ESAIM: M2AN 43 (2009) 1003–1026
DOI: 10.1051/m2an/2009015

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

A POSTERIORI ESTIMATES FOR THE CAHN–HILLIARD EQUATION WITH OBSTACLE FREE ENERGY *

L'UBOMÍR BAŇAS[1] AND ROBERT NÜRNBERG[2]

Abstract. We derive *a posteriori* estimates for a discretization in space of the standard Cahn–Hilliard equation with a double obstacle free energy. The derived estimates are robust and efficient, and in practice are combined with a heuristic time step adaptation. We present numerical experiments in two and three space dimensions and compare our method with an existing heuristic spatial mesh adaptation algorithm.

Mathematics Subject Classification. 65M60, 65M15, 65M50, 35K55.

Received January 14, 2008. Revised November 28, 2008.
Published online June 12, 2009.

1. INTRODUCTION

In this paper we derive spatial *a posteriori* error estimates for a piece-wise linear finite element approximation of the following Cahn–Hilliard equation:

$$\begin{array}{rcll} \gamma \dfrac{\partial u}{\partial t} & = & \Delta w & \text{in} \quad \Omega_T := \Omega \times [0,T], \\ w & = & -\gamma \Delta u + \dfrac{1}{\gamma} \Psi'(u) & \text{in} \quad \Omega_T, \\ \nabla u \cdot \nu & = & \nabla w \cdot \nu = 0 & \text{on} \quad \partial\Omega \times (0,T], \\ u(\cdot,0) & = & u_0 & \text{in} \quad \Omega, \end{array} \tag{1.1}$$

where Ω is a convex polyhedral domain in $\mathbb{R}^d$, $d = 2, 3$, and $T > 0$ is a fixed positive time. Moreover, Ψ is a given energy potential, and in this paper we will take Ψ to be the so called double obstacle potential

$$\Psi(s) := \begin{cases} \frac{1}{2}(1-s^2) & \text{if} \quad s \in [-1,1]\,, \\ \infty & \text{if} \quad s \notin [-1,1]\,. \end{cases} \tag{1.2}$$

Keywords and phrases. Cahn–Hilliard equation, obstacle free energy, linear finite elements, *a posteriori* estimates, adaptive numerical methods.

* *Supported by the EPSRC grant EP/C548973/1.*

[1] Department of Mathematics and the Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK. L.Banas@hw.ac.uk

[2] Department of Mathematics, Imperial College London, London, SW7 2AZ, UK.

We note that other choices of Ψ are also possible, see *e.g.* (1.4) below. In addition, the parameter $\gamma > 0$ is an interaction length, which is small compared to the dimensions of Ω.

Equation (1.1) was originally introduced by Cahn and Hilliard to model spinodal decomposition and coarsening phenomena in binary alloys, see [11,12]. Here u is defined to be the difference of the local concentrations of the two components of an alloy and hence u is restricted to lie in the interval $[-1, 1]$. More recently, the Cahn–Hilliard equation has been used *e.g.* as a phase field model for sharp interface evolutions and to study phase transitions and interface dynamics in multiphase fluids, see *e.g.* [7,22,23] and the references therein. We note that in (1.1) we have used a time scaling, so that in the limit $\gamma \to 0$, we recover the well known sharp interface motions by Mullins–Sekerka. We recall that this limit was first formally shown in [26], and later proved rigorously in [1].

We note that as properties of commercially produced materials depend on microstructures which are generated using special processing techniques, such as phase separation and coarsening mechanisms, accurate predictions of microstructure or the evolution of pattern formation during phase separation and coarsening are of considerable interest in materials science. As it is difficult to obtain such information by real-life experiments, reliable numerical computations are very important. It is the aim of this paper to prove suitable *a posteriori* estimates for the discrete approximation of the considered problem that can be used to construct robust and reliable mesh refinement algorithms in two and three space dimensions, which allow for efficient and reliable numerical simulations.

The theory of Cahn and Hilliard is based on the following Ginzburg–Landau free energy

$$\mathcal{E}(u) := \int_\Omega \left(\tfrac{\gamma}{2}|\nabla u|^2 + \gamma^{-1}\,\Psi(u)\right) \mathrm{d}x. \tag{1.3}$$

The first term in the free energy penalizes large gradients and the second term is the homogeneous free energy. Then (1.1) can be derived from mass balance considerations as a gradient flow for the free energy $\mathcal{E}(u)$, with the chemical potential $w := \frac{\delta \mathcal{E}}{\delta u}$ being the variational derivative of the energy $\mathcal{E}$ with respect to u.

For notational convenience in (1.1) it was implicitly assumed that the free energy Ψ is differentiable. An example for such a potential function is

$$\Psi(s) = \tfrac{1}{4}\,(s^2 - 1)^2, \tag{1.4}$$

which has the advantage of being smooth but the disadvantage that physically non-admissible values with $|u| > 1$ can be attained during the evolution. Of course, the obstacle free energy (1.2) forces u to stay within the interval $[-1, 1]$ of physically meaningful values. This is a clear advantage over a formulation involving (1.4). Hence, in this paper we will from now on consider the obstacle free energy (1.2). Then the chemical potential w needs to be computed with the help of a variational inequality, see (2.1) below. It is this variational inequality which requires special attention in developing an *a posteriori* error estimate.

Typical evolutions of (1.1) starting from a well mixed initial state begin with a relatively short early phase, called spinodal decomposition, in which the local concentrations u grow towards the minimizers ± 1 of (1.2). This leads to a setup, where large parts of the domain are occupied by regions where $u = \pm 1$, which are separated by interfacial regions where $|u| < 1$, in which u smoothly varies from -1 to 1. Then follows a much slower evolution phase, in which the total volume of these interfacial regions is decreased. This phase is called coarsening. The thickness of the interfacial regions, *i.e.*, the region where $|u| < 1$, is asymptotically of order $\mathcal{O}(\gamma)$. As mentioned earlier, it can be shown that in the sharp interface limit (*i.e.*, when $\gamma \to 0$) the long time dynamics of equations (1.1) correspond to the Mullins–Sekerka equation.

Finite element methods for equation (1.1) with (1.2) have been proposed and analyzed in [9], see also [5,6]. In addition, existence and uniqueness of the solution u, w to (1.1), as well as regularity results, were shown in [8]. In [7] a finite element approximation for a related, so called degenerate, Cahn–Hilliard equation was considered, and in addition a heuristic adaptive mesh refinement algorithm was used for numerical simulations in two space dimensions, in order to increase the efficiency of the computations. This approximation and the corresponding mesh refinement have recently been extended to three space dimensions in [3], see also [4]. There exist numerous works on finite element approximations of (1.1) with smooth potentials such as (1.4).

Here we refer to *e.g.* [16–18] and the references therein. *A posteriori* estimates for the Cahn–Hilliard equation with the smooth potential (1.4) have very recently been obtained in [19], where the estimates for a continuous in time semi-discrete approximation only depend on polynomial powers of γ^{-1}, a result which crucially depends on the spectral estimate from [13]. To our knowledge, so far there is no work on *a posteriori* estimates for the Cahn–Hilliard equation with the obstacle potential (1.4), apart from the numerical results in [2], which are based on results related to the work in this paper.

It is the aim of this paper to prove *a posteriori* estimates and examine adaptive finite element methods for (1.1) in two and three space dimensions. Since there is no spectral estimate corresponding to that from [13] available for the non-smooth model, we only examine the error due to the spatial discretization. Therefore we restrict our analysis to spatial *a posteriori* error estimates for a discrete in time analogue of (1.1). In particular, we will derive estimates for a coupled system that consists of an elliptic variational inequality involving two constant obstacles, and a linear elliptic problem; see (2.1) below.

By using the ideas of [27], where error estimates for linear finite element approximations of elliptic obstacle problems are introduced, we are able to obtain an estimate with localized interior residual, which enables effective and reliable error control by refinement that is mainly concentrated in the interfacial region, where $|u| < 1$. The *a posteriori* analysis of elliptic obstacle problems is a relatively new field. A residual *a posteriori* estimate with non-localized interior residual was obtained in [14]. A sharper estimate with localized interior residual was constructed in [28] for constant obstacles and in [27] for general obstacles. A short review on *a posteriori* estimates for elliptic obstacle problems is given in [10]. *A posteriori* estimates for parabolic variational inequalities were derived in [24] by extending the ideas of [27]. We also refer to work in optimal control theory, where very recently an error estimator for a control problem with side constraints involving PDEs and inequality constraints has been introduced in [20,21]. However, we stress that a crucial difference between work on optimal control theory and work on obstacle problems involving variational inequalities is that the former only applies the inequality constraints on the right hand side of the control PDE, and that the localization of the interior residual is not essential to obtain a lower bound for the error, *e.g.* see [21].

The paper is organized as follows. In Section 2, we introduce the continuous in space and discrete in time Cahn–Hilliard equation and its finite element approximation by conforming piece-wise linear elements. In Section 3, we establish an *a posteriori* estimate with non-localized residual, which can potentially lead to extensive mesh refinement outside of the interfacial region, *i.e.* in the region where the solution u is constant. In Section 4, we construct upper and lower bounds for the error with localized interior residual. In Section 5, we discuss a number of adaptive algorithms for numerical computations. Finally, Section 6 is devoted to numerical experiments, where we examine the performance of the adaptive algorithms in two and three space dimensions.

2. Finite element approximation

We consider the following continuous in space semi-discrete counterpart of the Cahn–Hilliard equation obtained by a backward-Euler time discretization of (1.1): Find $u \in \mathcal{K} := \{v \in H^1(\Omega) : |v| \leq 1\}$ and $w \in H^1(\Omega)$ such that

$$\begin{aligned} (u,\phi) + \frac{\tau}{\gamma}(\nabla w, \nabla \phi) &= (f,\phi) \qquad \forall \phi \in H^1(\Omega), \\ \gamma\,(\nabla u, \nabla(\psi - u)) - (w, \psi - u) &\geq (g, \psi - u) \qquad \forall \psi \in \mathcal{K}, \end{aligned} \tag{2.1}$$

where $(\phi,\psi) = \int_\Omega \phi\,\psi$ is the L^2-inner product over Ω. Throughout this paper, we denote the L^2-norm over $D \subset \Omega$ by $\|\cdot\|_D$, and similarly use $\|\cdot\|_{1,D}$ for the H^1-norm. For notational convenience, we drop the subscript in the case $D = \Omega$. In addition, we denote the norm in the dual space $(H^1(\Omega))'$ by $\|\cdot\|_{-1}$ and use $\langle\cdot,\cdot\rangle$ for the duality pairing between $H^1(\Omega)$ and its dual.

On introducing the linear finite element space

$$V^h := \{\phi \in C(\overline{\Omega}) : \phi\,|_T \ \text{ is linear } \forall T \in \mathcal{T}^h\} \subset H^1(\Omega)\,,$$

where $\overline{\Omega} = \cup_{T\in\mathcal{T}^h} T$, we consider the following finite element approximation of (2.1): Find $u_h \in \mathcal{K}^h$ and $w_h \in V^h$ such that

$$\begin{aligned} (u_h,\phi) + \frac{\tau}{\gamma}(\nabla w_h, \nabla\phi) &= (f,\phi) \qquad \forall \phi \in V^h, \\ \gamma\,(\nabla u_h, \nabla(\psi - u_h)) - (w_h, \psi - u_h) &\geq (g, \psi - u_h) \qquad \forall \psi \in \mathcal{K}^h, \end{aligned} \tag{2.2}$$

where $\mathcal{K}^h := \mathcal{K} \cap V^h$. Note that in view of the derivation of (2.1), we usually have $f = u_h^{\text{old}}$, $g = \frac{1}{\gamma} u_h^{\text{old}}$ in (2.2), where u_h^{old} is the solution from the previous time step. Then (2.2) corresponds to one time step of the unconditionally stable, fully discrete approximation in [9]. Also, in that case $f, g \in V_{\text{old}}^h$ are piecewise linear functions, where V_{old}^h is the finite element space corresponding to the previous time step. This case will simplify some steps in the analysis below, in particular when $V_{\text{old}}^h \subset V^h$.

We denote by

$$\begin{aligned} e_u &= u - u_h, \\ e_w &= w - w_h. \end{aligned} \tag{2.3}$$

We recall the following well-known result concerning V^h:

$$|(\phi,\chi) - (\phi,\chi)_h| \leq C\|h^2\,\nabla\phi\|\,\|\nabla\chi\| \quad \forall \phi,\, \chi \in V^h\,; \tag{2.4}$$

where $(\phi,\chi)_h = \int_\Omega \mathcal{I}^h(\phi\,\chi)$ for $\phi,\, \chi \in C(\overline{\Omega})$ is the usual mass lumped inner product, and $\mathcal{I}^h$ is the usual Lagrange interpolation operator onto V^h.

