}(f(U^j) - P_h^j f(U^j), w) =: T_1 + T_2 + ... + T_6
\end{aligned}$$

and split the proof of (a) into three parts.
Part 1: time discretization residuals. Using $\ell_{j-1}(t) = 1 - \ell_j(t) \le 1$ and the definitions of w^j and w^{j-1} we have

$$\begin{aligned}
T_2 &= (\nabla[\ell_{j-1}(t)w^{j-1} + \ell_j(t)w^j - w^j], \nabla v) = \ell_{j-1}(t)(\nabla(w^{j-1} - w^j), \nabla v)\\
&= \ell_{j-1}(t)(-\Delta_h^{j-1}U^{j-1} + \Delta_h^j U^j, v) \le \|\Delta_h^{j-1}U^{j-1} - \Delta_h^j U^j\|\|v\|.
\end{aligned}$$

Similarly, using the identity

$$f(U) - f(U^j) = \Big(\int_0^1 f'(rU + (1 - r)U^{j-1})\,\mathrm{d}r\Big)(U - U^j)$$

we derive the estimate

$$T_5 = \varepsilon^{-2}(f(U) - f(U^j), v) \le \varepsilon^{-2}C_{f',I_j}\|U - U^j\|\|v\| \le \varepsilon^{-2}C_{f',I_j}\|U^{j-1} - U^j\|\|v\|.$$

Part 2: coarsening and oscillation residuals. For the contributions T_3 and T_6 we get for arbitrary $V \in \mathbb{V}_h^j$

$$\begin{aligned} T_3 + T_6 &= \tau_j^{-1}(U^{j-1} - P_h^j U^{j-1}, v - V) + \varepsilon^{-2}(f(U^j) - P_h^j f(U^j), v - V) \\ &\leq \big(\tau_j^{-1}\|h_j(U^{j-1} - P_h^j U^{j-1})\| + \varepsilon^{-2}\|h_j(f(U^j) - P_h^j f(U^j))\|\big)\|h_j^{-1}(v - V)\|. \end{aligned}$$

A minimization over V leads to the contribution $\eta_{\Gamma,1}^{\mathrm{c}}(t)$.
Part 3: space discretization residuals. Noting that

$$a(d_t w^j, v) = (d_t \xi_h^j, v)$$

for all $v \in \mathbb{V}$ and that owing to (COMP)

$$a_h^{j-1/2}(d_t U^j, V) = (d_t \xi_h^j, V)$$

for all $V \in \mathbb{V}_h^{j-1/2}$ we deduce with (EST_{L^p}) that

$$T_1 = (d_t w^j - d_t U^j, v) \leq \|d_t w^j - d_t U^j\|\|v\| \leq \mathcal{E}_{L^2}(d_t U^j, d_t \xi_h^j; \mathbb{V}_h^{j-1/2})\|v\|.$$

Moreover, we have

$$\begin{aligned} T_4 = \varepsilon^{-2}(f(w) - f(U), v) &\leq \varepsilon^{-2} C_{f',I_j}\|w - U\|\|v\| \\ &\leq \varepsilon^{-2} C_{f',I_j} \max_{k=j-1,j}\|w^k - U^k\|\|v\| \leq \varepsilon^{-2} C_{f',I_j} \max_{k=j-1,j} \mathcal{E}_{L^2}(U^k, \xi_h^k; \mathbb{V}_h^k)\|v\|. \end{aligned}$$

A combination of the estimates implies (a). The proofs of (b), (c), and (d) are analogous. □

Remark 4.4. *The computation of $P_h^j U^{j-1}$ and $P_h^j f(U^j)$ in the evaluation of $\eta_{\Gamma,1}^{\mathrm{c}}$ can be avoided by using a modified scheme which computes for $j = 1, 2, ..., J$ the function $U^j \in \mathbb{V}_h^j$ such that*

$$\tau_j^{-1}(U^j - I_h^j U^{j-1}, V) + a_h^j(U^j, V) = -\varepsilon^{-2}(I_h^j f(U^j), V)$$

for all $V \in \mathbb{V}_h^j$. Here, $I_h^j : C(\overline{\Omega}) \to \mathbb{V}_h^j$ is an appropriate mesh-transfer operator, e.g., the nodal interpolation operator related to $\mathbb{V}_h^j$ in case of a conforming method. The quantity $\eta_{\Gamma,1}^{\mathrm{c}}(t)$ of Proposition 4.3 is then substituted by

