

## A posteriori error controlled local resolution of evolving interfaces for generalized Cahn–Hilliard equations

SÖREN BARTELS

*Institut für Numerische Simulation, Universität Bonn, D-53115 Bonn, Germany*

*E-mail: bartels@ins.uni-bonn.de*

RÜDIGER MÜLLER

*Weierstraß-Institut für Angewandte Analysis und Stochastik,*

*Mohrenstr. 39, 10117 Berlin, Germany*

*E-mail: mueller@wias-berlin.de*

[Received 5 December 2008 and in revised form 11 December 2009]

For equations of generalized Cahn–Hilliard type we present an a posteriori error analysis that is robust with respect to a small interface length scale  $\gamma$ . We propose the solution of a fourth order elliptic eigenvalue problem in each time step to gain a fully computable error bound, which only depends polynomially (of low order) on the inverse of  $\gamma$ . A posteriori and a priori error bounds for the eigenvalue problem are also derived. In numerical examples we demonstrate that this approach extends the applicability of robust a posteriori error estimation as it removes restrictive conditions on the initial data. Moreover we show that the computation of the principal eigenvalue allows the detection of critical points during the time evolution that limit the validity of the estimate.

### 1. Introduction

The Cahn–Hilliard equation determines the evolution of interfaces in phase separation processes on a mesoscopic length scale. This phase field model depends on a length scale  $0 < \gamma \leq 1$  that is related to the width of the diffuse interfaces where the order parameter  $\rho$  interpolates smoothly between the values  $-1$  and  $1$ . Moreover, for  $\gamma \rightarrow 0$  the Cahn–Hilliard equation approximates the motion of surfaces in Hele–Shaw flow and the Mullins–Sekerka model [1, 22]. Phase separation processes are of great importance in engineering applications because the coarsening of microstructure leads to aging of materials and might result in failure of devices. To model the demixing in solder alloys, it is necessary to include elastic stresses that are due to a lattice misfit between pure materials. These stresses are neglected in the Cahn–Hilliard model but become dominant at later stages of the time evolution [16]. The Cahn–Larché equations describe a phase field model that includes elastic effects [16, 9, 17, 18]. Their derivation is based on the formulation of a Ginzburg–Landau free energy functional. Let  $\Omega \subseteq \mathbb{R}^d$  be the domain occupied by a binary alloy and let the order parameter  $\rho(x) \in [-1, 1]$  be given by the difference between the volume fractions of the two components. We consider the energy functional

$$E(\rho, \vec{u}) := \int_{\Omega} \left( \frac{\gamma^2}{2} |\nabla \rho|^2 + \mathcal{F}(\rho) + \mathcal{W}(\rho, \mathcal{E}(\vec{u})) \right) dx, \quad (1)$$

where  $\mathcal{F}(\rho)$  is a double well potential that defines the stable states at  $\rho = \pm 1$ . Elastic effects are modeled by the term  $\mathcal{W}(\rho, \mathcal{E}(\vec{u}))$ , where  $\vec{u}$  is the displacement and  $\mathcal{E}(\vec{u})$  denotes the strain. The

interface dynamics is modeled as the  $H^{-1}$  gradient flow of (1), leading to the fourth order semilinear parabolic equation

$$\partial_t \rho - \Delta \left( -\gamma \Delta \rho + \frac{1}{\gamma} f(\rho) + \frac{1}{\gamma} W(\rho, \mathcal{E}(\vec{u})) \right) = 0, \quad (2)$$

where  $f(\rho) := \mathcal{F}'(\rho)$  and  $W(\rho, \mathcal{E}(\vec{u})) = \partial_\rho \mathcal{W}(\rho, \mathcal{E}(\vec{u}))$ . Because the time scale of mechanical relaxation is much smaller than the scale at which diffusion takes place, we may assume an elastic equilibrium state. The Cahn–Larché system consists of (2), a stationary equation of the form  $\mathcal{D}\vec{u} = \rho$  for the elasticity problem, and initial and boundary conditions.

Within this paper, we assume that deformations are small and the elasticity tensor is homogeneous, i.e. independent of  $\rho$ . As a consequence, there is a linear solution operator of the elasticity problem such that  $\vec{u} = \mathcal{D}^{-1}\rho$  and we can formally eliminate  $\vec{u}$  by writing  $W(\rho, \mathcal{E}(\vec{u})) =: A\rho$  with some linear operator  $A$ . This allows us to reduce the Cahn–Larché system to the single equation

$$\partial_t \rho - \Delta \left( -\gamma \Delta \rho + \frac{1}{\gamma} f(\rho) + \frac{1}{\gamma} A\rho \right) = 0 \quad (3)$$

with additional initial and boundary conditions. The error analysis presented in this paper is valid for generalized Cahn–Hilliard type equations of the form (3) with a linear selfadjoint and possibly nonlocal operator  $A$  that does not have to be related to elastic stresses.

For a fixed size of the parameter  $\gamma$ , the numerical analysis is well established for the Cahn–Hilliard equations [10, 7] and Cahn–Larché equations [17, 24, 18]. When  $\gamma$  becomes small, the solution  $\rho$  is of low effective regularity. Numerical approximation takes great advantage of mesh adaptivity, which in turn requires error control based on a posteriori estimates. Unfortunately, a straightforward analysis leads to error estimates with constants that depend exponentially on  $\gamma^{-1}T$ , where  $T$  is a time horizon. For the approximation of the sharp interface limit,  $T$  has to be independent of  $\gamma$ , and we remark that the time scaling in (2) is chosen accordingly. But then the outlined estimates become useless for  $\gamma \rightarrow 0$ . Robust estimates have been derived in several simpler situations, i.e. estimates that depend on  $\gamma^{-1}$  only through a low order polynomial. For the Allen–Cahn equation a robust a priori estimate was established in [13] and a posteriori estimates followed in [19, 3]. For the Cahn–Hilliard equation a robust a priori estimate was established in [14] and only very recently a first a posteriori estimate was derived [15]. All these results are based on spectral estimates of the linearized operator. The principal eigenvalue  $-\lambda$  of the linearized generalized Cahn–Hilliard operator is defined as the largest number  $\lambda$  for which there exists a nonvanishing function  $q$  such that

$$\lambda = \Delta \left( -\gamma \Delta q + \frac{1}{\gamma} f'(\rho)q + \frac{1}{\gamma} Aq \right).$$

For the smooth evolution of interfaces it was proven in [5, 8] that the principal eigenvalue of the linearized Allen–Cahn and Cahn–Hilliard operator is bounded from below uniformly with respect to  $\gamma^{-1}$ . This holds provided that the order parameter  $\rho$  has an admissible profile across the interface, and only as long as no topological changes occur. For more general operators there are no spectral estimates available.

To avoid restrictive assumptions on the initial data, we follow the approach of [3] where, in contrast to the references given above, the linearization is taken at the approximate solution  $\rho_h$  instead of the exact solution  $\rho$ . We propose to numerically approximate the principal eigenvalue  $-\lambda$

in each time step and thereby measure the stability of the evolution. The numerical computation of the principal eigenvalue is of great importance for the Cahn–Larché system since, as mentioned above, we are lacking a theoretical spectral estimate for these equations. Moreover it fits well the methodology of a posteriori error estimation as it provides important information about the approximation such as the detection of critical points of the nonlinear evolution. Our main result states that robust a posteriori error estimation is possible as long as an approximation of  $-\lambda$  remains uniformly bounded from below in a time interval  $[\gamma, T]$  (cf. Theorem 3.4 and Remark 3.5 below). Formal results of [22] suggest that an initially perturbed phase field develops an admissible profile within a short time frame and our numerical experiments show that the principal eigenvalue is bounded uniformly with respect to  $\gamma^{-1}$  after a time period of order  $O(\gamma)$ . It turns out that this is sufficient to gain a useful error estimate on  $[0, \gamma]$ , which together with our main result allows robust error control in the time interval  $[0, T]$ .

In the context of this paper, there are several differences from the situation related to the Allen–Cahn equation, which is only of second order, is nonconservative, and obeys a maximum principle. Here, we have to measure the error in the weaker  $H^{-1}$  norm. This makes it significantly harder to get control on the superquadratic term in the error, so at least for space dimension  $d = 3$  we have to impose a nonstandard growth condition on the potential function  $f(\rho)$ . We remark that if a priori bounds are employed, then our error analysis could be modified to cover the case of more general potentials. This however is not in the spirit of a posteriori error estimation. Since total mass is conserved by the Cahn–Hilliard equation, i.e.

$$\bar{\rho}(t) := \frac{1}{|\Omega|} \int_{\Omega} \rho(t, x) \, dx = \frac{1}{|\Omega|} \int_{\Omega} \rho_0(x) \, dx =: \bar{\rho}_0,$$

we obtain different constraints on the solution and the particular representation of the error defined below. The eigenvalue problem which has to be approximated is also of fourth order and we derive a computable a posteriori error bound for this problem.

The outline of this article is as follows. In the next section we introduce the weak formulation of the problem in question and the relevant finite element method. The main result is stated in Section 3, where the aforementioned a posteriori error estimate is proven. This estimate is formulated in an abstract way that does not depend on a specific numerical scheme. The necessary ingredients to get a fully computable estimate are given in the subsequent sections. In Section 4 we present an a posteriori upper bound for the numerically computed eigenvalue. A finite element method for the Cahn–Larché system is given in Section 5 along with the corresponding residual estimators. Finally, the numerical experiments in Section 6 show that the computation of the principal eigenvalue allows the detection of critical points during the time evolution that limit the validity of the developed a posteriori estimate. The results also demonstrate that restrictive assumptions on the initial data are not needed because the principal eigenvalue is uniformly bounded after sufficiently short times.