In addition to the triangulation $\mathcal{T}^h$, we introduce the set of its nodes $\mathcal{P}^h$ and edges $\mathcal{E}^h$. We denote the nodal basis functions of V^h as $(\chi_h^p)_{p\in\mathcal{P}_h}$, where $\chi_h^p(q) = 1$ if $p = q$ and $\chi_h^p(q) = 0$ otherwise. Moreover, for each $T \in \mathcal{T}^h$ and $e \in \mathcal{E}^h$ we denote their diameter by h_T and h_e, respectively. We also introduce the local mesh size function $h : \Omega \to \mathbb{R}$, which is piecewise constant and such that $h_{|T} = h_T$ for all $T \in \mathcal{T}^h$. For any set $D \subset \Omega$, we define the discrete neighbourhood of D by $\widetilde{D} = \cup\{T \in \mathcal{T}^h;\, T \cap D \neq \emptyset\}$. In addition, in a slight abuse of notation, we also introduce the short hand notation $\|h^\alpha[\nabla u_h]\|^2 := \sum_{e\in\mathcal{E}_h} \|h_e^\alpha[\nabla u_h]_e\|_e^2$, where $\alpha \in \mathbb{R}$.

3. *A posteriori* estimate with positivity preserving interpolation

In this section we extend the ideas of [14], in order to show how it is possible to derive an upper bound for the error of the finite element approximation in a relatively simple manner. The obtained estimate, however, does not take into account certain special properties of the solution, and may lead to excessive mesh refinement in practice, in areas where the solution u is constant.

We recall the definition of the positivity preserving interpolation operator $\Pi_0^h : L^1(\Omega) \to V^h \cap H_0^1(\Omega)$ from [14], *i.e.*, we have that $u \geq 0 \Rightarrow \Pi_0^h u \geq 0$ for all $u \in L^1(\Omega)$. It is then a straightforward matter to extend this definition to the Neumann boundary condition and double obstacle present here, to obtain an analogous operator $\Pi^h : L^1(\Omega) \to V^h$ such that

$$u \in \mathcal{K} \quad \Rightarrow \quad \Pi^h u \in \mathcal{K}^h. \tag{3.1}$$

In fact, we can choose Π^h to be the operator given in [25], Example 1.1.

We have the following approximation properties of Π^h for $u \in H^1(\Omega)$ and $u_h \in V^h$:

$$\|\nabla\Pi^h u\| \leq \|\nabla u\|, \tag{3.2a}$$

$$\|u - \Pi^h u\|_T \leq C h_T \|\nabla u\|_{\tilde{T}}, \tag{3.2b}$$

$$\|u - \Pi^h u\|_e \leq C h_e^{1/2} \|\nabla u\|_{\tilde{e}}, \tag{3.2c}$$

$$\|u_h - \Pi^h u_h\|_T \leq C\|h^{3/2}[\nabla u_h]_e\|_{\tilde{T}}, \tag{3.2d}$$

$$\|u_h - \Pi^h u_h\|_e \leq C\|h_e[\nabla u_h]_e\|_{\tilde{e}}, \tag{3.2e}$$

cf. [14], where we recall that $\tilde{T}$ is the union of all the elements surrounding T, and similarly for $\tilde{e}$.

Choosing $\phi = e_w$ in (2.1) and $\phi = \Pi^h e_w$ in (2.2), we obtain that

$$(e_u, e_w) + \frac{\tau}{\gamma}\|\nabla e_w\|^2 = (f, e_w - \Pi^h e_w) - (u_h, e_w - \Pi^h e_w) - \frac{\tau}{\gamma}\left(\nabla w_h, \nabla(e_w - \Pi^h e_w)\right). \tag{3.3}$$

Next, we take $\psi = u_h$ in (2.1), leading to

$$\gamma\left(\nabla u, \nabla e_u\right) - (w, e_u) - (g, e_u) \leq 0, \tag{3.4}$$

and $\psi = \Pi^h u \in \mathcal{K}^h$, recall (3.1), in (2.2), which gives

$$\gamma\left(\nabla u_h, \nabla e_u\right) - (w_h, e_u) - (g, e_u) \geq \left(g, \Pi^h u - u\right) - \gamma\left(\nabla u_h, \nabla(\Pi^h u - u)\right) + \left(w_h, \Pi^h u - u\right). \tag{3.5}$$

We have the simple identity $\Pi^h u - u = (\Pi^h e_u - e_u) + (\Pi^h u_h - u_h)$. Therefore, after we subtract (3.5) from (3.4), we obtain

$$\begin{aligned} \gamma\|\nabla e_u\|^2 - (e_w, e_u) \quad \leq \quad & \left(g, e_u - \Pi^h e_u\right) - \gamma\left(\nabla u_h, \nabla(e_u - \Pi^h e_u)\right) + \left(w_h, e_u - \Pi^h e_u\right) \\ & + \left(g + w_h, \Pi^h u_h - u_h\right) - \gamma\left(\nabla u_h, \nabla(\Pi^h u_h - u_h)\right). \end{aligned} \tag{3.6}$$

Further, after employing integration by parts, since $\Delta u_h|_T = 0$, we observe that

$$\begin{aligned} (\nabla u_h, \nabla\psi) = \sum_{T\in\mathcal{T}_h}\int_T \nabla u_h\cdot\nabla\psi \quad &= \quad \sum_{T\in\mathcal{T}_h}\left\{\int_{\partial T}\nabla u_h\cdot\boldsymbol{\nu}_T\psi - \int_T \Delta u_h\psi\right\} \\ &= \quad \sum_{e\in\mathcal{E}_h}\int_e [\nabla u_h]_e\psi \qquad \forall\psi\in H^1(\Omega), \end{aligned} \tag{3.7}$$

where $\boldsymbol{\nu}_T$ denotes the outward unit vector to $T \in \mathcal{T}_h$.

Hence it follows from (3.3), on applying the Cauchy–Schwartz and Young inequalities together with (3.2b) and (3.7), that

$$(e_u, e_w) + \frac{\tau}{\gamma}\|\nabla e_w\|^2 \leq C\frac{\gamma}{\tau}\|h(u_h - f)\|^2 + \frac{\tau}{4\gamma}\|\nabla e_w\|^2 + C\frac{\tau}{\gamma}\|h^{1/2}[\nabla w_h]_e\|^2 + \frac{\tau}{4\gamma}\|\nabla e_w\|^2. \tag{3.8}$$

Similarly, it follows from (3.6) that

$$\begin{aligned} \gamma\|\nabla e_u\|^2 - (e_w, e_u) \quad \leq \quad & \frac{C}{\gamma}\|h(g + w_h)\|^2 + \frac{\gamma}{4}\|\nabla e_u\|^2 + \gamma C\|h^{1/2}[\nabla u_h]\|^2 + \frac{\gamma}{4}\|\nabla e_u\|^2 \\ & + \left(g + w_h, \Pi^h u_h - u_h\right) - \gamma\left(\nabla u_h, \nabla(\Pi^h u_h - u_h)\right). \end{aligned} \tag{3.9}$$

The last two terms on the right-had side of (3.9) can be estimated, on noting (3.7) and (3.2d)-(3.2e), as

$$\left(g + w_h, \Pi^h u_h - u_h\right) - \gamma\left(\nabla u_h, \nabla(\Pi^h u_h - u_h)\right) \leq C\left(\gamma\|h^{1/2}[\nabla u_h]\|^2 + \frac{1}{\gamma}\|h(g + w_h)\|^2\right).$$

By combining the previous equation with (3.8), (3.9) we arrive at

$$\frac{\tau}{\gamma}\|\nabla e_w\|^2 + \gamma\|\nabla e_u\|^2 \leq C\left(\frac{\gamma}{\tau}\|h(f - u_h)\|^2 + \frac{\tau}{\gamma}\|h^{1/2}[\nabla w_h]\|^2 + \frac{1}{\gamma}\|h(g + w_h)\|^2 + \gamma\|h^{1/2}[\nabla u_h]\|^2\right). \tag{3.10}$$

Upon subsequently rescaling we obtain

$$\tau\,\|\nabla e_w\|^2 + \gamma^2\,\|\nabla e_u\|^2 \leq C\left[\frac{\gamma^2}{\tau}\,\|h(u_h - f)\|^2 + \tau\,\|h^{1/2}[\nabla w_h]\|^2 + \|h(g + w_h)\|^2 + \gamma^2\|h^{1/2}[\nabla u_h]\|^2\right]. \tag{3.11}$$

Remark 3.1. The disadvantage of the above estimate is that the interior residual $\|h(g+w_h)\|$ corresponding to the variational inequality in (2.2) is not localized to the noncontact set (see definition in the next section), which can cause excessive mesh refinement in the contact set, where the solution u_h is constant and where w_h usually attains large values. However, as the variational inequality in (2.2) trivially holds in the contact set, ideally there should be no contribution from the interior residual to the *a posteriori* error estimate. This problem will be addressed in the next section.

4. *A posteriori* estimate with localized interior residual

In this section we derive an *a posteriori* estimate with an interior residual localized to the interface, *i.e.* the interior residual induced by the variational inequality in (2.2) is zero in the region where $|u_h| = 1$. This result gives rise to more efficient *a posteriori* error based mesh refinement strategies, and it is furthermore a theoretical justification for the construction of heuristical mesh adaptive algorithms, where the mesh refinement is concentrated in the interfacial area, *i.e.* where $|u_h| < 1$. We extend the ideas of [27,28] to the semi-discrete formulation (2.1) of the Cahn–Hilliard equation. Here we define the discrete functions $f_h := \mathcal{I}^h f$, $g_h := \mathcal{I}^h g$ and note that by definition we have $(g_h, \phi)_h = (g, \phi)_h$, $(f_h, \phi)_h = (f, \phi)_h$ for all $\phi \in C(\overline{\Omega})$.

Instead of the discrete formulation (2.2), we consider the following discrete problem: Find $u_h \in \mathcal{K}^h$ and $w_h \in V^h$ such that

$$\begin{aligned} (u_h, \phi)_h + \frac{\tau}{\gamma}(\nabla w_h, \nabla \phi) &= (f_h, \phi)_h \qquad \forall \phi \in V^h, \\ \gamma\,(\nabla u_h, \nabla(\psi - u_h)) - (w_h, \psi - u_h)_h &\geq (g_h, \psi - u)_h \qquad \forall \psi \in \mathcal{K}^h. \end{aligned} \tag{4.1}$$

The above formulation only differs from (2.2) in the zero order terms, where we use the reduced discrete inner product $(\cdot,\cdot)_h$.

Given the true solution u, and following the technique in [27] for a single obstacle, we obtain the partition of the domain

$$\Omega = \mathcal{C}(u) \cup \mathcal{N}(u) \cup \mathcal{F}(u), \tag{4.2}$$

where

- the contact set $\mathcal{C}(u)$ is the maximal open set $A \subset \Omega$ such that $|u| \equiv 1$ on A;
- the noncontact set $\mathcal{N}(u) := \cup_{\epsilon>0} B_\epsilon$; where B_ϵ is the maximal open set $B \subset \Omega$ such that $|u| < 1 - \epsilon$;
- the free boundary $\mathcal{F}(u)$ is the set $\Omega \setminus (\mathcal{C}(u) \cup \mathcal{N}(u))$.

The contact set can be further decomposed as $\mathcal{C}(u) = \mathcal{C}^+(u) \cup \mathcal{C}^-(u)$, where $u = \pm 1$ on $\mathcal{C}^\pm(u)$.

We define the continuous residual $\sigma(u) \in (H^1(\Omega))'$ as

$$\langle \sigma(u), \psi \rangle = (g, \psi) + (w, \psi) - \gamma(\nabla u, \nabla \psi) \qquad \forall \psi \in H^1(\Omega). \tag{4.3}$$

The following properties can be obtained from the definition of σ (note, $|u| \equiv 1$ in $\mathcal{C}(u)$)

$$\sigma \geq 0 \qquad \text{in } \mathcal{C}^+(u), \tag{4.4a}$$
$$\sigma \leq 0 \qquad \text{in } \mathcal{C}^-(u), \tag{4.4b}$$
$$\sigma = g + w \qquad \text{in } \mathcal{C}(u), \tag{4.4c}$$
$$\sigma = 0 \qquad \text{in } \mathcal{N}(u). \tag{4.4d}$$

The discrete residual is defined as $\sigma_h \in V^h$ such that

$$(\sigma_h, \psi)_h = (g_h, \psi)_h + (w_h, \psi)_h - \gamma(\nabla u_h, \nabla \psi) \qquad \forall \psi \in V^h\,. \tag{4.5}$$

Alternatively we can write

$$(\sigma_h, \psi)_h = (g_h, \psi)_h + (w_h, \psi)_h + \gamma(\Delta_h u_h, \psi)_h, \tag{4.6}$$

where $\Delta_h : V^h \to V^h$ is the usual discrete Laplacian on V^h.