$$\eta_{\Gamma,0}^{\mathrm{c}}(t) = \tau_j^{-1}\|U^{j-1} - I_h^j U^{j-1}\| + \varepsilon^{-2}\|f(U^j) - I_h^j f(U^j)\|$$

and the third line of the error estimate stated in the introduction is is interchanged with

$$\sum_{j=1}^{J} \tau_j\big(\tau_j^{-1}\|U^{j-1} - I_h^j U^{j-1}\| + \varepsilon^{-2}\|f(U^j) - I_h^j f(U^j)\|\big).$$

A significantly weaker version of (COMP) can be imposed. If the property (COMP') stated below is assumed then the quantity $\mathcal{E}_{L^2}(d_t U^j, d_t \xi_h^j; \mathbb{V}_h^{j-1/2})$ appearing in the estimate of Proposition 4.3 needs to be replaced by

$$\mathcal{C}_{L^2}(\tau_j^{-1} U^{j-1}, \tau_j^{-1} U^j, d_t \xi_h^j; \mathbb{V}_h^{j-1}, \mathbb{V}_h^j).$$

Assumption (COMP'). *Suppose that for $\xi^{j-1}, \xi^j \in L^2(\Omega)$ satisfying $\int_\Omega \xi^{j-1}\,\mathrm{d}x = \int_\Omega \xi^j\,\mathrm{d}x = 0$ and $W^{j-1} \in \mathbb{V}_h^{j-1}$ and $W^j \in \mathbb{V}_h^j$ we have that*

$$a_h^{j-1}(W^{j-1}, V^{j-1}) = (\xi^{j-1}, V^{j-1}) \qquad \text{and} \qquad a_h^j(W^j, V^j) = (\xi^j, V^j)$$

for all $V^{j-1} \in \mathbb{V}_h^{j-1}$ and $V^j \in \mathbb{V}_h^j$. Then

$$\|W^j - W^{j-1}\| \leq \mathcal{C}_{L^2}(V^{j-1}, V^j, \xi^j - \xi^{j-1}; \mathbb{V}_h^{j-1}, \mathbb{V}_h^j)$$

with a computable quantity $\mathcal{C}_{L^2}(V^{j-1}, V^j, \xi^j - \xi^{j-1}; \mathbb{V}_h^{j-1}, \mathbb{V}_h^j)$.

Appropriate error estimators are required to bound the approximation error in an extension of the seminorm of $L^2(0,T;H^1(\Omega))$. The following assumption and the conditions of Proposition 4.6 below hold for a large class of conforming and nonconforming finite element methods, e.g., with the broken H^1 seminorm on a partition of Ω that is a common refinement of all employed triangulations or partitions that define the spaces $\mathbb{V}_h^j$, $j = 0, 1, ..., J$.

Assumption ($\mathbf{EST}_{H^1}$). *The subspace $\mathbb{V}_h$, the bilinear form $a_h : \mathbb{V}_h \times \mathbb{V}_h \to \mathbb{R}$, and the seminorm $|||\cdot|||$ defined on the span of $\mathbb{V} \cup \mathbb{V}_h$ satisfy* condition (EST_{H^1}) *if for all $\xi \in L^2(\Omega)$ with $\int_\Omega \xi \,\mathrm{d}x = 0$ the following holds: If $w \in \mathbb{V}$ and $W \in \mathbb{V}_h$ are such that $\int_\Omega W \,\mathrm{d}x = \int_\Omega w \,\mathrm{d}x$ and*

$$a(w,v) = (\xi, v) \quad \text{and} \quad a_h(W,V) = (\xi, V)$$

for all $v \in \mathbb{V}$ and all $V \in \mathbb{V}_h$ then we have

$$|||w - W||| \leq \mathcal{E}_{H^1}(W, \xi; \mathbb{V}_h)$$

for a computable quantity $\mathcal{E}_{H^1}(W, \xi; \mathbb{V}_h)$.