## 2. Problem formulation and finite element spaces

Let  $\Omega \subset \mathbb{R}^d$ ,  $d = 2, 3$ , be a bounded Lipschitz domain and  $T > 0$  a time horizon. Since we are interested in robust estimates for small  $\gamma$  we assume  $0 < \gamma \leq 1$ . We use standard notation for Lebesgue and Sobolev spaces and denote the inner product of  $L^2(\Omega)$  by  $(v, w)$ . The duality pairing between a Banach space  $X$  and its dual  $X^*$  is written  $\langle \cdot, \cdot \rangle$ . We define

$$\bar{v} := \int_{\Omega} v \, dx, \quad \dot{H}^1(\Omega) := \{v \in H^1(\Omega) : \bar{v} = 0\}.$$

The inverse Laplacian with natural boundary conditions  $\Delta_N^{-1} : \dot{H}^1(\Omega) \rightarrow \dot{H}^1(\Omega)$  is defined by

$$(\nabla \Delta_N^{-1} v, \nabla \varphi) = -(v, \varphi) \quad \text{for all } v \in \dot{H}^1(\Omega) \text{ and } \varphi \in H^1(\Omega).$$

To reduce the regularity requirements on the solution  $\rho$ , we use the mixed form of [11] and introduce the *chemical potential*  $w$  defined as

$$w := -\gamma \Delta \rho + \gamma^{-1}(f(\rho) + A\rho).$$

Based on the solution space

$$X_{\text{CH}} := (L^2([0, T], H^1(\Omega)) \cap H^1([0, T], H^1(\Omega)^*)) \times L^2([0, T], H^1(\Omega)),$$

the mixed variational formulation of the Cahn–Hilliard type equation with natural boundary conditions reads as follows:

$$(P) \quad \begin{cases} \text{Given } \rho(0, x) = \rho_0(x) \in H^1(\Omega), \text{ find } (\rho, w) \in X_{\text{CH}} \text{ such that for almost all } t \in (0, T), \\ \langle \varphi, \partial_t \rho \rangle + (\nabla \varphi, \nabla w) = 0 & \text{for all } \varphi \in H^1(\Omega), \\ (\psi, w) - \gamma (\nabla \psi, \nabla \rho) = \gamma^{-1}(\psi, f(\rho) + A\rho) & \text{for all } \psi \in H^1(\Omega). \end{cases}$$

For the pure Cahn–Hilliard equation, i.e. for  $A = 0$ , the existence and uniqueness of a global solution to (P) has been established in [12, 11]. The existence of a solution to the Cahn–Larché system was proven in [17], together with a uniqueness result for the case of homogeneous elasticity. We restrict our analysis to continuous potential functions, of which the most common example is  $\mathcal{F}_4(\rho) := (\rho^2 - 1)^2/4$ . More realistic nonsmooth logarithmic potential functions have been proposed and quantitatively studied e.g. in [2, 7] but they require further research. The following assumptions on the double well potential and the linear operator  $A$  are essential for our analysis:

(A1)  $f \in C^1(\mathbb{R})$  and there is a constant  $C_f \geq 0$  such that  $-f' \leq C_f$ .

(A2) There are a function  $g^*$  and constants  $C_\delta \geq 0$  and  $\delta \in (0, 1]$ , with  $\delta \leq 4/5$  if  $d = 3$ , such that for all  $a, b \in \mathbb{R}$ ,

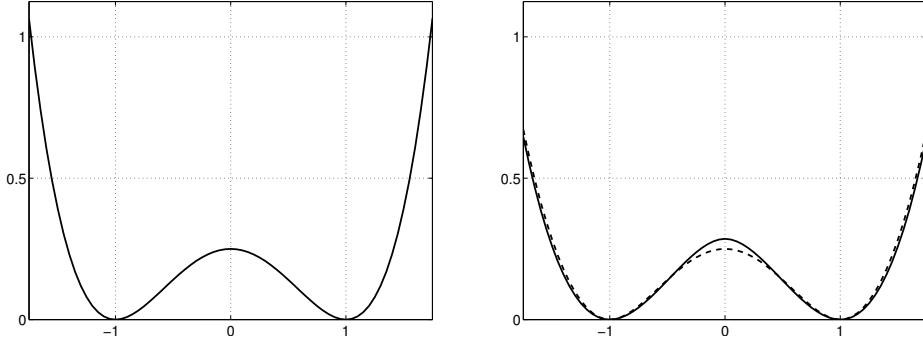
$$-(b - a)(f(b) - f(a)) \leq -f'(b)(b - a)^2 + g^*(b)C_\delta |b - a|^{2+\delta}.$$

(A3)  $A : L^2(\Omega) \rightarrow L^2(\Omega)$  is a bounded linear selfadjoint operator and there are constants  $\alpha, C_A > 0$  such that

$$\begin{aligned} -(\varphi, A\varphi) &\leq \alpha \|\varphi\|_{L^2(\Omega)}^2 && \text{for all } \varphi \in L^2(\Omega), \\ (\psi, A\varphi) &\leq C_A \|\nabla \psi\|_{L^2(\Omega)} \|\nabla \varphi\|_{L^2(\Omega)} && \text{for all } \psi, \varphi \in \dot{H}^1(\Omega). \end{aligned}$$

Because there is no maximum principle available for the Cahn–Hilliard equation, the properties of  $f$  outside of  $[-1, 1]$  are of relevance. If  $d = 3$  we cannot use  $f_4 := \mathcal{F}'_4$  on all of  $\mathbb{R}$  since  $f_4$  fails to satisfy (A2). Nevertheless, a proper modification of  $f_4$  outside of  $[-1, 1]$  is possible or alternatively a suitable replacement for the free energy function may be used.

LEMMA 2.1 (a) Let  $f = \mathcal{F}'_4$  for the quartic potential  $\mathcal{F}_4$ . If  $d = 2$ , then  $f$  satisfies (A1) and (A2) with  $C_f = 1$ ,  $\delta = 1$ ,  $C_\delta = 1$  and  $g^*(b) = 3b$ .

FIG. 1. Left: quartic potential  $\mathcal{F}_4$ ; right: alternative potentials:  $\mathcal{F}_1$  and  $\mathcal{F}_2$  (dashed).

- (b) Let  $0 < \delta \leq 1$  and  $f \in C^{1,\delta}(\mathbb{R})$ , a continuously differentiable function with a Hölder continuous derivative with exponent  $\delta$  or Lipschitz continuous derivative if  $\delta = 1$ . Then  $f$  satisfies (A2).
- (c) Define

$$\mathcal{F}_1(x) := \frac{1}{2+\delta}|x|^{2+\delta} - \frac{1}{2}x^2 + \frac{\delta}{4+2\delta},$$

$$\mathcal{F}_2(x) := \begin{cases} \frac{2}{1+\delta} \left( \frac{1}{2+\delta}|x|^{2+\delta} + x \right) + \frac{2}{2+\delta} & \text{if } x < -1, \\ \frac{1}{4}x^4 - \frac{1}{2}x^2 + \frac{1}{4} & \text{if } |x| \leq 1, \\ \frac{2}{1+\delta} \left( \frac{1}{2+\delta}|x|^{2+\delta} - x \right) + \frac{2}{2+\delta} & \text{if } x > 1. \end{cases}$$

Then  $f_1(x) := \mathcal{F}_1'(x) = x|x|^\delta - x$  and  $f_2(x) := \mathcal{F}_2'(x)$  satisfy (A1) and (A2).

*Proof.* (a) Since  $f(x) = x^3 - x$ , we have the lower bound  $f'(x) = 3x^2 - 1 \geq -1 =: -C_f$ . Moreover, we note  $f''(x) = 6x$ . Then, for  $a, b \in \mathbb{R}$ , Taylor expansion yields

$$f(b) - f(a) = f_4'(b)(b-a) + f''(b)(b-a)^2/2 + f'''(b)(b-a)^3/6.$$

We multiply this identity by  $-(b-a)$  to obtain

$$\begin{aligned} -(b-a)(f(b) - f(a)) &= -f'(b)(b-a)^2 - f''(b)(b-a)^3/2 - (b-a)^4 \\ &\leq -f'(b)(b-a)^2 - 6b(b-a)^3/2. \end{aligned}$$

(b) Let  $a \neq b \in \mathbb{R}$  and  $f \in C^{1,\delta}(\mathbb{R})$ . By the mean value theorem  $f(b) - f(a) = f'(\xi)(b-a)$  for some  $\xi \in (\min(a, b), \max(a, b))$  and there is a constant  $C_\delta$  such that

$$f'(\xi) - f'(b) \geq -|f'(\xi) - f'(b)| \geq -C_\delta|\xi - b|^\delta \geq -C_\delta|a - b|^\delta.$$

Thus the assertion follows upon multiplying

$$\frac{f(b) - f(a)}{b-a} = f'(\xi) \geq f'(b) - C_\delta|b-a|^\delta$$

by  $-(b-a)^2$ .

(c) We first note  $f_1'(x) = (1 + \delta)|x|^\delta - 1 \geq -1$  and then check  $f_1 \in C^{1,\delta}(\mathbb{R})$ : Without loss of generality assume  $|a| \leq |b|$ ; then

$$\begin{aligned} |f_1'(b) - f_1'(a)| &= (1 + \delta)|b|^\delta - |a|^\delta = (1 + \delta)(|b|^\delta - |a|^\delta) \\ &\leq (1 + \delta)(|b - a| + |a|)^\delta - |a|^\delta \leq (1 + \delta)(|b - a|^\delta + |a|^\delta - |a|^\delta) \\ &= (1 + \delta)|b - a|^\delta, \end{aligned}$$

where in the last inequality we used that for  $\delta \leq 1$  the mapping  $x \mapsto x^\delta$  is concave. We easily verify that  $f_2$  and  $f_2'$  are continuous at  $x = \pm 1$ . Moreover  $f_2'$  is Lipschitz continuous on  $[-1, 1]$ , hence  $|f_2'(b) - f_2'(a)| \leq \tilde{C}|b - a| = \tilde{C}|b - a|^{1-\delta}|b - a|^\delta \leq 2^{1-\delta}\tilde{C}|b - a|^\delta =: C|b - a|^\delta$  for all  $a, b \in [-1, 1]$ . By the same arguments as above, we see that  $f_2'$  is Hölder continuous with exponent  $\delta$  on the intervals  $(-\infty, -1]$  and  $[1, \infty)$ , and by the triangle inequality we conclude that  $f_2 \in C^{1,\delta}(\mathbb{R})$ .  $\square$

We consider discretizations of (P) using piecewise affine continuous finite elements for the quantities  $\rho$  and  $w$ . Owing to the boundary conditions employed, the proof of an inf-sup condition is a straightforward matter. For related estimates for the biharmonic equation with different boundary conditions and Ciarlet–Raviart mixed finite elements, we refer the reader to [6]. Based on this stable pair, we specify in Section 5 a discretization of (P) for a particular case (CL), where  $A$  is related to linear elasticity.