We define the jump across an inner element edge/face $e = T_1 \cap T_2 \in \mathcal{E}_h$ as

$$[\nabla u_h]_e = \frac{1}{2}(\nabla u_h|_{T_1} - \nabla u_h|_{T_2}) \cdot \boldsymbol{\nu}_e,$$

where $\boldsymbol{\nu}_e$ is a unit normal vector of e pointing from T_1 to T_2. For a Neumann boundary edge $e \in \mathcal{E}_h \cap \partial\Omega \subset T$ we define

$$[\nabla u_h]_e = \nabla u_h|_T \cdot \boldsymbol{\nu},$$

where $\boldsymbol{\nu}$ is he outward unit vector to the boundary $\partial\Omega$.

Similarly to (4.2), the domain Ω can be decomposed into

$$\Omega = \mathcal{C}_h(u_h) \cup \mathcal{F}_h(u_h) \cup \mathcal{N}_h(u_h), \tag{4.7}$$

where

$$\begin{aligned}
\mathcal{C}_h(u_h) &:= \mathcal{C}_h(u_h)^+ \cup \mathcal{C}_h(u_h)^-, \quad \mathcal{C}_h(u_h)^\pm := \bigcup\{T \in \mathcal{T}^h;\ |u_h| = \pm 1 \text{ on } T\},\\
\mathcal{N}_h(u_h) &:= \bigcup\{T \in \mathcal{T}^h;\ |u_h| < 1 \text{ on } T\},\\
\mathcal{F}_h(u_h) &:= \Omega \setminus [\mathcal{C}_h(u_h) \cup \mathcal{N}_h(u_h)].
\end{aligned}$$

In our context, $\mathcal{C}_h$ denotes the subdomains with pure materials, $\mathcal{N}_h$ denotes the diffuse interface and $\mathcal{F}_h$ is the so-called discrete free boundary between $\mathcal{C}_h$ and $\mathcal{N}_h$.

Similarly as in (4.4), one can establish for all nodes $p \in \mathcal{P}_h$ that

$$\sigma_h(p) \geq 0 \qquad \text{if } p \in \mathcal{C}_h^+, \tag{4.8a}$$

$$\sigma_h(p) \leq 0 \qquad \text{if } p \in \mathcal{C}_h^-, \tag{4.8b}$$

$$\sigma_h(p) = g_h(p) + w_h(p) \qquad \text{if } |u_h| = 1 \text{ on } \operatorname{supp} \chi_h^p, \tag{4.8c}$$

$$\sigma_h(p) = 0 \qquad \text{if } |u_h(p)| < 1. \tag{4.8d}$$

Note that $\Delta_h u_h = 0$ in $\mathcal{C}_h(u_h)$.

Following [27], we define the Galerkin functional $\mathcal{G}_h \in (H^1(\Omega))'$ as

$$\langle \mathcal{G}_h, \psi \rangle = \gamma(\nabla(u_h - u), \nabla\psi) - (w_h, \psi) + (w, \psi) + (\sigma_h - \sigma, \psi) \qquad \forall \psi \in H^1(\Omega)\,. \tag{4.9}$$

We directly have from (4.3) that

$$\langle \mathcal{G}_h, \psi \rangle = \gamma(\nabla u_h, \nabla\psi) - (w_h + g, \psi) + (\sigma_h, \psi) \qquad \forall \psi \in H^1(\Omega)\,. \tag{4.10}$$

Lemma 4.1 (perturbed Galerkin orthogonality)**.** *There exists a constant C depending only on the mesh regularity, such that*

$$\|\mathcal{G}_h\|_{-1,h} := \sup_{\psi_h \in V^h, \|\nabla\psi_h\| = 1} \langle \mathcal{G}_h, \psi_h \rangle = C\left(\gamma \|h^2 \nabla \Delta_h u_h\| + \|g_h - g\|_{-1,h}\right).$$

Proof. On recalling the definitions of $(\cdot,\cdot)_h$, σ_h and $\mathcal{G}_h$, we have for any $\psi_h \in V^h$ that

$$\begin{aligned}
\langle \mathcal{G}_h, \psi_h \rangle &= (g, \psi_h)_h + [\gamma(\nabla u_h, \nabla\psi_h) - (g + w_h, \psi_h)_h] + (w_h, \psi_h)_h - (w_h + g, \psi_h) + (\sigma_h, \psi_h)\\
&= (g, \psi_h)_h - (g, \psi_h) + (w_h, \psi_h)_h - (w_h, \psi_h) + (\sigma_h, \psi_h) - (\sigma_h, \psi_h)_h\\
&= (g + w_h - \sigma_h, \psi_h)_h - (g_h + w_h - \sigma_h, \psi_h) + (g_h - g, \psi_h)\\
&= -\gamma(\Delta_h u_h, \psi_h)_h + \gamma(\Delta_h u_h, \psi_h) + (g_h - g, \psi_h).
\end{aligned} \tag{4.11}$$

Furthermore, it follows from (2.4) that

$$-\gamma(\Delta_h u_h, \psi_h)_h + \gamma(\Delta_h u_h, \psi_h) \leq \gamma \|h^2 \nabla \Delta_h u_h\| \|\nabla \psi_h\|,$$

which yields the desired result. □

Also the following is just a generalisation of [27], Lemma 3.4, except that the term $\|\sigma_h - \sigma\|_{-1}^2$ does not appear on the left hand side of (4.12).

Lemma 4.2. *The following inequality holds*

$$\gamma \|\nabla(u_h - u)\|^2 - 2\,(w_h - w, u_h - u) \leq \frac{C_1}{\gamma} \|\mathcal{G}_h\|_{-1}^2 - C_2(\sigma_h - \sigma, u_h - u). \tag{4.12}$$

Proof. It follows from (4.9) that

$$\begin{aligned} \gamma \|\nabla(u_h - u)\|^2 - (w_h - w, u_h - u) &= \langle \mathcal{G}_h, u_h - u \rangle - (\sigma_h - \sigma, u_h - u) \\ &\leq \|\mathcal{G}_h\|_{-1} \|\nabla(u_h - u)\| - (\sigma_h - \sigma, u_h - u). \end{aligned}$$

Hence by Young's inequality we get

$$\gamma \|\nabla(u_h - u)\|^2 - (w_h - w, u_h - u) \leq \frac{1}{2\gamma} \|\mathcal{G}_h\|_{-1}^2 + \frac{\gamma}{2} \|\nabla(u_h - u)\|^2 - (\sigma_h - \sigma, u_h - u). \tag{4.13}$$

The assertion of the lemma then easily follows from the last inequality. □

4.1. **Global upper bound**

In the following lemma we estimate the Galerkin functional.

Lemma 4.3. *There exists a constant C depending only on the mesh regularity, such that*

$$\|\mathcal{G}_h\|_{-1} \leq C \left[\gamma \left(\sum_{e \in \mathcal{E}_h} \left\| h_e^{1/2} [\nabla u_h]_e \right\|_e^2 \right)^{1/2} + \|h(g + w_h - \sigma_h)\| + \gamma \left\| h^2 \nabla \Delta_h u_h \right\| + \|g - g_h\|_{-1,h} \right].$$

Proof. For $\varphi \in H^1(\Omega)$ we write

$$\langle \mathcal{G}_h, \varphi \rangle = \langle \mathcal{G}_h, \varphi - I^h \varphi \rangle + \langle \mathcal{G}_h, I^h \varphi \rangle,$$

where $I^h \varphi$ denotes the Clément interpolant for φ, see [15]. The second term in the above equation can be estimated using Lemma 4.1 (the perturbed Galerkin orthogonality) and the properties of I^h as

$$\langle \mathcal{G}_h, I^h \varphi \rangle \leq C \|\mathcal{G}_h\|_{-1,h} \|\nabla I^h \varphi\| \leq C(\gamma \|h^2 \nabla \Delta_h u_h\| + \|g - g_h\|_{-1,h}) \|\nabla \varphi\|.$$

Similarly, on recalling (4.10), we can estimate the first term using standard arguments of *a posteriori* estimation as

$$\begin{aligned} \langle \mathcal{G}_h, \varphi - I^h \varphi \rangle &= \gamma(\nabla u_h, \nabla(\varphi - I^h \varphi)) - (w_h + g - \sigma_h, \varphi - I^h \varphi) \\ &= -\gamma \sum_{e \in \mathcal{E}_h} \int_e [\nabla u_h]_e (\varphi - I^h \varphi) - \sum_{T \in \mathcal{T}_h} \int_T (w_h + g - \sigma_h)\,(\varphi - I^h \varphi) \\ &\leq C \left[\gamma \left(\sum_{e \in \mathcal{E}_h} \left\| h_e^{1/2} [\nabla u_h]_e \right\|_{L_2(e)}^2 \right)^{1/2} + \|h(g + w_h - \sigma_h)\| \right] \|\nabla \varphi\|, \end{aligned}$$

which concludes the proof. □

The following lemma is an adaption of [27], Proposition 3.7.

Lemma 4.4. *The following inequality holds for the solutions u and u_h of (2.1) and (4.1), respectively.*

$$(\sigma_h - \sigma, u_h - u) \geq -C \left[\gamma \sum_{T \in \mathcal{T}_h'} h_T^4 \|\nabla \Delta_h u_h\|_T^2 + \frac{1}{\gamma} \sum_{T \in \mathcal{T}_h'} h_T^2 \|w_h + g_h\|_T^2 + \gamma \sum_{e \in \mathcal{E}_h'} h_e \|[\nabla u_h]_e\|_e^2 \right],$$

where

$$\begin{aligned} \mathcal{T}_h' &= \{T \in \mathcal{T}_h;\, \exists\, p_1, p_2 \in \mathcal{P}_h \cap T,\, |u_h(p_1)| = 1 \text{ and } |u_h(p_2)| < 1\}, \\ \mathcal{E}_h' &= \{e \in \mathcal{E}_h;\, e \cap \mathcal{P}_h' \neq \emptyset\} \text{ with } \mathcal{P}_h' = \{p \in \mathcal{P}_h;\, |u_h(p)| = 1 \text{ and } |u_h| \not\equiv 1 \text{ on } \operatorname{supp} \chi_h^p\}. \end{aligned}$$

Proof. We rewrite

$$(\sigma_h - \sigma, u_h - u) = (\sigma_h, u_h - u) + (\sigma, u - u_h).$$

Since $u_h \in \mathcal{K}$, we can estimate the second term using (2.1)

$$(\sigma, u - u_h) = \gamma(\nabla u, \nabla(u_h - u)) - (g, u_h - u) - (w, u_h - u) \geq 0.$$

Next, on noting that $\Omega = \mathcal{C}_h \cup \mathcal{N}_h \cup \mathcal{T}_h'$, we rewrite the first term on the right-hand side as

$$(\sigma_h, u_h - u) = \int_{\mathcal{C}_h^-} \sigma_h(-1 - u) + \int_{\mathcal{C}_h^+} \sigma_h(1 - u) + \int_{\mathcal{N}_h} \sigma_h(u_h - u) + \int_{\mathcal{T}_h'} \sigma_h(u_h - u).$$

Using (4.8a)-(4.8b) we get

$$\int_{\mathcal{C}_h^-} \sigma_h(-1 - u) \geq 0, \quad \int_{\mathcal{C}_h^+} \sigma_h(1 - u) \geq 0.$$

Recalling (4.8d) we have

$$\int_{\mathcal{N}_h} \sigma_h(u_h - u) = 0.$$

The remaining term is estimated as follows. Consider $T \in \mathcal{T}_h'$ and $p_1, p_2 \in \mathcal{P}_h \cap T$, with $u_h(p_1) = \pm 1$, $|u_h(p_2)| < 1$ and $\sigma_h|_T \geq 0$ we get

$$\begin{aligned} \int_T \sigma_h(u_h - u) &= \int_T \sigma_h(u_h - 1) + \int_T \sigma_h(1 - u) \\ &\geq \int_T \sigma_h(u_h - 1) \geq -\|\sigma_h\|_T \|u_h - 1\|_T, \end{aligned}$$

if $\sigma_h|_T \leq 0$ we get

$$\begin{aligned} \int_T \sigma_h(u_h - u) &= \int_T \sigma_h(u_h + 1) + \int_T \sigma_h(-1 - u) \\ &\geq \int_T \sigma_h(u_h + 1) \geq -\|\sigma_h\|_T \|u_h + 1\|_T. \end{aligned}$$

From [27], Lemma 3.6, we obtain ($\mathcal{E}_h(p) := \{e \in \mathcal{E}_h;\, p \in e\}$)

$$\|u_h - 1\|_T \leq C h_T \left(\sum_{e \in \mathcal{E}_h(p_1)} h_e \|[u_h - 1]_e\|_e^2 \right)^{1/2} \leq C h_T \left(\sum_{e \in \mathcal{E}_h(p_1)} h_e \|[u_h]_e\|_e^2 \right)^{1/2}.$$

We have from

$$\sigma_h = \sigma_h - w_h - g_h + w_h + g_h \leq |\sigma_h - w_h - g_h| + |w_h + g_h|,$$

that

$$\|\sigma_h\|_T \le \|\sigma_h - w_h - g_h\|_T + \|w_h + g_h\|_T \le \gamma\, h_T \|\nabla \Delta_h u_h\|_T + \|w_h + g_h\|_T.$$

Finally we get

$$\int_T \sigma_h(u_h - u) \ge -C \left(\gamma\, h_T^4 \|\nabla \Delta_h u_h\|_T^2 + \frac{1}{\gamma}\, h_T^2 \|w_h + g_h\|_T^2 + \gamma \sum_{e \in \mathcal{E}_h(p_1)} h_e \|[u_h]_e\|_e^2 \right),$$

which, on noting that $\mathcal{E}_h' = \cup_{p \in \mathcal{P}_h'} \mathcal{E}_h(p)$, concludes the proof. □

The following lemma is a simple consequence of (4.12) and Lemmas 4.3 and 4.4.