Remark 4.5. *Assumption (EST_{H^1}) is well established for conforming methods, cf., e.g.,* [Ver96], *and also holds for mixed, nonconforming, and discontinuous Galerkin methods, cf., e.g.,* [Car97, RW03, Ain05] *with appropriate choices of extensions of the H^1 seminorm.*

Proposition 4.6 (Energy norm estimate)**.** *If (EST_{H^1}) is satisfied for all triples $(\mathbb{V}_h^j, a_h^j, |||\cdot|||)$, $j = 0, 1, ..., J$, with the same seminorm $|||\cdot|||$ then*

$$\int_0^T |||(U-w)(s)|||^2 \,\mathrm{d}s \leq \sum_{j=1}^{J} \frac{\tau_j}{2} \big(\mathcal{E}_{H^1}^2(U^{j-1}, \xi_h^{j-1}; \mathbb{V}_h^{j-1}) + \mathcal{E}_{H^1}^2(U^j, \xi_h^j; \mathbb{V}_h^j) \big).$$

Proof. For every $j = 1, 2, ..., J$ we deduce with Jensen's inequality that

$$\begin{aligned} |||(U-w)(s)|||^2 &= \big(\ell_{j-1}(s) |||U^{j-1} - w^{j-1}||| + (1 - \ell_{j-1}(s)) |||U^j - w^j||| \big)^2 \\ &\leq \ell_{j-1}(s) |||U^{j-1} - w^{j-1}|||^2 + (1 - \ell_{j-1}(s)) |||U^j - w^j|||^2. \end{aligned}$$

Noting $\int_{t_{j-1}}^{t_j} \ell_{j-1}(s) \,\mathrm{d}s = \tau_j/2$ and incorporating (EST_{H^1}) implies the assertion. □

5. Numerical experiments

We discuss our error estimate with numerical experiments leading to a generic topological change in an evolution process governed by (1) in two dimensions. Motivated by our estimates we employ the following strategy to efficiently simulate Allen-Cahn processes with a fixed time-step size. We let $\mathcal{I}_\mathcal{T}$ denote the nodal interpolation operator associated to the lowest order conforming finite element space defined by a triangulation $\mathcal{T}$.

Algorithm (ADAPT). Given a tolerance $\sigma > 0$ iterate for $j = 1, 2, ..., J$ the following steps:

(a) **Coarsen** elements in $\mathcal{T}_C \subseteq \mathcal{T}_{j-1}$ to obtain a triangulation $\mathcal{T}_{j,0}$ with

$$\eta_{\Gamma,1}^{\mathrm{c},1}(t_j) := \tau_j^{-1} \| h_j (U^{j-1} - \mathcal{I}_{\mathcal{T}_{j,k}} U^{j-1}) \| \leq \frac{\sigma}{10}.$$

Set $k := 0$.

(b) **Compute** $U^{j,k} \in \mathbb{V}_h^{j,k}$ such that for all $V \in \mathbb{V}_h^{j,k}$ we have

$$\tau^{-1}(U^{j,k} - \mathcal{I}_{\mathcal{T}_{j,k}} U^{j-1}, V) + (\nabla U^{j,k}, \nabla V) = -\varepsilon^{-2} (\mathcal{I}_{\mathcal{T}_{j,k}} f(U^{j,k}), V).$$

(c) **Refine** elements $K \in \mathcal{T}_{j,k}$ for which

$$\varepsilon^{-2} h_K^2 \|\tau_j^{-1}(U^{j,k} - \mathcal{I}_{\mathcal{T}_{j,k}} U^{j-1}) + \varepsilon^{-2} \mathcal{I}_{\mathcal{T}_{j,k}} f(U^{j,k})\| + h_K^{3/2} \|[\nabla U^{j,k} \cdot n_{\mathcal{F}_{j,k}}]\|_{L^2(K \cap (\cup \mathcal{F}_{j,k}))}$$
$$=: \eta^{\mathrm{s},2}_{\Gamma,0}(t_j)|_K \geq (1/2) \max_{K' \in \mathcal{T}_{j,k}} \eta^{\mathrm{s},2}_{\Gamma,0}(t_j)|_{K'},$$

set $k := k+1$, and go to (b) if $\sum_{K \in \mathcal{T}_{j,k}} \left(\eta^{\mathrm{s},2}_{\Gamma,0}(t_j)|_K\right)^2 \geq \sigma^2$.