We consider shape regular meshes  $\mathcal{T}$  without hanging nodes, which consist of simplicial elements and define the  $\mathcal{T}$ -elementwise constant function  $h_{\mathcal{T}} : \Omega \rightarrow \mathbb{R}$  by  $h_{\mathcal{T}}|_K := \text{diam}(K)$  for all  $K \in \mathcal{T}$ . The set of all element faces within  $\mathcal{T}$  is denoted by  $\mathcal{E}(\mathcal{T})$ , and  $\bigcup \mathcal{E}(\mathcal{T})$  is the skeleton  $\{x \in \overline{\Omega} : x \in E, E \in \mathcal{E}(\mathcal{T})\}$  of  $\mathcal{T}$ . To each edge  $E \in \mathcal{E}(\mathcal{T})$  we assign its diameter  $h_E$  and introduce the function  $h_{\mathcal{E}} \in L^\infty(\bigcup \mathcal{E}(\mathcal{T}))$  that satisfies  $h_{\mathcal{E}}|_E = h_E$  for all  $E \in \mathcal{E}$ . The space of elementwise affine continuous functions is denoted by  $\mathcal{S}$ . On the subspace  $\mathring{\mathcal{S}}$  of finite element functions having mean value zero, the discrete inverse Laplacian  $\Delta_{Nh}^{-1} : \mathring{\mathcal{S}} \rightarrow \mathring{\mathcal{S}}$  satisfies

$$(\nabla \Delta_{Nh}^{-1} v_h, \nabla \varphi_h) = -(v_h, \varphi_h) \quad \text{for all } v_h \in \mathring{\mathcal{S}} \text{ and } \varphi_h \in \mathcal{S}.$$

The operators  $\Delta_{\mathcal{T}}$  and  $\text{div}_{\mathcal{T}}$  satisfy  $\Delta_{\mathcal{T}}\varphi_h|_K = \Delta(\varphi_h|_K)$  for all  $\varphi \in \mathcal{S}$  and  $K \in \mathcal{T}$  and  $\text{div}_{\mathcal{T}} \xi_h|_K = \text{div}(\xi_h|_K)$  for all  $\mathcal{T}$ -elementwise affine vector fields  $\xi$ . To each face  $E \in \mathcal{E}$  we assign a unique normal vector  $\vec{n}$  and denote the neighboring elements  $K^+, K^- \in \mathcal{T}$  in such a way that  $E = K^+ \cap K^-$  and the normal  $\vec{n}$  points from  $K^+$  to  $K^-$ . Then the jump operator is defined by  $[[\varphi]] := \varphi|_{K^+} - \varphi|_{K^-}$  for all  $\mathcal{T}$ -elementwise affine functions  $\varphi$ . Let  $0 = t_0 < t_1 < \dots < t_M = T$  be a partition of the time interval  $[0, T]$ . At time step  $j$  the mesh is denoted by  $\mathcal{T}^{(j)}$ . Throughout this article, we abbreviate  $\mathcal{S}^{(j)} := \mathcal{S}(\mathcal{T}^{(j)})$  and  $\mathcal{E}^{(j)} := \mathcal{E}(\mathcal{T}^{(j)})$ .

**DEFINITION 2.2** (a) Let  $(\rho, w) \in X_{CH}$  be the solution of (P) and  $(\rho_h, w_h)$  a conforming finite element approximation. For almost all  $t \in [0, T]$ , we define the *errors*

$$\bar{e}_1 := \bar{\rho}_h - \bar{\rho}_0 \in \mathbb{R}, \quad \hat{e}_1 := \rho_h - \rho - \bar{e}_1 \in \mathring{H}^1(\Omega), \quad e_2 := w_h - w \in H^1(\Omega).$$

We set  $\widehat{\rho}_h := \rho_h - \bar{e}_1$  and introduce  $z := -\Delta_N^{-1} \hat{e}_1$ .

(b) For  $s \in (0, T)$  the *residuals*  $R_1(s), R_2(s) \in H^1(\Omega)^*$  of the approximation  $(\rho_h, w_h)$  are

$$\langle \varphi, R_1 \rangle := (\nabla \varphi, \nabla w_h) + \langle \varphi, \partial_t \rho_h \rangle \quad \text{for all } \varphi \in H^1(\Omega), \quad (4a)$$

$$\langle \psi, R_2 \rangle := \gamma(\nabla \psi, \nabla \rho_h) - (\psi, w_h) + \gamma^{-1}(\psi, f(\rho_h) + A\rho_h) \quad \text{for all } \psi \in H^1(\Omega). \quad (4b)$$

REMARK 2.3 (a) Because of  $\hat{e}_1 \in \hat{H}^1(\Omega)$ , by Poincaré's inequality there is a constant  $C_P$ , depending only on  $\Omega$ , such that  $\|\hat{e}_1\|_{L^2(\Omega)} \leq C_P \|\nabla \hat{e}_1\|_{L^2(\Omega)}$ . We will repeatedly apply Young's inequality with a positive factor  $\varepsilon > 0$  to insert appropriate powers of  $\gamma$ :

$$\|\hat{e}_1\|_{L^2(\Omega)}^2 = (\nabla \hat{e}_1, \nabla z) \leq \varepsilon \|\nabla \hat{e}_1\|_{L^2(\Omega)}^2 + \frac{1}{4\varepsilon} \|\nabla z\|_{L^2(\Omega)}^2. \quad (5)$$

(b) If  $\bar{\rho}_h^{(0)} = \bar{\rho}_0$  and a fixed triangulation  $\mathcal{T}^{(j)} = \mathcal{T}^{(0)}$  for  $j = 1, \dots, M$  is used, or if  $\mathcal{T}^{(j)}$  is a refinement of  $\mathcal{T}^{(j-1)}$  for all  $j \geq 0$ , then  $\bar{e}_1 = 0$ .

We note that if the finite element solution coincides with the exact solution then the residuals vanish. In order to derive an error equation, we choose  $\varphi = z$  in (4a), set  $\psi = \hat{e}_1$  in (4b) and subtract the weak formulation from the definition of the residuals, i.e.

$$\langle z, R_1 \rangle = (\nabla z, \nabla e_2) + \langle z, \partial_t \hat{e}_1 \rangle + \langle z, \partial_t \bar{\rho}_h \rangle, \quad (6a)$$

$$\langle \hat{e}_1, R_2 \rangle = \gamma \|\nabla \hat{e}_1\|_{L^2(\Omega)}^2 - (\hat{e}_1, e_2) + \frac{1}{\gamma} (\hat{e}_1, f(\rho_h) - f(\rho) + A\rho_h - A\rho). \quad (6b)$$

Note  $\langle z, \partial_t \hat{e}_1 \rangle = -\langle z, \partial_t \Delta z \rangle = \langle \nabla z, \partial_t \nabla z \rangle = \frac{1}{2} \frac{d}{dt} \|\nabla z\|_{L^2(\Omega)}^2$  and  $\langle z, \partial_t \bar{\rho}_h \rangle = 0$  because  $z \in \hat{H}^1(\Omega)$ . By definition of  $z$  we have  $(\hat{e}_1, e_2) = (\nabla z, \nabla e_2)$ . Thus, when adding (6a) and (6b), the mixed terms cancel and we obtain the identity

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla z\|_{L^2(\Omega)}^2 + \gamma \|\nabla \hat{e}_1\|_{L^2(\Omega)}^2 &= \langle z, R_1 \rangle + \langle \hat{e}_1, R_2 \rangle - \gamma^{-1} (\hat{e}_1, A\bar{e}_1) \\ &\quad - \gamma^{-1} (\hat{e}_1, f(\rho_h) - f(\rho) + A\hat{e}_1). \end{aligned} \quad (7)$$

### 3. A posteriori error estimate

Besides the error representation (7) our main result needs two more ingredients. First, we need residual estimators  $\eta_1$ ,  $\eta_{21}$  and  $\eta_{22}$  such that

$$\langle \varphi, R_1 \rangle \leq \eta_1 \|\nabla \varphi\|_{L^2(\Omega)}, \quad \langle \psi, R_2 \rangle \leq \eta_{21} \|\nabla \psi\|_{L^2(\Omega)} + \eta_{22} \|\psi\|_{L^2(\Omega)}, \quad (8)$$

for almost every  $t \in [0, T]$  and  $\varphi, \psi \in \hat{H}^1(\Omega)$ . The estimators depend on the specific discretization scheme used. When  $A$  is given by linear elasticity, i.e. the elasticity tensor is homogeneous, we derive in Section 5 computable estimates of the residuals  $R_1$  and  $R_2$ . Second, we need a lower bound  $\Lambda^*$  of the spectrum of the linearized operator,

$$-\Lambda^* := \inf_{s \in [0, T]} (-\Lambda - \eta_\Lambda) \leq \inf_{s \in [0, T]} -\lambda, \quad (9)$$

where  $-\lambda$  is the principal eigenvalue of the linearized operator at the corrected discrete approximation  $\widehat{\rho}_h$ , i.e.

$$-\lambda(s) := \inf_{\substack{q \in \hat{H}^1(\Omega) \setminus \{0\} \\ y = -\Delta_N^{-1} q}} \frac{\gamma \|\nabla q\|_{L^2(\Omega)}^2 + \gamma^{-1} (q, f'(\widehat{\rho}_h(s))q) + \gamma^{-1} (q, Aq)}{\|\nabla y\|_{L^2(\Omega)}^2}. \quad (10)$$

LEMMA 3.1 (a) If (A1) and (A3) hold then

$$\begin{aligned} -\gamma^{-1}(\dot{e}_1, f(\rho_h) - f(\rho) + A\dot{e}_1) &\leq \frac{(C_f + \alpha)^2}{\gamma^3} \|\nabla z\|_{L^2(\Omega)}^2 + \frac{1}{4}\gamma \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 \\ &\quad - \gamma^{-1}(\dot{e}_1, f(\widehat{\rho}_h) - f(\rho_h)). \end{aligned}$$