Lemma 4.5.

$$\begin{aligned} \gamma \|\nabla(u_h - u)\|^2 - 2\,(w_h - w, u_h - u) \quad \le \quad & C \frac{1}{\gamma} \Bigg[\gamma^2 \sum_{e \in \mathcal{E}_h} \left\| h_e^{1/2} [\nabla u_h]_e \right\|_e^2 + \|h(g + w_h - \sigma_h)\|^2 \\ & + \gamma^2 \left\| h^2 \nabla \Delta_h u_h \right\|^2 + \|g - g_h\|_{-1,h}^2 + \sum_{T \in \mathcal{T}_h'} \|h_T (w_h + g_h)\|_T^2 \Bigg]. \end{aligned}$$

The next lemma gives an estimate for the first equation in (4.1).

Lemma 4.6.

$$\begin{aligned} \frac{\tau}{\gamma} \|\nabla(w_h - w)\|^2 + 2\,(u_h - u, w_h - w) \quad \le \quad & C \frac{\gamma}{\tau} \Bigg[\frac{\tau^2}{\gamma^2} \sum_{e \in \mathcal{E}_h} \left\| h_e^{1/2} [\nabla w_h]_e \right\|_e^2 + \|h(u_h - f)\|^2 \\ & + \left\| h^2 \nabla (u_h - f_h) \right\|^2 + \|f - f_h\|_{-1,h}^2 \Bigg]. \end{aligned}$$

Proof. We start with the identity

$$\frac{\tau}{\gamma}(\nabla e_w, \nabla \phi) + (e_u, \phi) = \frac{\tau}{\gamma}(\nabla e_w, \nabla(\phi - I^h \phi)) + (e_u, \phi - I^h \phi) + \frac{\tau}{\gamma}(\nabla e_w, \nabla I^h \phi) + (e_u, I^h \phi)$$

for any $\phi \in H^1(\Omega)$. Next, according to (2.1), (4.1), we can rewrite the above equation as

$$\frac{\tau}{\gamma}(\nabla e_w, \nabla \phi) + (e_u, \phi) = -\frac{\tau}{\gamma}(\nabla w_h, \nabla(\phi - I^h \phi)) + (f - u_h, \phi - I^h \phi) + \frac{\tau}{\gamma}(\nabla e_w, \nabla I^h \phi) + (e_u, I^h \phi). \tag{4.14}$$

Similarly as in Lemma 4.1, on noting (2.1), (4.1) and (2.4), we obtain that

$$\frac{\tau}{\gamma}(\nabla e_w, \nabla I^h \phi) + (e_u, I^h \phi) \le \left\| h^2 \nabla (u_h - f_h) \right\| \|\nabla \phi\| + \|f - f_h\|_{-1,h} \|\nabla \phi\|, \tag{4.15}$$

and similarly to Lemma 4.3, we have that

$$-\frac{\tau}{\gamma}(\nabla w_h, \nabla(\phi - I^h \phi)) + (f - u_h, (\phi - I^h \phi)) \le C \left[\frac{\tau}{\gamma} \left(\sum_{e \in \mathcal{E}_h} \left\| h_e^{1/2} [\nabla w_h]_e \right\|_e^2 \right)^{1/2} + \|h(u_h - f)\| \right] \|\nabla \phi\|. \tag{4.16}$$

The proof can be concluded by combining (4.15), (4.16) and (4.14), and by subsequently applying a Young's inequality for $\phi = e_w$. □

The following corollary is a simple consequence of Lemmas 4.5 and 4.6.

Corollary 4.1. *The following estimate is valid for u_h, w_h:*

$$\begin{aligned}\gamma^2 \|\nabla(u_h-u)\|^2 + \tau\,\|\nabla(w_h-w)\|^2 \\ \le C\Bigg[\tau\sum_{e\in\mathcal{E}_h}\left\|h_e^{1/2}[\nabla w_h]_e\right\|_e^2 + \frac{\gamma^2}{\tau}\,\|h(u_h-f)\|^2 + \frac{\gamma^2}{\tau}\left\|h^2\nabla(u_h-f)\right\|^2 \\ + \gamma^2\sum_{e\in\mathcal{E}_h}\left\|h_e^{1/2}[\nabla u_h]_e\right\|_e^2 + \|h(g+w_h-\sigma_h)\|^2 + \sum_{T\in\mathcal{T}_h'}\|h_T(w_h+g_h)\|_T^2 \\ + \gamma^2\left\|h^2\nabla\Delta_h u_h\right\|^2 + \|g-g_h\|_{-1,h}^2 + \frac{\gamma^2}{\tau}\|f-f_h\|_{-1,h}^2\Bigg].\end{aligned} \tag{4.17}$$

Remark 4.1. The estimate (4.17) differs from (3.11) in the following:

- the interior residual, which is now $\|h(g+w_h-\sigma_h)\|^2 + \sum_{T\in\mathcal{T}_h'}\|h_T(w_h+g_h)\|_T^2$, is localized to the "discrete noncontact set" $\Omega\setminus\mathcal{C}_h$, on recalling (4.8c);
- for simplicity, we did not consider coarsening in the derivation of (3.11), *i.e.* the terms $\|g-g_h\|_{-1,h}^2$, $\frac{\gamma^2}{\tau}\|f-f_h\|_{-1,h}^2$ are not included in (3.11);
- the terms $\frac{\gamma^2}{\tau}\left\|h^2\nabla(u_h-f)\right\|^2$, $\gamma^2\left\|h^2\nabla\Delta_h u_h\right\|^2$ in (4.17) are due to the use of the discrete inner product $(\cdot,\cdot)_h$ and are therefore not present in (3.11).

Finally, we note that the quantity

$$\|h^2\nabla\Delta_h u_h\|$$

is 0 within the discrete contact set, *cf.* [27], Remark 3.7, and so it will not contribute to the *a posteriori* error estimate in that region.

4.2. **Local lower bounds**

To each function $f\in L_2(\Omega)$ we assign a piecewise constant function $\overline{f}$ defined as

$$\overline{f}_{|T} = \frac{1}{|T|}\int_T f \qquad \forall T\in\mathcal{T}^h.$$

Further, the so-called local data oscillation is defined as

$$\operatorname{osc}_h(f,T) = \|h_T(\overline{f}-f)\|_T.$$

Lemma 4.7. *The following local estimate holds for all $T\in\mathcal{T}^h$*

$$\begin{aligned}\left[\gamma^2\sum_{e\subset T}\left\|h_e^{1/2}[\nabla u_h]_e\right\|_e^2 + \|h_T(g+w_h-\sigma_h)\|_T^2 + \gamma^2\left\|h_T^2\nabla\Delta_h u_h\right\|_T^2\right]^{1/2} \le \\ C\Big\{\gamma\|\nabla(u_h-u)\|_T + \|h_T(g-g_h)\|_T + \|h_T(\sigma_h-\sigma)\|_T + \|h_T(w_h-w)\|_T + \operatorname{osc}_h(g+w_h-\sigma_h,T)\Big\}.\end{aligned}$$

Proof. The proof is based on the local argument of Verfürth [29].

With every $T\in\mathcal{T}_h$, $e\in\mathcal{E}_h$ we respectively associate the standard canonical bubble functions ψ_T, ψ_e. For technical reasons, we introduce the auxiliary function $z_h := \gamma\,u_h$. Then, following a similar argument in [14],

for any $T \in \mathcal{T}_h$, we can construct a function $\phi_T := \alpha_T \psi_T + \sum_{e \subset T} \beta_e \psi_e$, where α_T, β_e are chosen such that

$$([\nabla z_h]_e, \phi_T)_e = h_e \, \|[\nabla z_h]_e\|_e^2 \qquad \forall e \subset T,$$
$$(\overline{g} + \overline{w}_h - \overline{\sigma}_h, \phi_T)_T = h_T^2 \|\overline{g} + \overline{w}_h - \overline{\sigma}_h\|_T^2,$$

and

$$\|\nabla \phi_T\|_T \leq C \left[\sum_{e \subset T} \left\| h_e^{1/2} [\nabla z_h]_e \right\|_e^2 + \|h_T(\overline{g} + \overline{w}_h - \overline{\sigma}_h)\|_T^2 \right]^{1/2},$$
$$\|\phi_T\|_T \leq C h_T \left[\sum_{e \subset T} \left\| h_e^{1/2} [\nabla z_h]_e \right\|_e^2 + \|h_T(\overline{g} + \overline{w}_h - \overline{\sigma}_h)\|_T^2 \right]^{1/2}.$$

We have, on recalling (3.7) and (4.3), that

$$\begin{aligned}
\gamma^2 \sum_{e \subset T} \left\| h_e^{1/2} [\nabla u_h]_e \right\|_e^2 + \|h_T(\overline{g} + \overline{w}_h - \overline{\sigma}_h)\|_T^2 &= \sum_{e \subset T} \left\| h_e^{1/2} [\nabla z_h]_e \right\|_e^2 + \|h_T(\overline{g} + \overline{w}_h - \overline{\sigma}_h)\|_T^2 \\
&= \sum_{e \subset T} ([\nabla z_h]_e, \phi_T)_e + (\overline{g} + \overline{w}_h - \overline{\sigma}_h, \phi_T)_T \\
&= \gamma \sum_{e \subset T} ([\nabla u_h]_e, \phi_T)_e + (\overline{g} + \overline{w}_h - \overline{\sigma}_h, \phi_T)_T \\
&= -\gamma (\nabla u_h, \nabla \phi_T)_T + (g + w_h - \sigma_h, \phi_T)_T + (\overline{g} + \overline{w}_h - \overline{\sigma}_h - (g + w_h - \sigma_h), \phi_T)_T \\
&= \gamma (\nabla (u - u_h), \nabla \phi_T)_T - (w - w_h, \phi_T)_T + (\sigma - \sigma_h, \phi_T)_T \\
&\quad + (\overline{g} + \overline{w}_h - \overline{\sigma}_h - (g + w_h - \sigma_h), \phi_T)_T \\
&\leq \gamma \|\nabla (u - u_h)\|_T \|\nabla \phi_T\|_T + \|w - w_h\|_T \|\phi_T\|_T + \|\sigma - \sigma_h\|_T \|\phi_T\|_T \\
&\quad + \|\overline{g} + \overline{w}_h - \overline{\sigma}_h - (g + w_h - \sigma_h)\|_T \|\phi_T\|_T \\
&\leq C \Big[\gamma \|\nabla (u - u_h)\|_T + h_T \|w - w_h\|_T + h_T \|\sigma - \sigma_h\|_T + \mathrm{osc}_h(g + w_h - \sigma_h, T) \Big] \\
&\quad \times \left[\gamma^2 \sum_{e \subset T} \left\| h_e^{1/2} [\nabla u_h]_e \right\|_e^2 + \|h_T(\overline{g} + \overline{w}_h - \overline{\sigma}_h)\|_T^2 \right]^{1/2}.
\end{aligned} \tag{4.18}$$

Next, we have from (4.6) and an inverse inequality that

$$\begin{aligned}
\gamma \, \left\| h_T^2 \nabla \Delta_h u_h \right\|_T &= h_T^2 \|\nabla (g_h + w_h - \sigma_h)\|_T \leq C \, h_T \, \|g_h + w_h - \sigma_h\|_T \\
&\leq C \, [h_T \|\overline{g} + \overline{w}_h - \overline{\sigma}_h\|_T + \, h_T \|g - g_h\|_T + \mathrm{osc}_h(g + w_h - \sigma_h, T)].
\end{aligned} \tag{4.19}$$