(d) **Update** $U^j := U^{j,k}$, set $j := j+1$, and go to (a).

We tried Algorithm (ADAPT) with initial data that define a circular initial interface.

Example 5.1 (Vanishing particle). *Let $\Omega := (-2,2)^2$, set $r := 1$, and define $d(x) := |x| - r$ for $x \in \Omega$. For given $\varepsilon > 0$ and $x \in \Omega$ let*

$$u_0(x) := -\tanh\left(d(x)/(\sqrt{2}\varepsilon)\right).$$

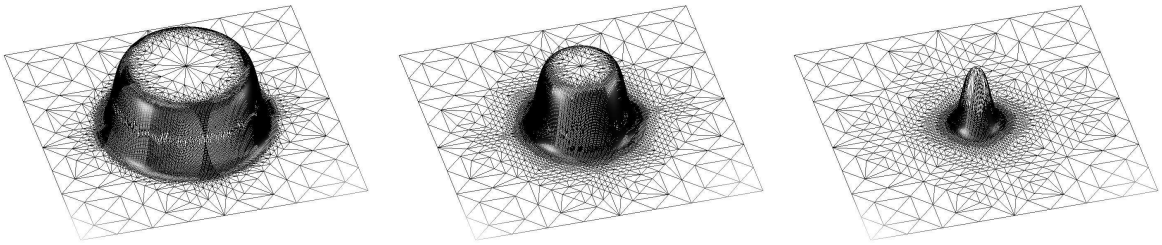

FIGURE 1. Evolving interface and adaptively refined and coarsened triangulations for $t = 0$, 0.31, and 0.48 obtained with Algorithm (ADAPT) in Example 5.1 with $\varepsilon = 1/16$ and $\sigma = \varepsilon/10$.

Snapshots of the evolution defined by the initial data of Example 5.1 for $\varepsilon = 1/16$ together with adaptively generated triangulations are shown in Figure 1. The approximations were obtained with the uniform time-step size $\tau = \varepsilon^3/16$ and the parameter $\sigma = \varepsilon/10$. We see that the interface Γ_t undergoes a topological change at $t \approx 0.49$ when the particle vanishes. The employed adaptive strategy refines the grid locally around the interface Γ_t where large gradients occur and coarsens the triangulations when the interface has advanced.

For $\varepsilon = 2^{-\ell}$, $\ell = 2, 3, ..., 6$ we plotted in Figure 2 the numerically computed eigenvalue $\Lambda_{AC}(t)$ (left plot) as a function of t and the integral over $(0,t)$ of its positive part (right plot), i.e., the functions

$$t \mapsto \Lambda_{AC}(t), \qquad t \mapsto \int_0^t \Lambda^+_{AC}(s)\,\mathrm{d}s.$$

The results of the experiment show that a uniform bound for $\Lambda_{AC}(t)$ breaks down when the topological change occurs and we observe $\max_{t\in(0,T)} \Lambda_{AC}(t) \sim \varepsilon^{-2}$. In contrast, the integrated eigenvalue grows logarithmically in ε^{-1}, i.e., we have

$$\int_0^T \Lambda^+_{AC}(t)\,\mathrm{d}t \sim C_0 + \log(\varepsilon^{-\kappa}). \tag{12}$$

Therefore, robust a posteriori error estimation in $L^\infty(0,T;L^2(\Omega))$ is possible past topological changes in this prototypical example.

For fixed $\varepsilon = 1/8$ and decreasing tolerances $\sigma = 2^{-\ell}\varepsilon/10$, $\ell = 0, 1, 2, 3$, we plotted in Figure 3 the error estimator $\eta_{L^\infty(L^2)}$ defined through the approximate solution obtained with Algorithm (ADAPT) as a function of $t \in [0, 0.6]$ and the number of degrees of freedom required to reduce the spatial discretization residuals below the tolerance σ. Consequently, we observe a