(b) Let  $-\Lambda^*$  as in (9) be a lower bound for the principal eigenvalue  $-\lambda$  in (10). If (A2) holds then

$$\begin{aligned} -\gamma^{-1}(\dot{e}_1, f(\rho_h) - f(\rho) + A\dot{e}_1) &\leq \Lambda^* \|\nabla z\|_{L^2(\Omega)}^2 - \gamma^{-1}(\dot{e}_1, f(\widehat{\rho}_h) - f(\rho_h)) \\ &\quad + \gamma \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 + \|g^*(\rho_h)\|_{L^\infty(\Omega)} \frac{C_\delta}{\gamma} \|\dot{e}_1\|_{L^{2+\delta}(\Omega)}^{2+\delta}. \end{aligned}$$

(c) Let  $\eta_1, \eta_{21}$  and  $\eta_{22}$  be the residual estimators as in (8) and define

$$\eta^2 := \frac{1}{4}\eta_1^2 + \frac{2}{\gamma^4}\eta_{21}^2 + \frac{4}{\gamma^2}\eta_{22}^2 + \frac{C_P^2}{\gamma^6} \|f(\rho_h) - f(\widehat{\rho}_h)\|_{L^2(\Omega)}^2 + \frac{4}{\gamma^4} \|A\bar{e}_1\|_{L^2(\Omega)}^2.$$

Then

$$\begin{aligned} \langle z, R_1 \rangle + \langle \dot{e}_1, R_2 \rangle - \gamma^{-1}(\dot{e}_1, f(\widehat{\rho}_h) - f(\rho_h) + A\bar{e}_1) \\ \leq \eta^2 + \left(1 + \frac{1}{32}\right) \|\nabla z\|_{L^2(\Omega)}^2 + \frac{1}{2}\gamma^4 \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2. \end{aligned}$$

*Proof.* (a) From the fundamental theorem of calculus and assumption (A1) we get

$$-\gamma^{-1}\dot{e}_1[f(\widehat{\rho}_h) - f(\rho)] = \gamma^{-1}\dot{e}_1 \int_\rho^{\widehat{\rho}_h} -f'(\xi) d\xi \leq \gamma^{-1}C_f \dot{e}_1^2.$$

Integrating this identity over  $\Omega$  and using (A3) we deduce

$$\begin{aligned} -\gamma^{-1}(\dot{e}_1, f(\rho_h) - f(\rho) + A\dot{e}_1) &= -\gamma^{-1}(\dot{e}_1, f(\widehat{\rho}_h) - f(\rho) + A\dot{e}_1) - \gamma^{-1}(\dot{e}_1, f(\rho_h) - f(\widehat{\rho}_h)) \\ &\leq \gamma^{-1}(C_f + \alpha) \|\dot{e}_1\|_{L^2(\Omega)}^2 - \gamma^{-1}(\dot{e}_1, f(\rho_h) - f(\widehat{\rho}_h)), \end{aligned}$$

and then (5) with  $\varepsilon = (C_f + \alpha)\gamma^2/4$  proves the first assertion. For (b) we set  $q = \dot{e}_1 \in \dot{H}^1(\Omega)$  in (10) and thus by definition  $y = z$ . To prove (c), we apply Young's and Poincaré's inequality to obtain

$$\begin{aligned} \gamma^{-1}(\dot{e}_1, f(\rho_h) - f(\widehat{\rho}_h)) &\leq \gamma^{-1} \|f(\rho_h) - f(\widehat{\rho}_h)\|_{L^2(\Omega)} \|\dot{e}_1\|_{L^2(\Omega)} \\ &\leq \frac{C_P^2}{\gamma^6} \|f(\rho_h) - f(\widehat{\rho}_h)\|_{L^2(\Omega)}^2 + \frac{\gamma^4}{4} \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2. \end{aligned}$$

By Hölder's and Young's inequalities,  $-\frac{1}{\gamma}(\dot{e}_1, A\bar{e}_1) \leq \frac{1}{4\varepsilon_0\gamma^4} \bar{\eta}_A^2 + \varepsilon_0\gamma^2 \|\dot{e}_1\|_{L^2(\Omega)}^2$  with some  $\varepsilon_0 > 0$ , and applying (5) with  $\varepsilon = \gamma^2$ , we have

$$-\frac{1}{\gamma}(\dot{e}_1, A\bar{e}_1) \leq \frac{1}{4\varepsilon_0\gamma^4} \bar{\eta}_A^2 + \varepsilon_0\gamma^4 \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 + \frac{\varepsilon_0}{4} \|\nabla z\|_{L^2(\Omega)}^2.$$



In the same way we treat the residuals  $R_1$  and  $R_2$  to deduce that

$$\begin{aligned} \langle z, R_1 \rangle + \langle \dot{e}_1, R_2 \rangle &\leq \frac{1}{4}\eta_1^2 + \|\nabla z\|_{L^2(\Omega)}^2 + \frac{1}{4\varepsilon_{21}\gamma^4}\eta_{21}^2 + \varepsilon_{21}\gamma^4\|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 \\ &\quad + \frac{1}{4\varepsilon_{22}\gamma^2}\eta_{22}^2 + \varepsilon_{22}\gamma^4\|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 + \frac{\varepsilon_{22}}{4}\|\nabla z\|_{L^2(\Omega)}^2. \end{aligned}$$

Combining the last three estimates and choosing  $\varepsilon_{21} = 1/8$ ,  $\varepsilon_0 = \varepsilon_{22} = 1/16$ , we get the assertion.  $\square$

REMARK 3.2 Applying (a) and (c) of Lemma 3.1 to (7) we get

$$\frac{1}{2} \frac{d}{dt} \|\nabla z\|_{L^2(\Omega)}^2 + \frac{1}{4}\gamma \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 \leq \eta^2 + \left(1 + \frac{1}{32} + \frac{(C_f + \alpha)^2}{\gamma^3}\right) \|\nabla z\|_{L^2(\Omega)}^2.$$

Integration over  $[0, t]$  and application of Gronwall's lemma leads to

$$\sup_{s \in [0, T]} \|\nabla z(s)\|_{L^2(\Omega)}^2 + \frac{\gamma}{2} \int_0^T \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 ds \leq 2 \left[ \|\nabla z_0\|_{L^2(\Omega)}^2 + 2 \int_0^T \eta^2 ds \right] e^{2C_0\gamma^{-3}T},$$

with  $C_0 = (1 + 1/32)\gamma^3 + (C_f + \alpha)^2$ . This estimate has an exponential dependence on  $\gamma^{-1}$  and is only useful for  $T \lesssim \gamma^3$ . In [15] it was pointed out that this would be sufficient for robust error control if the principal eigenvalue is uniformly bounded after a certain time of order  $O(\gamma^3)$ . Our numerical experiments in Section 6 show that for general initial data such a bound can only be expected after some time of order  $O(\gamma)$ .

LEMMA 3.3 Suppose that  $\delta \in (0, 1]$ , with  $\delta \leq 4/5$  if  $d = 3$ . Then there is a constant  $C_{S,\delta}$  such that

$$\|\dot{e}_1\|_{L^{2+\delta}(\Omega)}^{2+\delta} \leq C_{S,\delta}^{2-\delta} \|\nabla z\|_{L^2(\Omega)}^\delta \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2.$$

*Proof.* If  $\delta < 1$ , then Hölder's inequality with exponents  $1/\delta$  and  $1/(1-\delta)$  implies

$$\int_{\Omega} |\dot{e}_1|^{2+\delta} dx \leq \|\dot{e}_1\|_{L^{1/\delta}(\Omega)}^{2\delta} \|\dot{e}_1\|_{L^{1/(1-\delta)}(\Omega)}^{2-2\delta} = \|\dot{e}_1\|_{L^2(\Omega)}^{2\delta} \|\dot{e}_1\|_{L^{(2-\delta)/(1-\delta)}(\Omega)}^{2-2\delta}. \quad (11)$$

If  $d = 2$ , there is a constant  $C_{S,\delta} > 0$  such that

$$\|v\|_{L^{(2-\delta)/(1-\delta)}(\Omega)} \leq C_{S,\delta} \|\nabla v\|_{L^2(\Omega)} \quad \text{for all } v \in \dot{H}^1(\Omega). \quad (12)$$

If  $d = 3$  and  $\delta \leq 4/5$ , then  $(2-\delta)/(1-\delta) \leq 6$ , and again there is a constant  $C_{S,\delta} > 0$  such that (12) holds. Noting  $\|\dot{e}_1\|_{L^2(\Omega)}^{2\delta} \leq \|\nabla \dot{e}_1\|_{L^2(\Omega)}^\delta \|\nabla z\|_{L^2(\Omega)}^\delta$ , we conclude in either case

$$\int_{\Omega} |\dot{e}_1|^{2+\delta} dx \leq \|\dot{e}_1\|_{L^2(\Omega)}^{2\delta} (C_{S,\delta} \|\nabla \dot{e}_1\|_{L^2(\Omega)})^{2-2\delta} \leq C_{S,\delta}^{2-\delta} \|\nabla z\|_{L^2(\Omega)}^\delta \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2.$$

Let  $d = 2$  and  $\delta = 1$ . According to the multiplicative Sobolev inequality [20, (2.10)] there is a constant  $C_{S,\delta} > 0$  such that

$$\|v\|_{L^4(\Omega)}^2 \leq C_{S,\delta} \|v\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \quad \text{for all } v \in \dot{H}^1(\Omega)$$

and we deduce with Hölder's inequality that

$$\begin{aligned} \int_{\Omega} |\dot{e}_1|^3 \, dx &\leq \|\dot{e}_1\|_{L^2(\Omega)} \|\dot{e}_1\|_{L^4(\Omega)}^2 \leq \|\dot{e}_1\|_{L^2(\Omega)} C_{S,\delta} \|\dot{e}_1\|_{L^2(\Omega)} \|\nabla \dot{e}_1\|_{L^2(\Omega)} \\ &\leq C_{S,\delta}^{2-\delta} \|\nabla z\|_{L^2(\Omega)}^\delta \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2, \end{aligned}$$

where in the last line we used  $\|\dot{e}_1\|_{L^2(\Omega)}^2 = (\nabla z, \nabla \dot{e}_1) \leq \|\nabla z\|_{L^2(\Omega)} \|\nabla \dot{e}_1\|_{L^2(\Omega)}$ .  $\square$