Finally, the assertion of the lemma follows on combining (4.18) and (4.19) and on noting that

$$\|h_T(g + w_h - \sigma_h)\|_T^2 \leq \|h_T(\overline{g} + \overline{w}_h - \overline{\sigma}_h)\|_T^2 + \mathrm{osc}_h(g + w_h - \sigma_h, T). \qquad \square$$

Lemma 4.8. *The following estimate holds for all* $T \in \mathcal{T}^h$

$$\left[\frac{\tau^2}{\gamma^2}\sum_{e\subset T}\left\|h_e^{1/2}[\nabla w_h]_e\right\|_e^2 + \|h_T(u_h-f)\|_T^2 + \left\|h_T^2\nabla(u_h-f)\right\|_T^2\right]^{1/2} \leq$$
$$C\left(\frac{\tau}{\gamma}\|\nabla(w_h-w)\|_T + \|h_T(u_h-u)\|_T + \|h_T(f-f_h)\| + \mathrm{osc}_h(u_h-f,T)\right). \quad (4.20)$$

Proof. The proof is similar to the proof of the previous lemma. Similarly to before, we can construct a function $\phi_T := \alpha_T\psi_T + \sum_{e\subset T}\beta_e\psi_e$, where α_T, β_e are chosen such that

$$([\nabla w_h]_e, \phi_T)_e = \frac{\tau}{\gamma}\,h_e\|[\nabla w_h]_e\|_e^2 \qquad \forall e\subset T,$$
$$(\overline{f}-\overline{u}_h, \phi_T)_T = h_T^2\|\overline{u}_h-\overline{f}\|_T^2,$$

and

$$\|\nabla\phi_T\|_T \leq C\left[\frac{\tau^2}{\gamma^2}\sum_{e\subset T}\left\|h_e^{1/2}[\nabla w_h]_e\right\|_e^2 + \left\|h_T(\overline{u}_h-\overline{f})\right\|_T^2\right]^{1/2},$$
$$\|\phi_T\|_T \leq Ch_T\left[\frac{\tau^2}{\gamma^2}\sum_{e\subset T}\left\|h_e^{1/2}[\nabla w_h]_e\right\|_e^2 + \left\|h_T(\overline{u}_h-\overline{f})\right\|_T^2\right]^{1/2}.$$

We can write, on recalling (2.1), that

$$\begin{aligned}
\frac{\tau^2}{\gamma^2}\sum_{e\subset T}\left\|h_e^{1/2}[\nabla w_h]_e\right\|_e^2 + \left\|h_T(\overline{u}_h-\overline{f})\right\|_T^2 &= \frac{\tau}{\gamma}\sum_{e\subset T}([\nabla w_h]_e,\phi_T)_e - (\overline{u}_h-\overline{f},\phi_T)_T\\
&= -\frac{\tau}{\gamma}(\nabla w_h,\nabla\phi_T)_T - (u_h-f,\phi_T)_T - (\overline{u}_h-\overline{f}-(u_h-f),\phi_T)_T\\
&= \frac{\tau}{\gamma}(\nabla(w-w_h),\nabla\phi_T)_T + (u-u_h,\phi_T)_T - (\overline{u}_h-\overline{f}-(u_h-f),\phi_T)_T\\
&\leq \frac{\tau}{\gamma}\|\nabla(w-w_h)\|_T\|\nabla\phi_T\|_T + \|u-u_h\|_T\|\phi_T\|_T + \|\overline{u}_h-\overline{f}-(u_h-f)\|_T\|\phi_T\|_T\\
&\leq C\left[\frac{\tau}{\gamma}\|\nabla(w-w_h)\|_T + h_T\|u-u_h\|_T + \mathrm{osc}_h(u_h-f,T)\right]\\
&\quad\times\left[\frac{\tau^2}{\gamma^2}\sum_{e\subset T}\left\|h_e^{1/2}[\nabla w_h]_e\right\|_e^2 + \left\|h_T(\overline{u}_h-\overline{f})\right\|_T^2\right]^{1/2}. \quad (4.21)
\end{aligned}$$

Finally, similarly to (4.19), we have from an inverse inequality that

$$\left\|h_T^2\nabla(u_h-f)\right\|_T \leq C\left\|h_T(u_h-f)\right\|_T \leq C\left[\|h_T(\overline{u}_h-\overline{f})\|_T + \mathrm{osc}_h(u_h-f,T)\right]. \quad (4.22)$$

Combining (4.21) and (4.22) concludes the proof. □

Remark 4.2. The error quantities in Lemmas 4.7 and 4.8 contain additional terms, *e.g.* $\|h_T(w_h-w)\|$ and $\|h_T(u_h-u)\|_T$, which are not present in the error expression for the upper bound, *cf.* Corollary 4.1. Therefore, we are not able to combine these two lemmas in order to obtain a lower bound that corresponds precisely to the upper error estimate in Corollary 4.1, *i.e.* a lower bound for the error

$$\gamma^2\,\|\nabla(u_h-u)\|^2 + \tau\,\|\nabla(w_h-w)\|^2.$$

Naturally, such a lower bound would be desirable, as it would give a theoretical proof of the efficiency of the derived *a posteriori* estimator.

5. Adaptive algorithms

In this section we introduce several mesh adaption strategies, that are based on the *a posteriori* error estimator derived in Section 4. Throughout this section, we assume that $f = u_h^{\text{old}}$, $g = \frac{1}{\gamma} u_h^{\text{old}}$ arise from a fully discrete approximation of (1.1), where u_h^{old} is the discrete solution from the previous time level. Hence f, g are piecewise linear functions on V_{old}^h, the finite element space from the previous time level, and they will only differ from $f_h = \mathcal{I}^h f$ and $g_h = \mathcal{I}^h g$, respectively, if $V_{\text{old}}^h \not\subseteq V^h$, *i.e.*, when mesh coarsening is employed.

We define the following local error indicators:

- $\eta_{u,T} = \frac{1}{2} \sum_{e \subset T} \left\| h_e^{1/2} [\nabla u_h]_e \right\|_e^2 + \frac{1}{\gamma^2} \|h_T(g + w_h - \sigma_h)\|_T^2 + \frac{1}{\gamma^2} \|h_T(g + w_h)\|_{T \cap (\cup_{S \in \mathcal{T}_h'} S)}^2$;
- $\eta_{w,T} = \frac{\tau}{2\gamma^2} \sum_{e \subset T} \left\| h_e^{1/2} [\nabla w_h]_e \right\|_e^2 + \frac{1}{\tau} \|h_T(u_h - f)\|_T^2$;
- $\eta_{c,T} = \left\| h^2 \nabla \Delta_h u_h \right\|_T^2 + \frac{1}{\tau} \left\| h^2 \nabla (u_h - f) \right\|_T^2$.

The global error indicators are then defined as a corresponding sum of local error indicators, *i.e.*, $\eta_u = \sum_{T \in \mathcal{T}^h} \eta_{u,T}$, $\eta_w = \sum_{T \in \mathcal{T}^h} \eta_{w,T}$, $\eta_c = \sum_{T \in \mathcal{T}^h} \eta_{c,T}$.
By using the above definition of the error indicators, Corollary 4.1 can be reformulated as

$$\|\nabla(u_h - u)\|^2 + \frac{\tau}{\gamma^2} \|\nabla(w_h - w)\|^2 \leq C \left[\eta_u + \eta_w + \eta_c + \frac{1}{\gamma^2} \|g - g_h\|_{-1,h}^2 + \frac{1}{\tau} \|f - f_h\|_{-1,h}^2 \right]. \tag{5.1}$$

Further, in the numerical experiments we measured the relative error by the indicator defined as:

$$\eta_{\text{rel}} = \frac{\eta_u + \eta_w + \eta_c}{\|u_h\|_1}.$$

Remark 5.1. The error contributions in (5.1) can be classified as follows

- η_u corresponds to the discretization error of u;
- η_w corresponds to the discretization error of w;
- η_c corresponds to the consistency error caused by the use of the mass lumped product $(\cdot, \cdot)_h$;
- the terms $\frac{1}{\gamma^2} \|g - g_h\|_{-1,h}^2$, $\frac{1}{\tau} \|f - f_h\|_{-1,h}^2$ correspond to the error in the approximation of the solution from the previous time-level caused by mesh coarsening, *i.e.*, they are zero if no elements are coarsened;
- we introduce a heuristic indicator η_τ for time step control as follows $\eta_\tau = \frac{1}{\gamma^2} \|u_h - g_h\|_1^2$.

Remark 5.2. The discrete dual norm $\|\cdot\|_{-1,h}$ is difficult to compute in practice, *cf.* [24], Remark 5.2. Instead of using the dual norm we define a simple coarsening indicator using the L_2 norm as follows:

$$\eta_{h,T} = \frac{1}{\gamma^2} \|g - I^h g\|_T^2 \geq \frac{1}{\gamma^2} \|g - I^h g\|_{-1,T}^2 \geq \frac{1}{\gamma^2} \|g - g_h\|_{-1,h,T}^2.$$

Also note that for our choice $f = u_h^{\text{old}}$, $g = \frac{1}{\gamma} u_h^{\text{old}}$, we have that $\frac{1}{\gamma^2} \|g - g_h\|_{-1,h}^2 = \frac{\tau}{\gamma^4} \left(\frac{1}{\tau} \|f - f_h\|_{-1,h}^2 \right)$. Hence the term $\frac{1}{\tau} \|f - f_h\|_{-1,h}^2$ can be neglected, when $\tau = \mathcal{O}(\gamma^2)$, which is generally the case in our experiments.

Below we outline the detailed definitions of the adaptive algorithms that we used for our numerical experiments. The heuristic adaptive algorithm $(VOL1)$ was used in [3,7] for computations for the degenerate Cahn–Hilliard equation. The idea of the algorithm $(VOL1)$ is to locally refine the mesh in such a way, that one has uniformly small elements of a prescribed volume $\mathrm{vol}(T) \leq vol_f$ for $T \in \mathcal{T}_h \setminus \mathcal{C}_h(u_h)$. The elements $T \in \mathcal{C}_h(u_h)$ are coarsened if $\mathrm{vol}(T) \leq vol_c/2$ and refined if $\mathrm{vol}(T) > vol_c$. Note, that in our implementation we set $vol_f = h_{\min}^2/2$, $vol_c = h_{\max}^2/2$ in 2d and $vol_f = h_{\min}^3/6$, $vol_c = h_{\max}^3/6$ in 3D, where $h_{\min}$ and $h_{\max}$ are given desired minimum and maximum mesh sizes, respectively.

The second adaptive algorithm $(VOL2)$ is based on the observation that the estimator attains maximum values at the elements from the discrete boundary $\mathcal{F}_h(u_h)$. The $(VOL2)$ algorithm is similar to the $(VOL1)$ algorithm with the addition of an adaptive control of the constants $vol_f < vol_c$ to keep the value of η_u below a prescribed tolerance.

Algorithm (VOL2)

(1) compute u_h;
(2) for all $T \in \mathcal{T}_h$;
  if $T \in \mathcal{F}_h$ and $\mathrm{vol}(T) > vol_f$ mark T for refinement;
  if $T \in \mathcal{N}_h$ and $\mathrm{vol}(T) > 2vol_f$ mark T for refinement, else if $\mathrm{vol}(T) \leq vol_f$ mark T for coarsening;
  if $T \in \mathcal{C}_h$ and $\mathrm{vol}(T) \leq vol_c/2$ mark T for coarsening, else if $\mathrm{vol}(T) > vol_c$ mark T for refinement;
(3) refine/coarsen mesh; if no elements were refined/coarsened proceed with step 4 else proceed with step 2;
(4) compute η_u, if $\eta_u > TOL$ set $vol_f := vol_f/2$ and proceed with step 1, else proceed with step 5;
(5) if $\eta_\tau > TOL_\tau$ decrease time step $\tau := \tau/2$; if $\eta_\tau < 0.01\,TOL_\tau$ increase time step $\tau := \min\{2 \times \tau, \tau_{\max}\}$;
(6) proceed to the next time level.

The adaptive algorithm (MAX) is similar to the maximum error adaptive strategy from [2] and is described below. For given tolerances TOL and TOL_τ, and coarsening/refinement parameters ϵ_c, ε_c, ε_r, vol_f, vol_c we start with the mesh from the previous time step, *i.e.*, $\mathcal{T}^h = \mathcal{T}^h_{\mathrm{old}}$, and improve the mesh for the next time level with the following steps, where we use the notation $\eta_{\max} := \max_{T \in \mathcal{T}^h} \eta_{u,T}$.