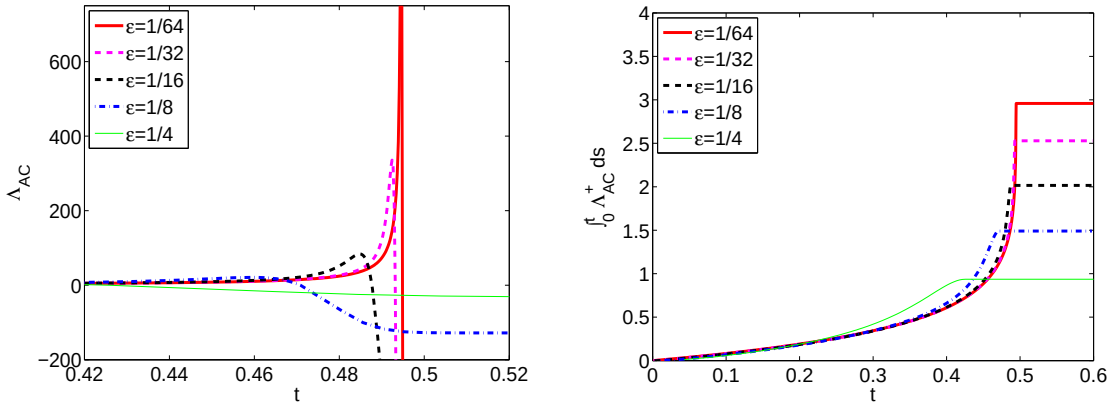

FIGURE 2. Approximated eigenvalue $\Lambda_{AC}(t)$ in Example 5.1 as a function of $t \in [0.42, 0.52]$ (left) and the integral of its positive part over $(0, t)$ as a function of $t \in [0, 0.6]$ (right). The eigenvalue grows like ε^{-2} at the time of the topological change while its temporal integral only grows logarithmically in ε^{-1}.

linear relation between $\eta_{L^\infty(L^2)}$ and σ. The numbers of degrees of freedom shown in the right plot of Figure 3 depend inverse proportionally on σ, i.e., twice as many degrees of freedom are required to decrease the approximation error by a factor $1/2$. This relation corresponds to the quadratic scaling $\eta_{L^\infty(L^2)} \sim h^2$ and the fact that the theoretical mesh-size is $h^2 = N_h^{-1}$ for the number of nodes N_h in a two-dimensional triangulation $\mathcal{T}_h$. To illustrate the significant increase in efficiency of the proposed adaptive method, we checked that to decrease the error estimator below the largest tolerance $\sigma = \varepsilon/10$ using uniform grids, roughly eight times as many nodes are required as in the case of an adaptive approach. We remark that in order to guarantee the mesh compatibility condition (COMP) we either refined or coarsened the mesh in each time step. Once the particle has disappeared at $t \approx 0.49$, the grid is maximally coarsened.

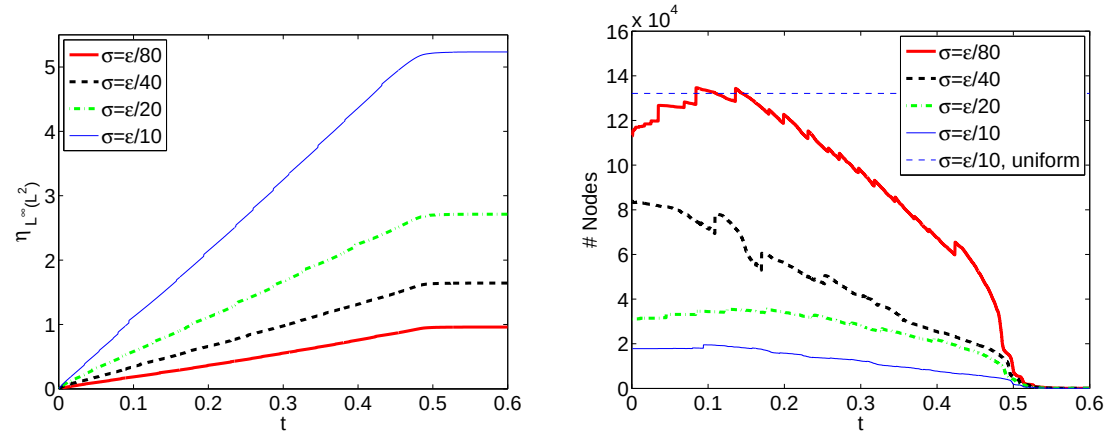

FIGURE 3. Estimator $\eta_{L^\infty(L^2)}$ as function of $t \in [0, 0.6]$ (left) and degrees of freedom of adaptively generated triangulations needed to reduce spatial discretization residuals below the tolerance σ (right) for fixed $\varepsilon = 1/8$ and $\tau = 0.00024$.