**THEOREM 3.4** Let  $\Lambda_o := 1 + 1/32 + (C_f + \alpha)^2$ ,  $\eta_{g^*} := \max(1, \sup_{t \in [0, T]} \|g^*(\rho_h)\|_{L^\infty(\Omega)})$  and define

$$\mu_1^2 := \left( \frac{e^{-2 \max(0, \Lambda_o + \Lambda^*)T}}{16 \eta_{g^*} C_\delta C_{S,\delta}^{2-\delta}} \right)^{1/\delta} \quad \text{and} \quad \mu_2^2 := \frac{3}{16} e^{-2 \max(0, \Lambda_o + \Lambda^*)T}.$$

Given a tolerance  $\theta \leq \gamma^{5/\delta} \mu_1^2$ , suppose the approximation error of the initial values and the residuals can be controlled by this tolerance  $\theta$  in the sense that

$$\|\nabla z_0\|_{L^2(\Omega)} + \sqrt{2} \|\eta\|_{L^2([0, T])} < \mu_2 \theta \leq \mu_1^2 \mu_2 \gamma^{5/\delta}, \quad (13)$$

where  $z_0 := z(0) = \Delta_N^{-1}(\rho_h^{(0)} - \rho_0 - (\bar{\rho}_h^{(0)} - \bar{\rho}_0))$ . Then

$$\sup_{s \in [0, T]} \|\nabla z(s)\|_{L^2(\Omega)}^2 + \frac{1}{2} \gamma^4 \int_0^T \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 \, ds \leq \theta^2.$$

*Proof.* Starting from (7) and taking from Lemma 3.1 a convex combination of  $\gamma^3$  times (a) plus  $(1 - \gamma^3)$  times (b) yields

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla z\|_{L^2(\Omega)}^2 + \frac{3}{4} \gamma^4 \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 &\leq \langle z, R_1 \rangle + \langle \dot{e}_1, R_2 \rangle - \gamma^{-1} (\dot{e}_1, f(\widehat{\rho}_h) - f(\rho_h) + A \bar{e}_1) \\ &\quad + ((C_f + \alpha)^2 + (1 - \gamma^3) \Lambda^*) \|\nabla z\|_{L^2(\Omega)}^2 \\ &\quad + \frac{(1 - \gamma^3) C_\delta}{\gamma} \sup_{t \in [0, T]} \|g^*(\rho_h)\|_{L^\infty(\Omega)} \|\dot{e}_1\|_{L^{2+\delta}(\Omega)}^{2+\delta}. \end{aligned}$$

By Lemma 3.1(c) the residual terms are bounded, so that

$$\frac{1}{2} \frac{d}{dt} \|\nabla z\|_{L^2(\Omega)}^2 + \frac{1}{4} \gamma^4 \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 \leq \eta^2 + (\Lambda_o + \Lambda^*) \|\nabla z\|_{L^2(\Omega)}^2 + \frac{\eta_{g^*} C_\delta}{\gamma} \|\dot{e}_1\|_{L^{2+\delta}(\Omega)}^{2+\delta}. \quad (14)$$

To apply a continuation argument as in [19], we define the time interval

$$I_\theta := \left\{ t \in [0, T] : \Gamma(t) := \sup_{s \in (0, t)} \|\nabla z(s)\|_{L^2(\Omega)}^2 + \frac{1}{2} \gamma^4 \int_0^t \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 \, ds \leq \theta^2 \right\}. \quad (15)$$

Since  $\|\nabla z_0\|_{L^2(\Omega)} < \theta$ , the interval  $I_\theta$  is nonempty, and because  $\Gamma(t)$  is continuous,  $I_\theta$  is closed. To establish  $I_\theta = [0, T]$  we need to show that  $I_\theta$  is also relatively open in  $[0, T]$ . Let  $t \in I_\theta$ . Then by definition of  $I_\theta$  we have  $\|\nabla z(t)\|_{L^2(\Omega)} \leq \theta$ , as well as

$$\int_0^t \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 \, ds \leq 2\gamma^{-4} \theta^2. \quad (16)$$

Together with Lemma 3.1, this allows us to bound the superquadratic term with an extra power of  $\theta$ , i.e.

$$\begin{aligned} \gamma^{-1} \int_0^t \|\dot{e}_1\|_{L^{2+\delta}(\Omega)}^{2+\delta} ds &\leq \gamma^{-1} C_{S,\delta}^{2-\delta} \theta^\delta \int_0^t \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 ds \\ &\leq 2\gamma^{-5} C_{S,\delta}^{2-\delta} \theta^\delta \theta^2 \leq 2C_{S,\delta}^{2-\delta} \left(\frac{\theta}{\gamma^{5/\delta}}\right)^\delta \theta^2 \leq 2C_{S,\delta}^{2-\delta} \mu_1^{2\delta} \theta^2. \end{aligned}$$

By assumption (13) we have  $\|\nabla z_0\|_{L^2(\Omega)}^2 + 2 \int_0^T \eta^2 ds \leq (\|\nabla z_0\|_{L^2(\Omega)} + \sqrt{2}\|\eta\|_{L^2([0,T])})^2 \leq (\mu_2\theta)^2$ . Upon integrating the estimate (14) in time over  $[0, t]$  we deduce, with these two inequalities,

$$\begin{aligned} \|\nabla z(t)\|_{L^2(\Omega)}^2 + \frac{1}{2}\gamma^4 \int_0^t \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 ds &\leq 2 \max(0, \Lambda_\circ + \Lambda^*) \int_0^t \|\nabla z\|_{L^2(\Omega)}^2 ds \\ &\quad + \mu_2^2 \theta^2 + 4\eta_{g^*} C_\delta C_{S,\delta}^{2-\delta} \mu_1^{2\delta} \theta^2. \end{aligned} \quad (17)$$

We are now in a position to apply Gronwall's lemma in such a way that the resulting exponent  $2 \max(0, \Lambda_\circ + \Lambda^*)T$  is independent of  $\gamma^{-1}$ . This yields

$$\begin{aligned} \|\nabla z(t)\|_{L^2(\Omega)}^2 + \frac{1}{2}\gamma^4 \int_0^t \|\nabla \dot{e}_1\|_{L^2(\Omega)}^2 ds &\leq [(\mu_2^2 + 4\eta_{g^*} C_\delta C_{S,\delta}^{2-\delta} \mu_1^{2\delta})\theta^2] e^{2 \max(0, \Lambda_\circ + \Lambda^*)T} \\ &\leq \left[\frac{3}{16} + \frac{1}{4}\right] \theta^2 < \frac{1}{2} \theta^2. \end{aligned}$$

Hence  $\Gamma(t) < \theta^2$  and  $I_\theta$  is also open. Altogether we have proved  $I_\theta = [0, T]$ .  $\square$

**REMARK 3.5** (a) The theorem above guarantees error bounds that do not depend exponentially on  $\gamma^{-1}$  provided that  $\Lambda^*$  is independent of  $\gamma^{-1}$  or  $T \lesssim \gamma$ . For the latter case we notice that  $\Lambda \lesssim \gamma^{-1}$  (cf. (27) below). Then we choose  $T_1 \lesssim \gamma$  and apply the theorem separately on  $[0, T_1]$  and  $[T_1, T]$ , in such a way that the error estimate on the first interval ensures the condition for the initial error on the second interval.

(b) From (13) we see that the minimal polynomial degree in which the error estimate depends on  $\gamma^{-1}$  is 5, if  $d = 2$  and  $\delta = 1$ . If  $d = 3$ , due to the requirement  $\delta \leq 4/5$  the minimal polynomial degree is limited by  $5/\delta > 6$ . The estimate in [15] yields the same polynomial orders.

#### 4. Estimates for the eigenvalue approximation

In this section we derive a version of estimate (9) following ideas for the a posteriori error estimation of eigenvalue problems in [21]. The principal eigenvalue  $-\lambda$  from (10) is well defined since, possibly after a constant shift, the numerator on the right-hand side is a strictly convex functional. Hence there is a minimizing  $q \in \dot{H}^1(\Omega) \setminus \{0\}$  with

$$\lambda(v, \Delta_N^{-1}q) = \gamma(\nabla v, \nabla q) + \gamma^{-1}(v, f'(\hat{\rho}_h)q) + \gamma^{-1}(v, Aq) \quad \text{for all } v \in \dot{H}^1(\Omega). \quad (18)$$

In the discrete eigenvalue problem, we are looking for the smallest number  $-\Lambda$  and a function  $q_h \in \dot{S} \setminus \{0\}$  such that

$$\Lambda(v_h, \Delta_{N_h}^{-1}q_h) = \gamma(\nabla v_h, \nabla q_h) + \gamma^{-1}(v_h, f'(\hat{\rho}_h)q_h) + \gamma^{-1}(v_h, A_h q_h) \quad \text{for all } v_h \in \dot{S}. \quad (19)$$

Here,  $\Delta_N^{-1}$  and  $A$  are replaced by approximations  $\Delta_{Nh}^{-1}$  and  $A_h$ , respectively. If  $A$  is related to linear elasticity, an estimator  $\eta_A$  for the residual  $(A_h - A)q_h$  is given in Lemma 5.3 below. Keeping the definition  $y := -\Delta_N^{-1}q_h$  in mind, the proof of the following lemma can be directly deduced from standard a posteriori estimates for the Laplace equation.