Algorithm (MAX)

(1) compute u_h and $\eta_{u,T}$, $\eta_{h,T}$, $\forall T \in \mathcal{T}^h$;
(2) for all $T \in \mathcal{T}^h$, if $\eta_u > TOL$ and $\eta_{u,T} > \varepsilon_r \eta_{\max}$ mark T for refinement; if $\eta_{u,T} + \eta_{h,T} < \varepsilon_c \eta_{\max}$ mark T for coarsening;
(3) if $\eta_\tau > TOL_\tau$ decrease time step $\tau := \tau/2$; if $\eta_\tau < 0.01\,TOL_\tau$ increase time step $\tau := \min\{2 \times \tau, \tau_{\max}\}$;
(4) proceed to the next time level.

The constants ε_r, ε_c were chosen as 0.6 and 0.05, respectively. Note, that the algorithm (MAX) really only uses the indicator η_u for the mesh refinement. As confirmed by the numerical experiments below, this also guarantees the control over the remaining error contributions in practice.

Remark 5.3. We note that the coarsening estimate η_h was not employed in the adaptive strategy in [2]. The coarsening estimate is critically important when computing spinodal decomposition, where mesh coarsening may lead to an excessive loss of information and an unphysical rise of the discrete analogue of the free energy (1.3).

We used a Uzawa-multigrid algorithm for the solution of the discrete system of nonlinear algebraic equations arising from (2.2). For more details on this iterative solver see [3,4].

6. Numerical results

6.1. Failure of the non-localized estimator

We demonstrate that a localized estimator is essential for efficient numerical computations. We compute an evolution of a square to a circle for $\gamma = \frac{1}{8\pi}$ on a time interval $(0, 10^{-4})$. We employ the adaptive strategy $(VOL1)$ with $h_{\min} = 1/32$, $h_{\max} = 1$. In Figure 1 we display for $t = 10^{-4}$ the computed solution u_h, the mesh,

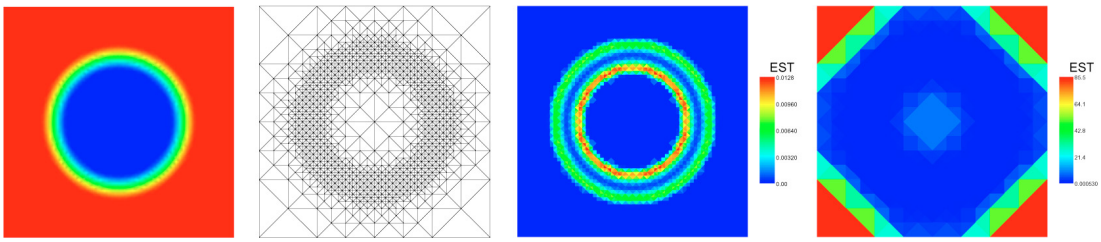

FIGURE 1. Solution u_h, the mesh and the indicators η_u, η_u^* at $t = 10^{-4}$.

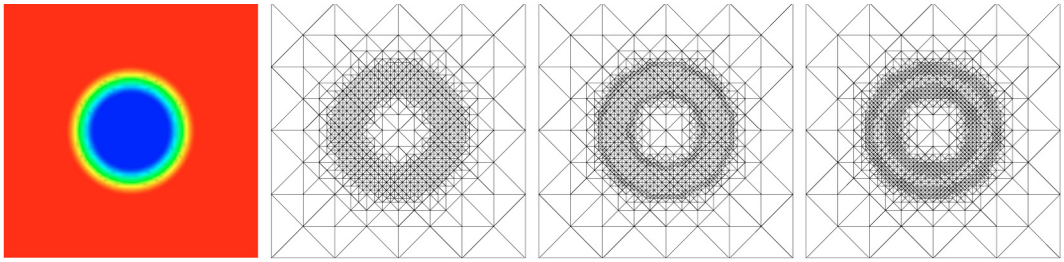

FIGURE 2. Solution u_h and adaptive meshes $VOL1$, $VOL2$, MAX.

the localized estimator η_u and non-localized estimator from Section 3 defined as:

$$\eta_u^* = \sum_{T\in\mathcal{T}^h} \left(\frac{1}{2} \sum_{e\subset T} \left\| h_e^{1/2} [\nabla u_h]_e \right\|_e^2 + \frac{1}{\gamma^2} \left\| h_T (g + w_h) \right\|_T^2 \right).$$

Clearly, the indicator η_u^* does not reflect the character of the solution properly and leads to a substantial overestimation of the error in the areas where the solution is constant. On the other hand, the localized indicator η_u is non-zero only in the interfacial region.

6.2. **Comparison of different adaptive strategies, discrete convergence**

We compare the adaptive algorithm ($VOL1$) with the adaptive algorithm with refinement along the free boundary ($VOL2$), the maximum strategy (MAX), and the uniform global mesh refinement.

In order to highlight the differences between the adaptive strategies ($VOL1$), ($VOL2$) and (MAX), we display in Figure 2 an example of meshes generated by the respective adaptive strategies.

We examine the convergence of η_u, η_w, η_c with respect to the number of degrees of freedom, with the help of an example computation for an established interface in the form of an ellipse and $\gamma = \frac{1}{8\pi}$. We computed with uniform time steps $\tau = \left(\frac{h_{\min}}{64}\right)^2 10^{-7}$, where $h_{\min}$ is the minimum mesh size in the respective computations. The behaviour was similar at all time levels, and we therefore only present the results at time $t = \hat{t} := 10^{-5}$. The profile of u_h at times $t = 0$ and $t = \hat{t}$ for a uniform mesh computation can be seen in Figure 3.

The graphs of the dependence of η_u, η_w, η_c on the number of vertices at time $\hat{t}$ are depicted in Figures 4–6, respectively. A logarithmic scaling is used in the figures, which allows us to interpret the slope $-\alpha$ as an experimental convergence rate of 2α, since $h \approx \#\mathcal{N}_h^{-2}$ in 2D. The above results support the assumption that the control of η_u in the presented adaptive algorithms (or in other words the refinement in the interfacial area only) is sufficient to guarantee the control over the remaining indicators η_w, η_c. The only qualitative difference between the uniform mesh refinement and adaptive strategies is in the convergence rates of η_w, which appears to be $\mathcal{O}(h^2)$ for uniform mesh refinement and "$\mathcal{O}(h)$" for adaptive mesh refinement. The difference can be

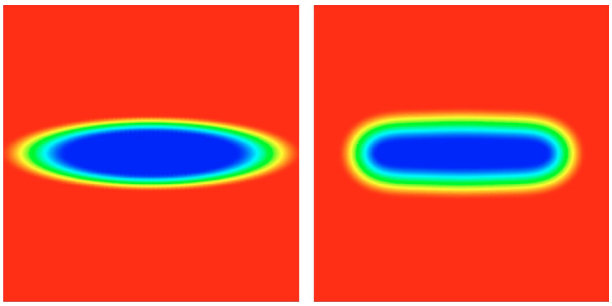

FIGURE 3. Solution u_h at $t = 0$ and $t = \hat{t}$ on a uniform mesh with $h = 1/64$.

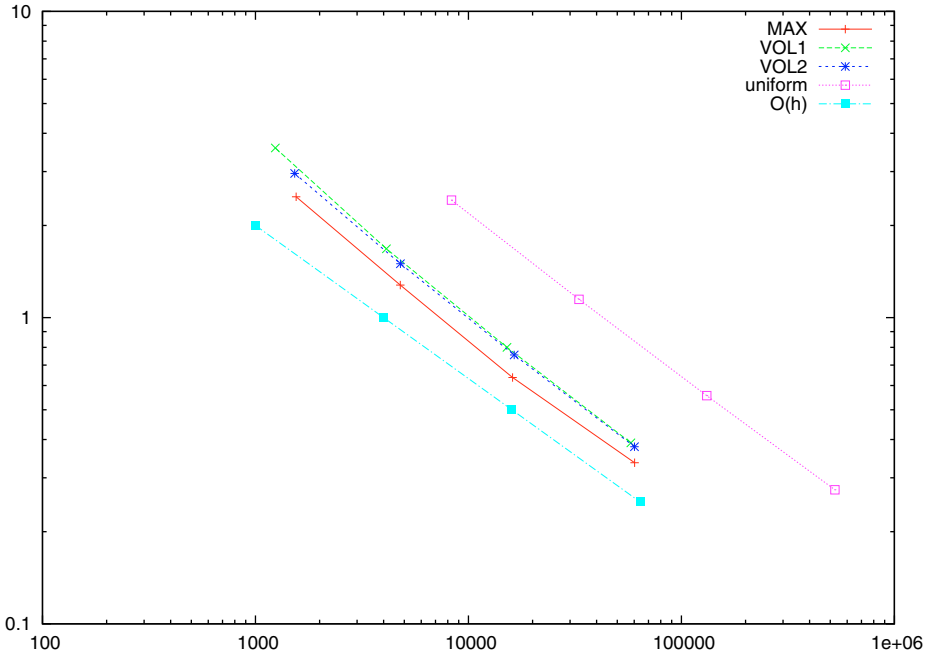

FIGURE 4. Convergence of η_u at $t = 10^{-5}$. Plot of estimator against number of degrees of freedom (dof).

accounted to the fact that η_w is not localized to the interfacial region, the region that is mainly refined by the adaptive methods. Note, that in the present case the worse convergence rates do not influence the overall convergence rate, which is $\mathcal{O}(h)$.

We conclude, that apart from the above disadvantage of the $(VOL1)$ algorithm there is no significant qualitative difference in the performance of the three adaptive algorithms. The algorithm (MAX) is perhaps the most flexible and effective of all three algorithms; however, its performance depends on the choice of the refinement/coarsening constants. The algorithm $(VOL1)$ is the simplest to implement.

6.3. Dependence of the estimator on γ

We study the efficiency of the adaptive algorithms with respect to the parameter γ. In order to obtain reliable results it is desired that the adaptive algorithm produces meshes for which the estimate $\eta_{\text{rel}}(\gamma) \leq TOL$,

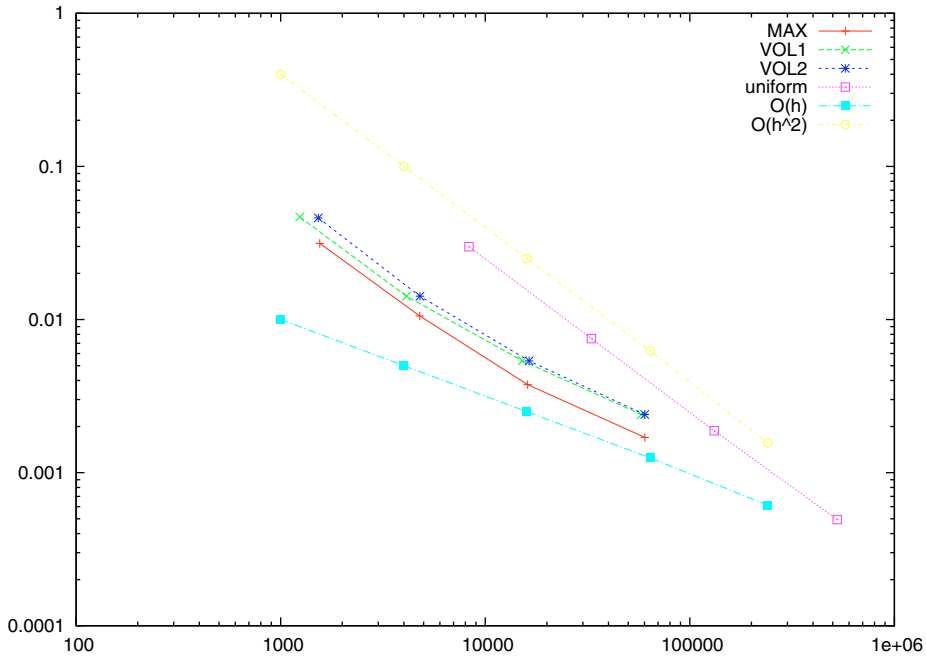

FIGURE 5. Convergence of η_w at $t = 10^{-5}$. Plot of estimator against dof.