To verify the expected scaling properties of the estimators $\eta_{L^\infty(L^2)}$ and $\eta_{L^2(H^1)}$ we ran experiments with uniform triangulations in which either ε or h was fixed. The results for fixed $\varepsilon = 1/8$ and decreasing discretization parameters $h = 2^{-\ell}$, $\ell = 5, 6, 7$ and $\tau = h^2/32$ shown in Figure 4 confirm that we have $\eta_{L^\infty(L^2)} \sim h^2$ and $\eta_{L^2(H^1)} \sim h$. These experimental convergence rates can be read from the slopes of the curves shown in the left plot of Figure 6 where we displayed

the total estimators at the final time $t = 0.6$ versus the mesh-size h of the underlying uniform triangulations with a logarithmic scaling used for both axes. We also observe in Figure 4 the linear accumulation of contributions to $\eta_{L^\infty(L^2)}$ while the estimator $\eta_{L^2(H^1)}$ grows proportionally to $t^{1/2}$ in time.

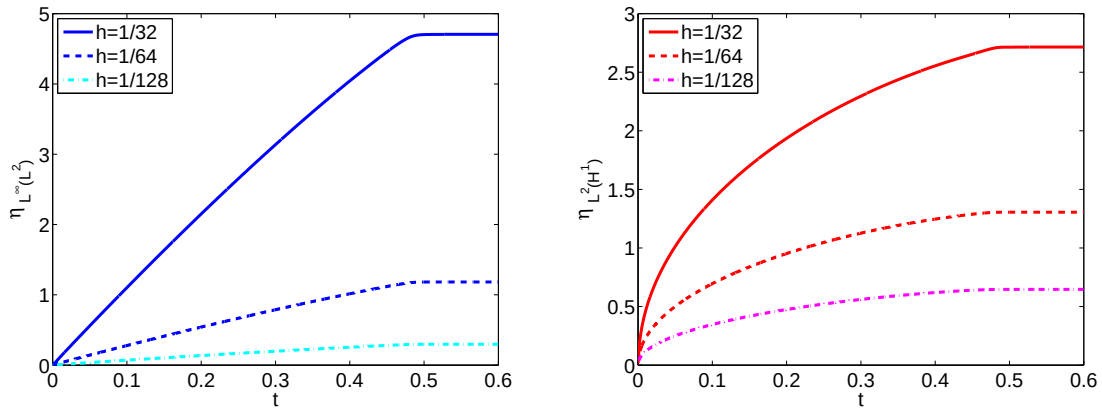

FIGURE 4. Estimators $\eta_{L^\infty(L^2)}$ and $\eta_{L^2(H^1)}$ as functions of $t \in [0, 0.6]$ for $h = 1/32$, $1/64$, $1/128$ and $\tau = h^2/32$ and fixed $\varepsilon = 1/8$.

We ran the same experiment with a fixed uniform triangulation of mesh-size $h = 1/64$ and fixed time-step size $\tau = 0.00003$ but varying $\varepsilon = 2^{-\ell}$, $\ell = 2, 3, 4, 5$. The corresponding values for $\eta_{L^\infty(L^2)}$ and $\eta_{L^2(H^1)}$ as functions of $t \in [0, 0.6]$ are shown in the left and right plot of Figure 5, respectively. The graphs reveal a polynomial dependence on ε^{-1} and the double-logarithmic scaling used in the right plot of Figure 6 shows that we have $\eta_{L^\infty(L^2)} \sim \varepsilon^{-7/2}$ and $\eta_{L^2(H^1)} \sim \varepsilon^{-5/2}$ in this example. This can also be understood directly from the definitions of the estimators since $\|D^2 u(t)\| \leq \varepsilon^{-3/2}$ if $u(t)$ represents a developed interface.