LEMMA 4.1 ([23]) Let  $C_{Cl}$  be the constant related to estimates for the Clément interpolation operator and define the residual estimator

$$\eta_{Nh} := C_{Cl} \|h_{\mathcal{T}}(\Delta_{\mathcal{T}}\Delta_{Nh}^{-1}q_h - q_h)\|_{L^2(\Omega)} + C_{Cl} \|h_{\mathcal{E}}^{1/2} \llbracket \partial_{\bar{n}} \Delta_{Nh}^{-1}q_h \rrbracket \|_{L^2(\cup \mathcal{E})}.$$

Then

$$\|\nabla(\Delta_N^{-1}q_h - \Delta_{Nh}^{-1}q_h)\|_{L^2(\Omega)} \leq \eta_{Nh} \quad \text{for all } q_h \in \mathring{S}.$$

In the same way, we also get a lower bound of the form  $\eta_{Nh} \leq C \|\nabla(\Delta_N^{-1}q_h - \Delta_{Nh}^{-1}q_h)\|_{L^2(\Omega)}$ , where no oscillation terms appears because  $q_h \in \mathring{S}$ . Hence, with the standard a priori estimates for the Laplace equation, we can always guarantee sufficient smallness of  $\eta_{Nh}$  for small mesh width  $h$ .

DEFINITION 4.2 Given  $(\Lambda, q_h) \in \mathbb{R} \times \mathring{S} \setminus \{0\}$  satisfying (19), the *residual*  $R_\Lambda$  is defined to be such that for all  $v \in \mathring{H}^1(\Omega)$ ,

$$\langle v, R_\Lambda \rangle := \Lambda(v, \Delta_N^{-1}q_h) - \gamma(\nabla v, \nabla q_h) - \gamma^{-1}(v, f'(\widehat{\rho}_h)q_h) - \gamma^{-1}(v, Aq_h).$$

LEMMA 4.3 Let  $(\Lambda, q_h) \in \mathbb{R} \times \mathring{S} \setminus \{0\}$  be a solution of (19) and define the residual estimator

$$\eta_{q_h} := C_{Cl} \|h_{\mathcal{T}}(\gamma \Delta_{\mathcal{T}}q_h + \gamma^{-1}(f'(\widehat{\rho}_h)q_h + A_hq_h) - \Lambda \Delta_{Nh}^{-1}q_h)\|_{L^2(\Omega)} + C_{Cl} \|h_{\mathcal{E}}^{1/2} \gamma \llbracket \partial_{\bar{n}} q_h \rrbracket \|_{L^2(\cup \mathcal{E})},$$

with the constant  $C_{Cl}$  related to estimates for the Clément interpolation operator. If there is an estimator  $\eta_A$  such that  $\langle v, (A_h - A)q_h \rangle \leq \eta_A \|\nabla v\|$  for all  $v \in \mathring{H}^1(\Omega)$ , then

$$|\langle v, R_\Lambda \rangle| \leq (\eta_{q_h} + \gamma^{-1}\eta_A) \|\nabla v\|_{L^2(\Omega)} + \Lambda \eta_{Nh} \|\nabla \Delta_N^{-1}v\|_{L^2(\Omega)} \quad \text{for all } v \in \mathring{H}^1(\Omega). \quad (20)$$

*Proof.* For all  $v \in \mathring{H}^1(\Omega)$  and  $v_h \in \mathring{S}$  we subtract (19) from the definition of the residual  $R_\Lambda$  to find that

$$\begin{aligned} \langle v, R_\Lambda \rangle &= \gamma(\nabla(v_h - v), \nabla q_h) + \gamma^{-1}(v_h - v, f'(\widehat{\rho}_h)q_h + A_hq_h) - \Lambda(v_h - v, \Delta_{Nh}^{-1}q_h) \\ &\quad - \Lambda(v, \Delta_{Nh}^{-1}q_h - \Delta_N^{-1}q_h) + \gamma^{-1}(v, (A_h - A)q_h). \end{aligned}$$

An elementwise integration by parts and Hölder's inequality imply

$$\begin{aligned} |\langle v, R_\Lambda \rangle| &\leq \sum_{K \in \mathcal{T}} \|h_{\mathcal{T}}^{-1}(v - v_h)\|_{L^2(K)} \|h_{\mathcal{T}}(\gamma \Delta q_h - \gamma^{-1}(f'(\widehat{\rho}_h)q_h + A_hq_h) + \Lambda \Delta_{Nh}^{-1}q_h)\|_{L^2(K)} \\ &\quad + \sum_{E \in \mathcal{E}} \|h_{\mathcal{E}}^{-1/2}(v - v_h)\|_{L^2(E)} \|h_{\mathcal{E}}^{1/2} \gamma \llbracket \partial_{\bar{n}} q_h \rrbracket \|_{L^2(E)} \\ &\quad + \Lambda |\langle v, \Delta_N^{-1}q_h - \Delta_{Nh}^{-1}q_h \rangle| + \gamma^{-1} |\langle v, (A_h - A)q_h \rangle|. \end{aligned}$$

Choosing  $v_h = \Pi_h v$  to be the Clément interpolant, from standard estimates and Lemma 4.1 we deduce (20).  $\square$

Let  $P_\lambda$  denote the  $L^2$  projection onto the eigenspace related to the eigenvalue  $-\lambda$ , i.e. the space of all  $q \in \dot{H}^1(\Omega)$  satisfying (18). Choosing  $v = q_h$  and  $q = P_\lambda q_h$  in (18) leads to

$$0 = -\lambda(q_h, \Delta_N^{-1} P_\lambda q_h) + \gamma(\nabla q_h, \nabla P_\lambda q_h) + \gamma^{-1}(q_h, f'(\widehat{\rho}_h) P_\lambda q_h) + \gamma^{-1}(q_h, A P_\lambda q_h).$$

We add this equation to the residual in Definition 4.2, where we choose  $v = P_\lambda q_h$ . Because  $\Delta_N^{-1}$  is selfadjoint, i.e.  $(P_\lambda q_h, \Delta_N^{-1} q_h) = -(\nabla \Delta_N^{-1} P_\lambda q_h, \nabla \Delta_N^{-1} q_h) = (\Delta_N^{-1} P_\lambda q_h, q_h)$ , and  $A$  is assumed to be selfadjoint, we get a representation of the error in the eigenvalue approximation,

$$\Lambda - \lambda = \frac{\langle P_\lambda q_h, R_\Lambda \rangle}{(P_\lambda q_h, \Delta_N^{-1} q_h)}, \quad (21)$$

provided that the denominator does not vanish. Thus the numerical approximation space has to be large enough to resolve the eigenvectors related to the principal eigenvalue, i.e.  $P_\lambda \dot{S} \neq \{0\}$ .

**THEOREM 4.4** Let  $-\lambda$  be the smallest number for which there exists a nontrivial  $q$  satisfying (18), and  $P_\lambda$  the  $L^2$  projection onto the eigenspace related to  $-\lambda$ . Let  $(\Lambda, q_h) \in (\mathbb{R}, \dot{S})$  solve (19) with  $\|\nabla \Delta_{Nh}^{-1} q_h\|_{L^2(\Omega)} = 1$ . Assume  $\dot{S}$  is sufficiently large, such that

$$\|\nabla(\Delta_N^{-1} q_h - \Delta_{Nh}^{-1} q_h)\|_{L^2(\Omega)} \leq \eta_{Nh} \leq 1/4, \quad (22)$$

and moreover

$$\|\nabla \Delta_N^{-1}(q_h - P_\lambda q_h)\|_{L^2(\Omega)}^2 \leq 1/4. \quad (23)$$

Then we have the following computable a posteriori error estimate for the eigenvalue:

$$\lambda - \Lambda \leq \eta_\Lambda := 8\gamma^{-1/2}(\eta_{q_h} + \eta_A)(\|f'(\widehat{\rho}_h)\|_\infty + \alpha)\|q_h\|^2 + 4 \max(0, -\Lambda)^{1/2} + 16\Lambda\eta_{Nh}. \quad (24)$$

*Proof.* For the denominator in (21), the assumed bound (23) implies

$$\begin{aligned} -2(P_\lambda q_h, \Delta_N^{-1} q_h) &= 2(\nabla \Delta_N^{-1} P_\lambda q_h, \nabla \Delta_N^{-1} q_h) \\ &= \|\nabla \Delta_N^{-1} P_\lambda q_h\|_{L^2(\Omega)}^2 + \|\nabla \Delta_N^{-1} q_h\|_{L^2(\Omega)}^2 - \|\nabla \Delta_N^{-1}(P_\lambda q_h - q_h)\|_{L^2(\Omega)}^2 \\ &\geq 0 + \|\nabla(\Delta_N^{-1} q_h - \Delta_{Nh}^{-1} q_h) + \nabla \Delta_{Nh}^{-1} q_h\|_{L^2(\Omega)}^2 - 1/4, \\ &\geq (\|\nabla(\Delta_N^{-1} q_h - \Delta_{Nh}^{-1} q_h)\|_{L^2(\Omega)} - \|\nabla \Delta_{Nh}^{-1} q_h\|_{L^2(\Omega)})^2 - 1/4 \\ &= \|\nabla(\Delta_N^{-1} q_h - \Delta_{Nh}^{-1} q_h)\|_{L^2(\Omega)}^2 + \|\nabla \Delta_{Nh}^{-1} q_h\|_{L^2(\Omega)}^2 \\ &\quad - 2\|\nabla(\Delta_N^{-1} q_h - \Delta_{Nh}^{-1} q_h)\|_{L^2(\Omega)} \|\nabla \Delta_{Nh}^{-1} q_h\|_{L^2(\Omega)} - 1/4 \\ &\geq 0 + 1 - 2 \cdot \frac{1}{4} \cdot 1 - \frac{1}{4} = \frac{1}{4}. \end{aligned}$$

Hence  $\Lambda - \lambda \leq 8|\langle P_\lambda q_h, R_\Lambda \rangle|$ . Now we apply Lemma 4.3, where we choose  $v = P_\lambda q_h$  to conclude

$$\Lambda - \lambda \leq 8(\eta_{q_h} + \gamma^{-1}\eta_A)\|\nabla P_\lambda q_h\|_{L^2(\Omega)} + 8\Lambda\eta_{Nh}\|\nabla \Delta_N^{-1} P_\lambda q_h\|_{L^2(\Omega)}.$$