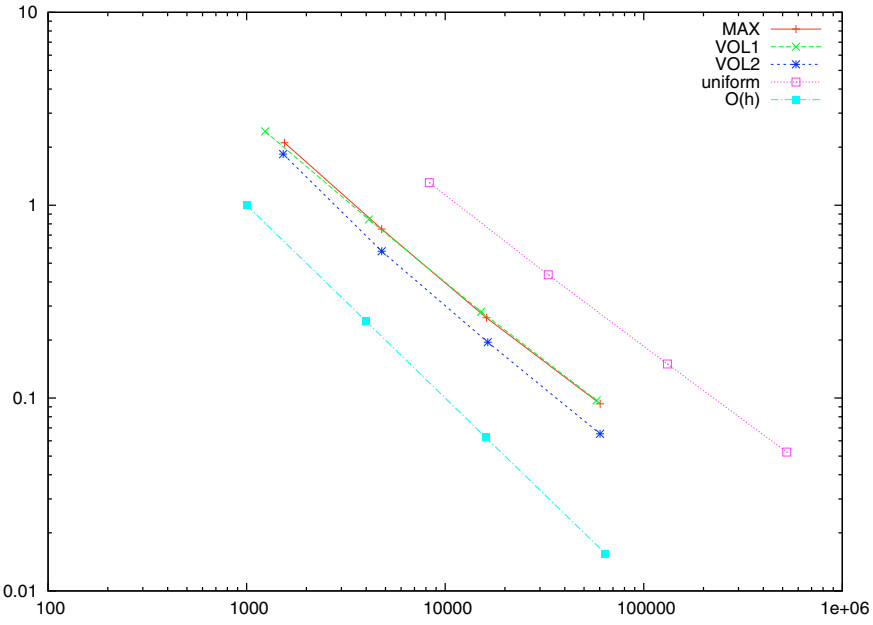

FIGURE 6. Convergence of η_c at $t = 10^{-5}$. Plot of estimator against dof.

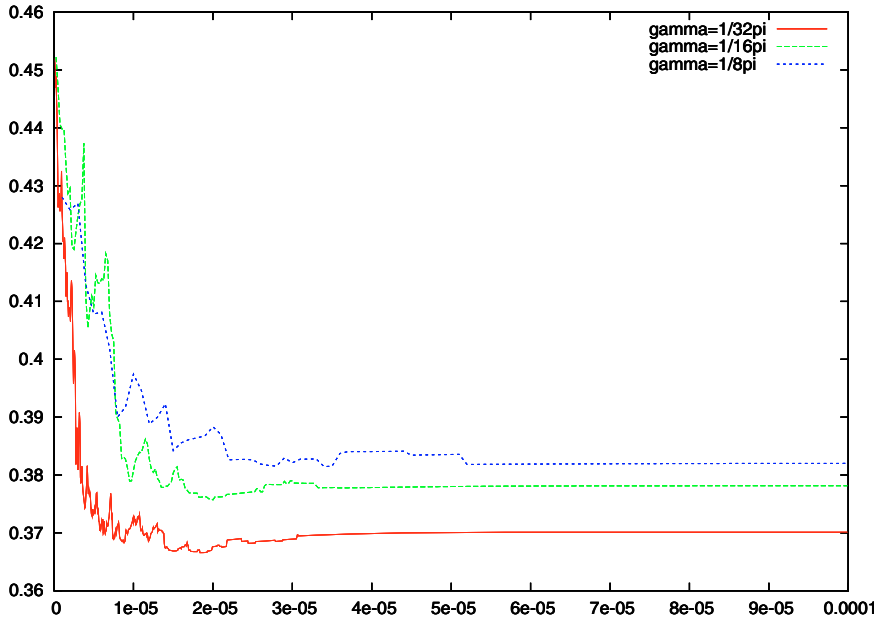

FIGURE 7. Evolution of η_{rel} for $\gamma = \frac{1}{8\pi}, \frac{1}{16\pi}, \frac{1}{32\pi}$ for the (MAX) algorithm.

where TOL is a tolerance independent of γ. On the other hand, in order to obtain an efficient adaptive mesh refinement, the minimum mesh size $h_{\min}(\gamma)$, needed in order to keep the error below a given tolerance, should have a linear dependence on γ.

We computed an evolution of a square using the adaptive algorithms $(VOL1)$ and (MAX). In Figures 7 and 8 we display the time evolution of η_{rel} for $\gamma_i = \frac{1}{2^i 8\pi}$, $i = 0, 1, 2$ for the two algorithms. The parameters for the adaptive mesh refinement in the $(VOL1)$ algorithm were chosen as $h_{\min} = \frac{\gamma_i}{\gamma_0 128}$, $h_{\max} = \frac{\gamma_i}{\gamma_0}$ for $i = 0, 1, 2$. The tolerance in the (MAX) algorithm was $TOL = 0.45$, which resulted in similar maximum and minimum mesh sizes in both algorithms. In order to exclude the influences of the adaptive time stepping on the error we used a fixed time step $\tau = \frac{(\gamma^i)^2 10^{-6}}{(\gamma^0)^2}$. The results show that the evolution of η_{rel} is similar for different values of γ if the number of mesh points in the interface is kept constant (*i.e.* for the above choices of $h_{\min/\max}$). This is a natural requirement, which underlines the efficiency of the adaptive mesh refinement.

In Figure 9 we display the computed solution u_h and the underlying adaptive mesh obtained by the (MAX) algorithm for $\gamma = \frac{1}{8\pi}$.

6.4. **Spinodal decomposition**

In the next experiment we perform an example computation of spinodal decomposition. The initial data is obtained by defining a coarse solution $\tilde{u}_0$ as a random perturbation around 0 on a uniform mesh with $h = 1/20$. A smooth initial condition u_0 is then obtained as

$$u_0(x) = \int_\Omega \tilde{u}_0(y) \mathrm{e}^{-1000\,|y-x|^2} \mathrm{d}y.$$

Note, that the above integral is computed approximately.

We computed the example for $\gamma = \frac{1}{8\pi}$ using a uniform mesh with $h = 1/32$ and using the adaptive mesh refinement strategy (MAX). We used adaptive time-stepping based on the indicator η_τ, giving a time step size $10^{-12} \leq \tau \leq 1.1 \times 10^{-9}$. The solution and adaptive mesh at different time levels is displayed in Figure 10.

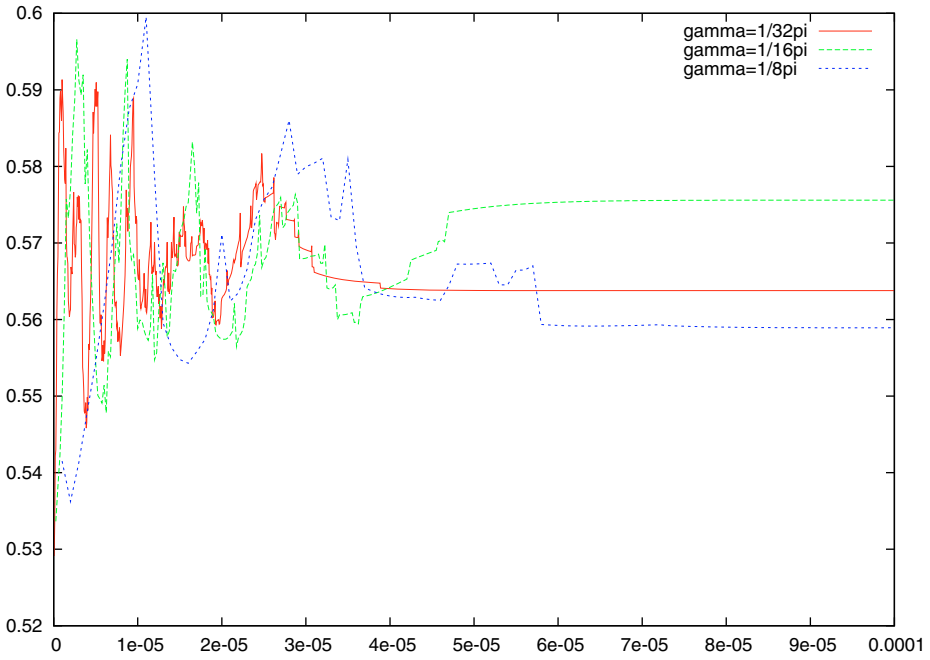

FIGURE 8. Evolution of η_{rel} for $\gamma = \frac{1}{8\pi}, \frac{1}{16\pi}, \frac{1}{32\pi}$ for the $(VOL1)$ algorithm.

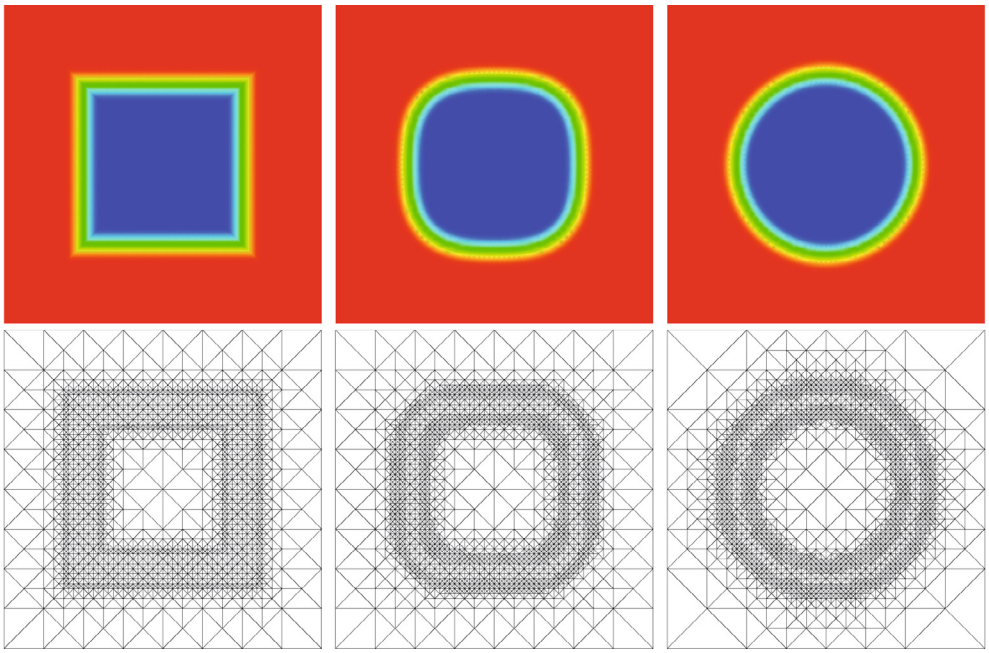

FIGURE 9. u_h at times $t = 0, 10^{-5}, 10^{-4}$ for $\gamma = \frac{1}{8\pi}$.

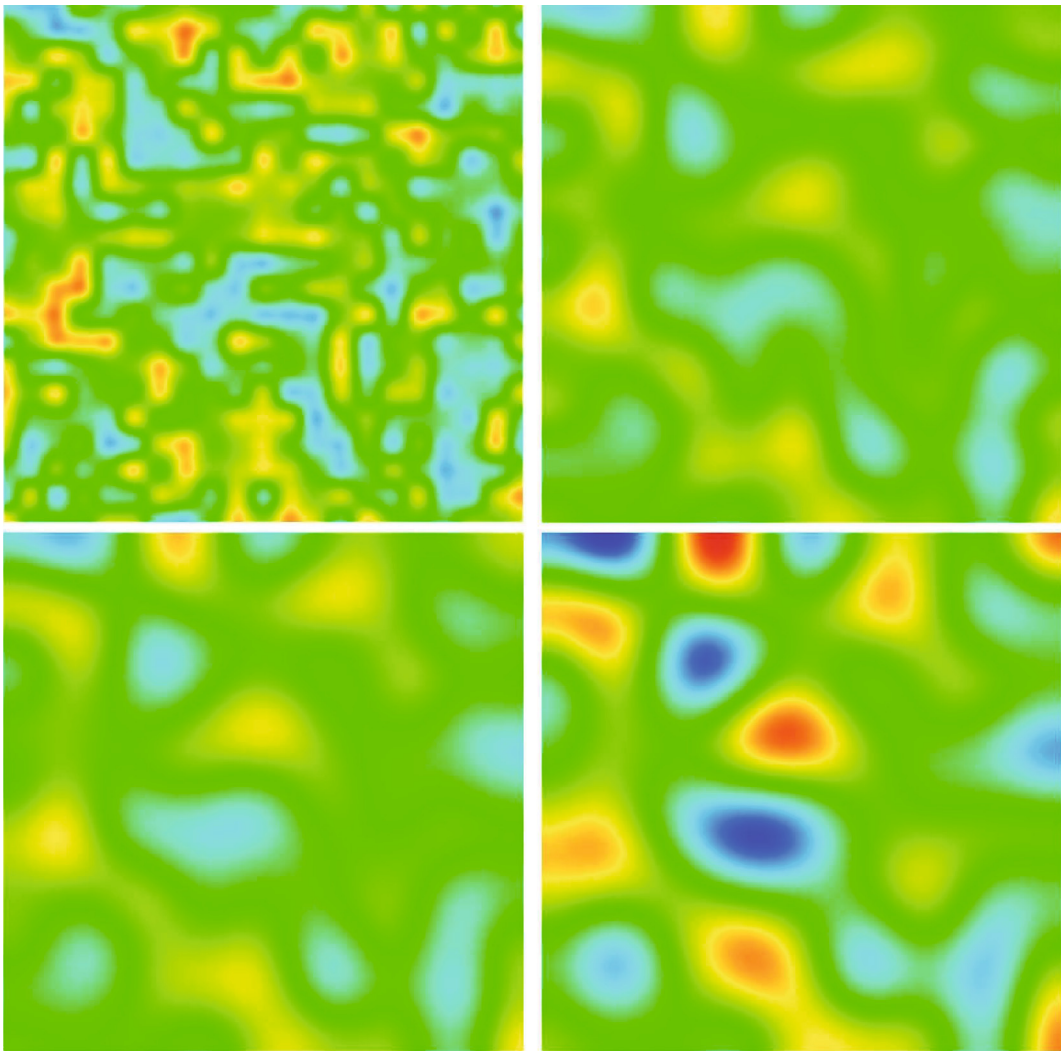

FIGURE 10. Spinodal decomposition: u_h at times $t = 0$, 10^{-6}, 3×10^{-6}, 10^{-5}.