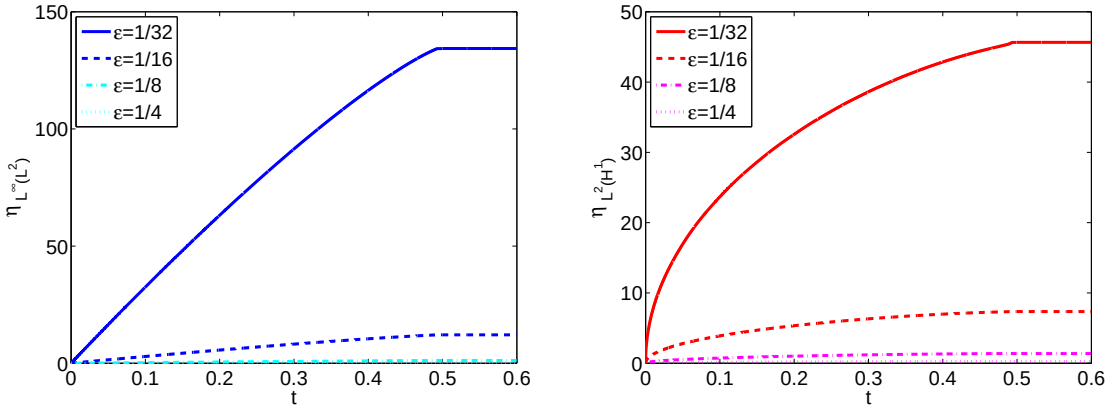

FIGURE 5. Estimators $\eta_{L^\infty(L^2)}$ and $\eta_{L^2(H^1)}$ as functions of $t \in [0, 0.6]$ for $\varepsilon = 1/4$, $1/8$, $1/16$, $1/32$ and fixed mesh-size $h = 1/64$, $\tau = 0.00003$.

Although the proposed estimator $\eta_{L^\infty(L^2)}$ has a worse dependence on ε^{-1} than $\eta_{L^2(H^1)}$, its quadratic convergence in h makes it superior since a reasonable resolution of interfaces requires $h \ll \varepsilon$.

Acknowledgements: The authors acknowledge stimulating discussions with Christoph Ortner.

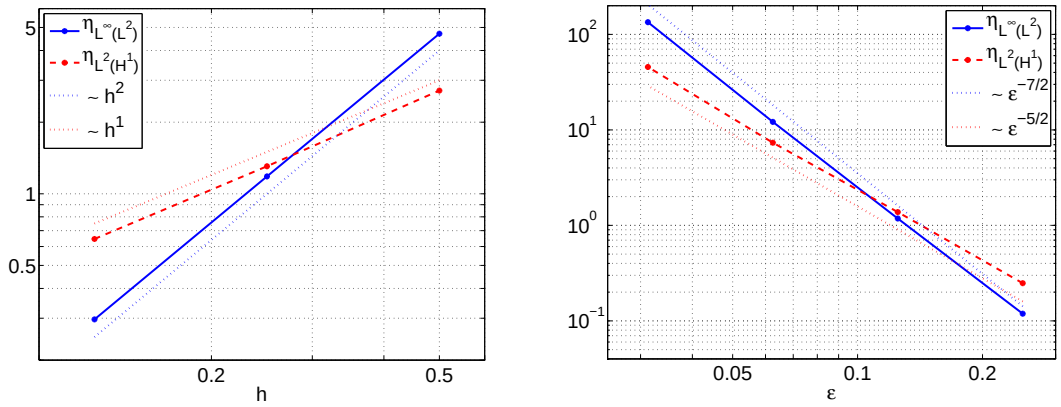

FIGURE 6. Dependence of $\eta_{L^\infty(L^2)}$ and $\eta_{L^2(H^1)}$ on h (left) and ε (right).

REFERENCES

[AC79] Sam Allen and John Cahn. A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. *Acta Metall.*, 27:1084–1095, 1979.

[AF93] Nicholas D. Alikakos and Giorgio Fusco. The spectrum of the Cahn-Hilliard operator for generic interface in higher space dimensions. *Indiana Univ. Math. J.*, 42(2):637–674, 1993.

[Ain05] Mark Ainsworth. A synthesis of a posteriori error estimation techniques for conforming, non-conforming and discontinuous Galerkin finite element methods. In *Recent advances in adaptive computation*, volume 383 of *Contemp. Math.*, pages 1–14. Amer. Math. Soc., Providence, RI, 2005.

[Bar05] Sören Bartels. A posteriori error analysis for time-dependent Ginzburg-Landau type equations. *Numer. Math.*, 99(4):557–583, 2005.