With (22), (23) and Lemma 4.1 we find that

$$\begin{aligned} \|\nabla \Delta_N^{-1} P_\lambda q_h\|_{L^2(\Omega)} &\leq \|\nabla \Delta_N^{-1}(P_\lambda q_h - q_h)\|_{L^2(\Omega)} + \|\nabla(\Delta_N^{-1} q_h - \Delta_{Nh}^{-1} q_h)\|_{L^2(\Omega)} \\ &\quad + \|\nabla \Delta_{Nh}^{-1} q_h\|_{L^2(\Omega)} \\ &\leq 1/2 + \eta_{Nh} + 1 \leq 2. \end{aligned} \quad (25)$$

To bound  $\|\nabla P_\lambda q_h\|_{L^2(\Omega)}$  we choose  $v = q = P_\lambda q_h$  in (18). This yields

$$\gamma \|\nabla P_\lambda q_h\|_{L^2(\Omega)}^2 = -\lambda \|\nabla \Delta_N^{-1} P_\lambda q_h\|_{L^2(\Omega)}^2 - \gamma^{-1} (P_\lambda q_h, f'(\widehat{\rho}_h) P_\lambda q_h + A P_\lambda q_h).$$

If  $-\lambda \leq -\Lambda$  we use  $\|P_\lambda q_h\|_{L^2(\Omega)} \leq \|q_h\|_{L^2(\Omega)}$  and (25) to obtain

$$\gamma \|\nabla P_\lambda q_h\|_{L^2(\Omega)}^2 \leq 4 \max(0, -\Lambda) + \gamma^{-1} (\|f'(\widehat{\rho}_h)\|_\infty + \alpha) \|q_h\|^2$$

and deduce (24). Otherwise, if  $\lambda \leq \Lambda$ , nothing remains to be shown since the right-hand side of (24) is nonnegative.  $\square$

The *saturation assumption* (23) in Theorem 4.4 is quite commonly used to derive error estimates for eigenvalue approximation, but it is not clear how it can be verified in practice. To close this theoretical gap we present an explicit a priori estimate that only requires that the Laplace operator subject to homogeneous Neumann boundary conditions is  $H^2$  regular on  $\Omega$ , i.e. there is a constant  $C_{H^2} > 0$  such that if  $q, y \in \dot{H}^1(\Omega)$  satisfy  $-\Delta y = q$  in  $\Omega$  and  $\partial_n y = 0$  on  $\partial\Omega$ , then  $\|y\|_{H^2(\Omega)} \leq C_{H^2} \|q\|_{L^2(\Omega)}$ . Instead of the difference  $q_h - P_\lambda q$  in the a posteriori estimate, we now consider  $q - \mathcal{I}_T q$ , where  $\mathcal{I}_T$  denotes the nodal interpolation operator. Then we have the interpolation estimate

$$\|q - \mathcal{I}_T q\|_{L^2(\Omega)} + h \|\nabla(q - \mathcal{I}_T q)\|_{L^2(\Omega)} \leq C_{\mathcal{I}_T} h^2 \|D^2 q\|_{L^2(\Omega)}. \quad (26)$$

Analogous to the above, we need a suitable a priori estimate for  $(A_h - A)q_h$ . This can be obtained by standard methods, if  $A$  is the operator related to linear elasticity (cf. Lemma 5.3). From standard a priori estimates for the Laplace equation we get a constant  $C_{Nh} > 0$  such that  $\|\nabla(\Delta_N^{-1} q_h - \Delta_{Nh}^{-1} q_h)\|_{L^2(\Omega)} \leq C_{Nh} h \|q_h\|_{L^2(\Omega)}$  for all  $q_h \in \dot{S}$ . The optimal constant  $C_{-1,1}$  such that  $\|\nabla \Delta_N^{-1} v\|_{L^2(\Omega)} \leq C_{-1,1}^{-1} \|\nabla v\|_{L^2(\Omega)}$  for all  $v \in \dot{H}^1(\Omega)$  is needed to get an a priori estimate for the principal eigenvalue  $-\lambda$ . We note that  $C_{-1,1}^{-1} \leq C_P^2$ .

**THEOREM 4.5** Assume that the Laplace operator subject to homogeneous Neumann boundary conditions is  $H^2$  regular on  $\Omega$  and suppose there is a constant  $C_{Ah} > 0$  such that  $(q_h, (A_h - A)q_h) \leq C_{Ah} h \|\nabla q_h\|_{L^2(\Omega)}^2$  for all  $q_h \in \dot{S}$ . Set

$$\begin{aligned} \varepsilon_0 &:= \gamma^2 + C_A + \max(C_P, C_P^2) \|f'(\widehat{\rho}_h)\|_{L^\infty(\Omega)}, \\ \varepsilon_1 &:= (2\gamma^2 \varepsilon_0 C_{-1,1}^2 + 4(C_f + \alpha)^2)^{1/2}, \\ \varepsilon_2 &:= (\varepsilon_0 C_{H^2} (\gamma^2 C_{-1,1}^2 C_P + \varepsilon_1) C_{\mathcal{I}_T} C_P / 2)^{1/2}. \end{aligned}$$

Assume  $\dot{S}$  is sufficiently large so that  $2\gamma^{-2} \varepsilon_2 h \leq 1/2$  as well as  $\gamma^{-1} (4\varepsilon_1^{1/2} C_P + \gamma) C_{Nh} h \leq C_P$ , and define

$$\begin{aligned} \varepsilon_3 &:= 2\varepsilon_2 + C_{Nh} (4\gamma \varepsilon_1^{1/2} + \gamma^2 / C_P) / 2, \\ \varepsilon_4 &:= \varepsilon_1 \varepsilon_3 + \varepsilon_2 (4\varepsilon_2 + \varepsilon_3 / C_P). \end{aligned}$$

Then the error of the computed eigenvalue is bounded by

$$\lambda - \Lambda \leq \gamma^{-7} [(2\gamma^2 + 6\varepsilon_0)(\varepsilon_1 + \varepsilon_4 / (4\varepsilon_2)) \varepsilon_4 + \gamma^2 C_{Ah} (\varepsilon_1 + \varepsilon_4 / (4\varepsilon_2))^2] h.$$

*Proof.* In a first step we derive a priori bounds for  $q$  and  $|\lambda|$ . From the definition of the principal eigenvalue  $-\lambda$  in (10) and (A3) we get

$$|\lambda| \leq \gamma^{-1}(\gamma^2 + C_P^2 \|f'(\widehat{\rho}_h)\|_{L^\infty(\Omega)} + C_A) \inf_{v \in \dot{H}^1(\Omega) \setminus \{0\}} \frac{\|\nabla v\|_{L^2(\Omega)}^2}{\|\nabla \Delta_N^{-1} v\|_{L^2(\Omega)}^2} \leq \gamma^{-1} \varepsilon_0 C_{-1,1}^2. \quad (27)$$

Let  $(\lambda, q)$  satisfy (18) with  $\|\nabla \Delta_N^{-1} q\|_{L^2(\Omega)} = 1$ . We choose  $v = q$  in (18) to infer with (A1), (A3) and (5) that

$$\begin{aligned} \gamma \|\nabla q\|_{L^2(\Omega)}^2 &= -\lambda \|\nabla \Delta_N^{-1} q\|_{L^2(\Omega)}^2 - \gamma^{-1} (q, f'(\widehat{\rho}_h)q) - \gamma^{-1} (q, Aq) \\ &\leq |\lambda| + \gamma^{-1} (C_f + \alpha) \|q\|_{L^2(\Omega)}^2 \\ &\leq |\lambda| + 2\gamma^{-3} (C_f + \alpha)^2 \|\nabla \Delta_N^{-1} q\|_{L^2(\Omega)}^2 + \frac{\gamma}{2} \|\nabla q\|_{L^2(\Omega)}^2. \end{aligned}$$

From (27) we deduce the estimates

$$\|\nabla q\|_{L^2(\Omega)}^2 \leq 2\gamma^{-2} \varepsilon_0 C_{-1,1}^2 + 4\gamma^{-4} (C_f + \alpha)^2 = \gamma^{-4} \varepsilon_1^2, \quad (28)$$

$$\|q\|_{L^2(\Omega)}^2 = -(\nabla \Delta_N^{-1} q, \nabla q) \leq \|\nabla \Delta_N^{-1} q\|_{L^2(\Omega)} \|\nabla q\|_{L^2(\Omega)} \leq \gamma^{-2} \varepsilon_1. \quad (29)$$

We note  $\|\Delta_N^{-1} q\|_{L^2(\Omega)} \leq C_P \|\nabla \Delta_N^{-1} q\|_{L^2(\Omega)} = C_P$  to conclude from the strong form  $-\lambda \Delta_N^{-1} q = \gamma \Delta q - \gamma^{-1} (f'(\widehat{\rho}_h)q - Aq)$  and the assumed  $H^2$  regularity that

$$\begin{aligned} \|D^2 q\|_{L^2(\Omega)} &\leq \gamma^{-1} C_{H^2} [|\lambda| \|\Delta_N^{-1} q\|_{L^2(\Omega)} + \gamma^{-1} \|f'(\widehat{\rho}_h)q + Aq\|_{L^2(\Omega)}] \\ &\leq \gamma^{-2} C_{H^2} [\gamma |\lambda| C_P + (C_P \|f'(\widehat{\rho}_h)\|_{L^\infty(\Omega)} + C_A) \|\nabla q\|_{L^2(\Omega)}] \\ &\leq \gamma^{-2} C_{H^2} (\varepsilon_0 C_{-1,1}^2 C_P + \gamma^{-2} \varepsilon_0 \varepsilon_1) = \gamma^{-4} 2\varepsilon_2^2 / (C_{\mathcal{I}_T} C_P). \end{aligned}$$

The second step consists in the construction of a discrete approximate eigenfunction. We define the mean value corrected nodal interpolant  $\widetilde{\mathcal{I}}_T q := \mathcal{I}_T q - |\Omega|^{-1} (1, \mathcal{I}_T q) \in \dot{\mathcal{S}}$ . Then the triangle inequality and  $q \in \dot{H}^1(\Omega)$  lead to

$$\|q - \widetilde{\mathcal{I}}_T q\|_{L^2(\Omega)} \leq 2C_{\mathcal{I}_T} h^2 \|D^2 q\|_{L^2(\Omega)} \leq 2h^2 \gamma^{-4} 2\varepsilon_2^2 / C_P.$$