The evolution of the indicator η_u is displayed in Figure 11. Clearly, the error on the uniform mesh is almost two times larger than the tolerance in the initial part of the computation, while the error on adaptive meshes is always below the tolerance. We found that too much coarsening in the computations of spinodal decomposition could lead to an unphysical rise in the discrete energy, which underlines the importance of the coarsening estimate.

6.5. **Coarsening in 3D**

The last experiment is to demonstrate the performance of adaptive mesh refinement in 3D computations. The zero level set of the initial condition consisted of two cubes of slightly different sizes. We computed the example using the adaptive strategy $(VOL1)$ with fixed time step $\tau = 10^{-6}$. The evolution of the zero level set of the computed solution and a cut through the adaptive mesh at $x_3 = 0$ are displayed in Figure 12. The evolution of η_{rel} in Figure 13 indicates a good control of the approximation error.

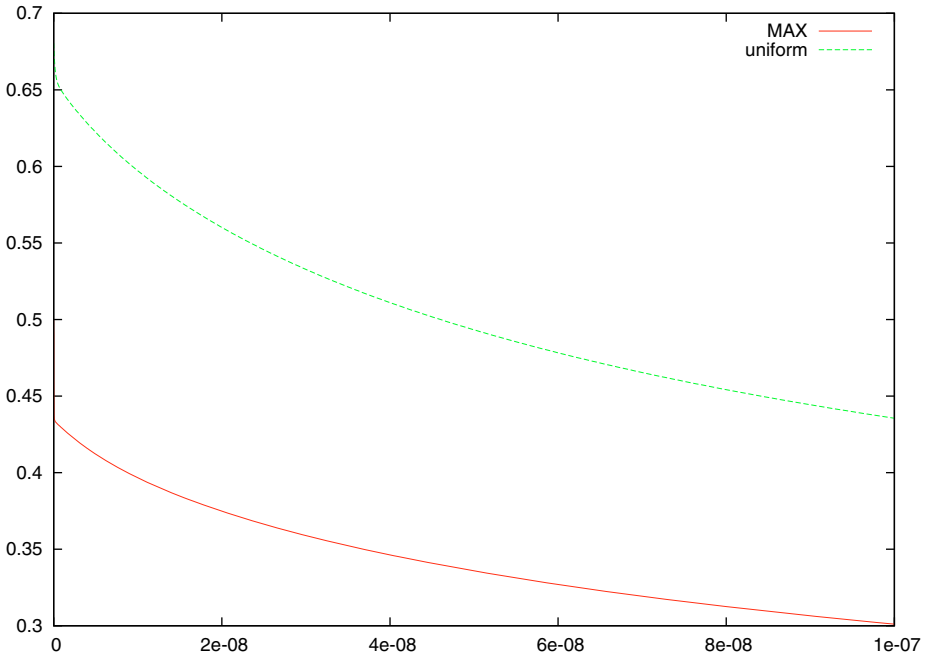

FIGURE 11. Evolution of the estimate η_u for a uniform mesh of fixed size $h = 1/32$ and the algorithm (MAX).

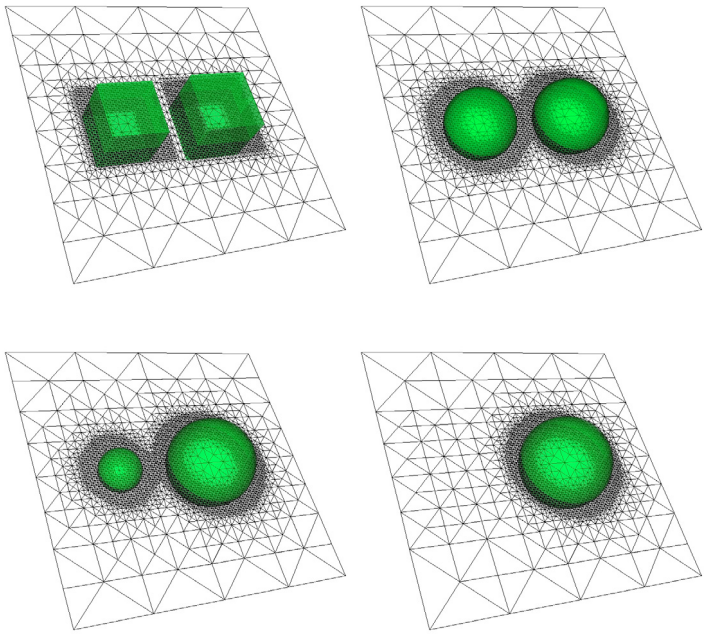

FIGURE 12. 3D coarsening: zero level set of u_h and cut through the mesh at $x_3 = 0$ at times $t = 0$, 10^{-5}, 10^{-4}, 1.2×10^{-4}.

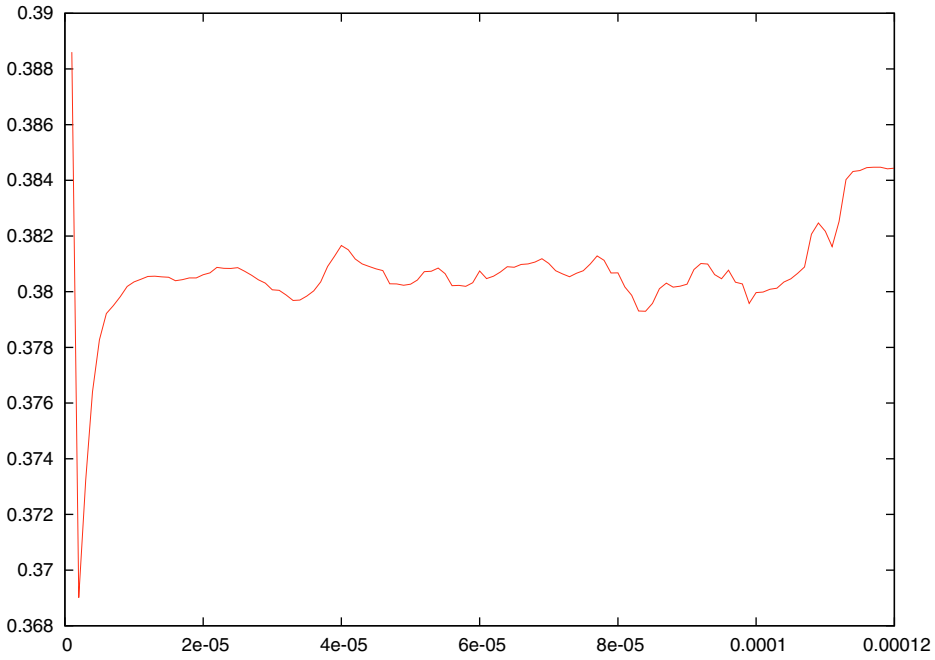

FIGURE 13. 3D coarsening: evolution of η_{rel}.

REFERENCES

[1] N.D. Alikakos, P.W. Bates and X.F. Chen, The convergence of solutions of the Cahn–Hilliard equation to the solution of the Hele–Shaw model. *Arch. Rational Mech. Anal.* **128** (1994) 165–205.
[2] Ľ. Baňas and R. Nürnberg, Adaptive finite element methods for Cahn–Hilliard equations. *J. Comput. Appl. Math.* **218** (2008) 2–11.
[3] Ľ. Baňas and R. Nürnberg, Finite element approximation of a three dimensional phase field model for void electromigration. *J. Sci. Comp.* **37** (2008) 202–232.
[4] Ľ. Baňas and R. Nürnberg, Phase field computations for surface diffusion and void electromigration in $\mathbb{R}^3$. *Comput. Vis. Sci.* (2008), doi: 10.1007/s00791-008-0114-0.
[5] J.W. Barrett and J.F. Blowey, Finite element approximation of a model for phase separation of a multi-component alloy with non-smooth free energy. *Numer. Math.* **77** (1997) 1–34.
[6] J.W. Barrett, J.F. Blowey and H. Garcke, Finite element approximation of the Cahn–Hilliard equation with degenerate mobility. *SIAM J. Numer. Anal.* **37** (1999) 286–318.
[7] J.W. Barrett, R. Nürnberg and V. Styles, Finite element approximation of a phase field model for void electromigration. *SIAM J. Numer. Anal.* **42** (2004) 738–772.
[8] J.F. Blowey and C.M. Elliott, The Cahn–Hilliard gradient theory for phase separation with non-smooth free energy. Part I: Mathematical analysis. *European J. Appl. Math.* **2** (1991) 233–279.
[9] J.F. Blowey and C.M. Elliott, The Cahn–Hilliard gradient theory for phase separation with non-smooth free energy. Part II: Numerical analysis. *European J. Appl. Math.* **3** (1992) 147–179.
[10] D. Braess, A posteriori error estimators for obstacle problems – another look. *Numer. Math.* **101** (2005) 415–421.
[11] J.W. Cahn, On spinodal decomposition. *Acta Metall.* **9** (1961) 795–801.
[12] J.W. Cahn and J.E. Hilliard, Free energy of a non-uniform system. I. Interfacial free energy. *J. Chem. Phys.* **28** (1958) 258–267.
[13] X. Chen, Spectrum for the Allen–Cahn, Cahn–Hilliard, and phase-field equations for generic interfaces. *Comm. Partial Differ. Equ.* **19** (1994) 1371–1395.
[14] Z. Chen and R.H. Nochetto, Residual type a posteriori error estimates for elliptic obstacle problems. *Numer. Math.* **84** (2000) 527–548.
[15] P. Clément, Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* **9** (1975) 77–84.
[16] C.M. Elliott and Z. Songmu, On the Cahn–Hilliard equation. *Arch. Rational Mech. Anal.* **96** (1986) 339–357.

[17] C.M. Elliott, D.A. French and F.A. Milner, A second order splitting method for the Cahn–Hilliard equation. *Numer. Math.* **54** (1989) 575–590.
[18] X. Feng and A. Prohl, Error analysis of a mixed finite element method for the Cahn–Hilliard equation. *Numer. Math.* **99** (2004) 47–84.
[19] X. Feng and H. Wu, A posteriori error estimates for finite element approximations of the Cahn–Hilliard equation and the Hele–Shaw flow. *J. Comput. Math.* **26** (2008) 767–796.
[20] M. Hintermüller and R.H.W. Hoppe, Goal-oriented adaptivity in control constrained optimal control of partial differential equations. *SIAM J. Control Optim.* **47** (2008) 1721–1743.
[21] M. Hintermüller, R.H.W. Hoppe, Y. Iliash and M. Kieweg, An *a posteriori* error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM: COCV* **14** (2008) 540–560.
[22] J. Kim, K. Kang and J. Lowengrub, Conservative multigrid methods for Cahn–Hilliard fluids. *J. Comput. Phys.* **193** (2004) 511–543.
[23] L. Modica, Gradient theory of phase transitions with boundary contact energy. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **4** (1987) 487–512.
[24] K.-S. Moon, R.H. Nochetto, T. von Petersdorff and C.-S. Zhang, *A posteriori* error analysis for parabolic variational inequalities. *ESAIM: M2AN* **41** (2007) 485–511.
[25] R.H. Nochetto and L.B. Wahlbin, Positivity preserving finite element approximation. *Math. Comp.* **71** (2002) 1405–1419.
[26] R.L. Pego, Front migration in the nonlinear Cahn-Hilliard equation. *Proc. Roy. Soc. London Ser. A* **422** (1989) 261–278.
[27] A. Veeser, Efficient and reliable a posteriori error estimators for elliptic obstacle problems. *SIAM J. Numer. Anal.* **39** (2001) 146–167.
[28] A. Veeser, On a posteriori error estimation for constant obstacle problems, in *Numerical methods for viscosity solutions and applications (Heraklion, 1999)*, M. Falcone and C. Makridakis Eds., *Ser. Adv. Math. Appl. Sci.* **59**, World Sci. Publ., River Edge, USA (2001) 221–234.
[29] R. Verfürth, *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Teubner-Wiley, New York (1996).