[BM08] Sören Bartels and Rüdiger Müller. Error controlled local resolution of evolving interfaces for generalized Cahn-Hilliard equations. Preprint available at `www.bartels.ins.uni-bonn.de`, 2008.

[BMO09a] Sören Bartels, Rüdiger Müller, and Christoph Ortner. Robust a posteriori error estimates for the approximation of Cahn-Hilliard equations past topological changes. In preparation, 2009.

[BMO09b] Sören Bartels, Rüdiger Müller, and Christoph Ortner. Robust a priori and a posteriori error analysis for the approximation of Allen-Cahn and Ginzburg-Landau equations past topological changes. Preprint available at `www.bartels.ins.uni-bonn.de`, 2009.

[Bra78] Kenneth A. Brakke. *The motion of a surface by its mean curvature*, volume 20 of *Mathematical Notes*. Princeton University Press, Princeton, N.J., 1978.

[Car97] Carsten Carstensen. A posteriori error estimate for the mixed finite element method. *Math. Comp.*, 66(218):465–476, 1997.

[Che94] Xinfu Chen. Spectrum for the Allen-Cahn, Cahn-Hilliard, and phase-field equations for generic interfaces. *Comm. Partial Differential Equations*, 19(7-8):1371–1395, 1994.

[DDP00] Enzo Dari, Ricardo G. Durán, and Claudio Padra. Maximum norm error estimators for three-dimensional elliptic problems. *SIAM J. Numer. Anal.*, 37(2):683–700 (electronic), 2000.

[dMS95] Piero de Mottoni and Michelle Schatzman. Geometrical evolution of developed interfaces. *Trans. Amer. Math. Soc.*, 347(5):1533–1589, 1995.

[EJ95a] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM J. Numer. Anal.*, 32(3):706–740, 1995.

[EJ95b] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems. IV. Nonlinear problems. *SIAM J. Numer. Anal.*, 32(6):1729–1749, 1995.

[FP03] Xiaobing Feng and Andreas Prohl. Numerical analysis of the Allen-Cahn equation and approximation for mean curvature flows. *Numer. Math.*, 94(1):33–65, 2003.

[FW05] Xiaobing Feng and Hai-jun Wu. A posteriori error estimates and an adaptive finite element method for the Allen-Cahn equation and the mean curvature flow. *J. Sci. Comput.*, 24(2):121–146, 2005.

[GL08] Emmanuil H. Georgoulis and Omar Lakkis. A posteriori error control for discontinuous Galerkin methods for parabolic problems, 2008.

[Ilm93] Tom Ilmanen. Convergence of the Allen-Cahn equation to Brakke's motion by mean curvature. *J. Differential Geom.*, 38(2):417–461, 1993.

[IT79] Alexander D. Ioffe and Vladimir M. Tihomirov. *Theory of extremal problems*, volume 6 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1979.

[KNS04] Daniel Kessler, Ricardo H. Nochetto, and Alfred Schmidt. A posteriori error control for the Allen-Cahn problem: circumventing Gronwall's inequality. *M2AN Math. Model. Numer. Anal.*, 38(1):129–142, 2004.

[LM06] Omar Lakkis and Charalambos Makridakis. Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems. *Math. Comp.*, 75(256):1627–1658 (electronic), 2006.

[MN03] Charalambos Makridakis and Ricardo H. Nochetto. Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, 41(4):1585–1594 (electronic), 2003.

[Noc95] Ricardo H. Nochetto. Pointwise a posteriori error estimates for elliptic problems on highly graded meshes. *Math. Comp.*, 64(209):1–22, 1995.

[NSSV06] Ricardo H. Nochetto, Alfred Schmidt, Kunibert G. Siebert, and Andreas Veeser. Pointwise a posteriori error estimates for monotone semi-linear equations. *Numer. Math.*, 104(4):515–538, 2006.

[RW03] Béatrice Rivière and Mary F. Wheeler. A posteriori error estimates for a discontinuous Galerkin method applied to elliptic problems. Log number: R74. *Comput. Math. Appl.*, 46(1):141–163, 2003.

[Ver96] Rüdiger Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques.* Wiley & Teubner, 1996.

[Whe73] Mary F. Wheeler. A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations. *SIAM J. Numer. Anal.*, 10:723–759, 1973.