Since for any  $v \in \dot{H}^1(\Omega)$  we have  $\|\nabla \Delta_N^{-1} v\|_{L^2(\Omega)} \leq C_P \|v\|_{L^2(\Omega)}$ , the assumptions on the mesh width  $h$  yield

$$\|\nabla \Delta_N^{-1} (q - \widetilde{\mathcal{I}}_T q)\|_{L^2(\Omega)} \leq (2\gamma^{-2} \varepsilon_2 h)^2 \leq 1/4. \quad (30)$$

On the other hand, from the a priori estimate for  $\Delta_N^{-1}$  we get

$$\begin{aligned} \|\nabla (\Delta_N^{-1} \widetilde{\mathcal{I}}_T q - \Delta_{Nh}^{-1} \widetilde{\mathcal{I}}_T q)\|_{L^2(\Omega)} &\leq C_{Nh} h \|\widetilde{\mathcal{I}}_T q\|_{L^2(\Omega)} \leq C_{Nh} h (\|q\|_{L^2(\Omega)} + \|q - \widetilde{\mathcal{I}}_T q\|_{L^2(\Omega)}) \\ &\leq C_{Nh} h (\gamma^{-1} \varepsilon_1^{1/2} + 1/(4C_P)) \leq 1/4. \end{aligned}$$

Combining both estimates, from the triangle inequality we get  $1/2 \leq \|\nabla \Delta_{Nh}^{-1} \widetilde{\mathcal{I}}_T q\|_{L^2(\Omega)} \leq 3/2$  so

that we can define  $\tilde{q}_h := \tilde{\mathcal{I}}_{\mathcal{T}}q / \|\nabla \Delta_{N_h}^{-1} \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)}$  and deduce

$$\begin{aligned}
\|\nabla(\tilde{q}_h - q)\|_{L^2(\Omega)} &\leq \frac{|1 - \|\nabla \Delta_{N_h}^{-1} \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)}|}{\|\nabla \Delta_{N_h}^{-1} \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)}} \|\nabla \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)} + \|\nabla(\tilde{\mathcal{I}}_{\mathcal{T}}q - q)\|_{L^2(\Omega)} \\
&\leq 2|\|\nabla \Delta_N^{-1}q\|_{L^2(\Omega)} - \|\nabla \Delta_{N_h}^{-1} \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)}| \|\nabla \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)} \\
&\quad + \|\nabla(\tilde{\mathcal{I}}_{\mathcal{T}}q - q)\|_{L^2(\Omega)} \\
&\leq \|\nabla(\tilde{\mathcal{I}}_{\mathcal{T}}q - q)\|_{L^2(\Omega)} + 2\|\nabla \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)} \\
&\quad \times (\|\nabla \Delta_N^{-1}(q - \tilde{\mathcal{I}}_{\mathcal{T}}q)\|_{L^2(\Omega)} + \|\nabla(\Delta_N^{-1} \tilde{\mathcal{I}}_{\mathcal{T}}q - \Delta_{N_h}^{-1} \tilde{\mathcal{I}}_{\mathcal{T}}q)\|_{L^2(\Omega)}) \\
&\leq 2((2\gamma^{-2}\varepsilon_2h)^2 + C_{N_h}h(\gamma^{-1}\varepsilon_1^{1/2} + 1/(4C_P)))\|\nabla \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)} \\
&\quad + \|\nabla(\tilde{\mathcal{I}}_{\mathcal{T}}q - q)\|_{L^2(\Omega)} \\
&=: \gamma^{-2}\varepsilon_3h\|\nabla \tilde{\mathcal{I}}_{\mathcal{T}}q\|_{L^2(\Omega)} + \|\nabla(\tilde{\mathcal{I}}_{\mathcal{T}}q - q)\|_{L^2(\Omega)},
\end{aligned}$$

where we applied (30) and the assumption on the mesh size in order to reduce the polynomial dependence on  $\gamma^{-1}$ . From (28) and (26), we conclude

$$\begin{aligned}
\|\nabla(\tilde{q}_h - q)\|_{L^2(\Omega)} &\leq \gamma^{-2}\varepsilon_3h\|\nabla q\|_{L^2(\Omega)} + (1 + \gamma^{-2}\varepsilon_3h)\|\nabla(\tilde{\mathcal{I}}_{\mathcal{T}}q - q)\|_{L^2(\Omega)} \\
&\leq \gamma^{-4}\varepsilon_1\varepsilon_3h + (1 + \varepsilon_3/(4\varepsilon_2))4\gamma^{-4}(\varepsilon_2^2/C_P)h =: \gamma^{-4}\varepsilon_4h.
\end{aligned}$$

In the third step, we use the definitions of  $\lambda$  and  $\Lambda$  to estimate their difference. Setting  $v_h = \tilde{q}_h$  in (19) and using the minimality of  $-\Lambda$  we get

$$-\Lambda \leq \gamma\|\nabla \tilde{q}_h\|_{L^2(\Omega)}^2 + (\tilde{q}_h, f'(\hat{\rho}_h)\tilde{q}_h) + (\tilde{q}_h, A\tilde{q}_h) + \gamma^{-1}(\tilde{q}_h, (A_h - A)\tilde{q}_h). \quad (31)$$

Upon adding (18), where we choose  $v = q$ , we obtain

$$\begin{aligned}
\lambda - \Lambda &\leq \gamma(\|\nabla \tilde{q}_h\|_{L^2(\Omega)}^2 - \|\nabla q\|_{L^2(\Omega)}^2) + \gamma^{-1}(\tilde{q}_h, (A_h - A)\tilde{q}_h) \\
&\quad + \gamma^{-1}((\tilde{q}_h, f'(\hat{\rho}_h)\tilde{q}_h) - (q, f'(\hat{\rho}_h)q) + (\tilde{q}_h, A\tilde{q}_h) - (q, Aq)).
\end{aligned}$$

In order to apply the binomial formula  $a^2 - b^2 \leq 2a(a - b)$ , we define the shifted, positive definite operator  $B := (f'(\hat{\rho}_h) + \varepsilon_0 C_P^{-2})\text{Id} + A$ . This yields

$$\begin{aligned}
\lambda - \Lambda &\leq \gamma(\|\nabla \tilde{q}_h\|_{L^2(\Omega)}^2 - \|\nabla q\|_{L^2(\Omega)}^2) + \gamma^{-1}((\tilde{q}_h, B\tilde{q}_h) - (q, Bq)) \\
&\quad - \gamma^{-1}\varepsilon_0 C_P^{-2}(\|\tilde{q}_h\|_{L^2(\Omega)}^2 - \|q\|_{L^2(\Omega)}^2) + \gamma^{-1}(\tilde{q}_h, (A_h - A)\tilde{q}_h).
\end{aligned}$$

For the term containing  $A_h - A$ , we use the given a priori estimate from the assumptions to estimate

$$\begin{aligned}
\lambda - \Lambda &\leq 2\gamma(\|\nabla \tilde{q}_h\|_{L^2(\Omega)}, \nabla(\tilde{q}_h - q)) + 2\gamma^{-1}(B\tilde{q}_h, \tilde{q}_h - q) \\
&\quad + 2\gamma^{-1}\varepsilon_0 C_P^{-2}(q, \tilde{q}_h - q) + \gamma^{-1}C_{A_h}h\|\nabla \tilde{q}_h\|_{L^2(\Omega)}^2 \\
&\leq 2\gamma\|\nabla \tilde{q}_h\|_{L^2(\Omega)}\|\nabla(\tilde{q}_h - q)\|_{L^2(\Omega)} + 2\gamma^{-1}2\varepsilon_0\|\nabla \tilde{q}_h\|_{L^2(\Omega)}\|\nabla(\tilde{q}_h - q)\|_{L^2(\Omega)} \\
&\quad + 2\gamma^{-1}\varepsilon_0 C_P^{-2}\|q\|_{L^2(\Omega)}\|\tilde{q}_h - q\|_{L^2(\Omega)} + \gamma^{-1}C_{A_h}h\|\nabla \tilde{q}_h\|_{L^2(\Omega)}^2 \\
&\leq 2[\gamma\|\nabla \tilde{q}_h\|_{L^2(\Omega)} + 2\gamma^{-1}\varepsilon_0\|\nabla \tilde{q}_h\|_{L^2(\Omega)} + \gamma^{-1}\varepsilon_0\|\nabla q\|_{L^2(\Omega)}]\gamma^{-4}\varepsilon_4h \\
&\quad + \gamma^{-1}C_{A_h}h\|\nabla \tilde{q}_h\|_{L^2(\Omega)}^2.
\end{aligned}$$

We finally note  $\|\nabla \tilde{q}_h\|_{L^2(\Omega)} \leq \gamma^{-2}(\varepsilon_1 + \varepsilon_4/(4\varepsilon_2))$  to deduce the desired estimate.  $\square$



## 5. Application to Cahn–Larché equations

The last missing parts to gain robust error control are the residual estimators. These depend on the concrete application with a particular linear operator  $A$  and on the implemented numerical method. In this section we present a practical method for a semi-implicit treatment of the general Cahn–Larché system and derive computable residual estimators in the case of a homogeneous elasticity tensor.

### 5.1 Homogeneous elasticity

We assume a small displacement field  $\vec{u}$  so that the strain is approximated by the symmetric gradient  $\mathcal{E}(\vec{u}) := \frac{1}{2}(\nabla\vec{u} + \nabla\vec{u}^T)$ . By  $\bar{\mathcal{E}}(\rho) := \kappa\mathbb{I}\rho$  we denote the *stress free strain*, where  $\mathbb{I}$  is the unit tensor and  $\kappa > 0$  is called misfit. For symmetric matrices  $\mathcal{A}, \mathcal{B}$  we define  $\mathcal{A} : \mathcal{B} := \text{trace}(\mathcal{A}^T \mathcal{B}) = \sum_{i,j} \mathcal{A}_{ij} \mathcal{B}_{ij}$  and the scalar product  $(\mathcal{A}, \mathcal{B})_{\mathcal{C}} := \int_{\Omega} \mathcal{A} : \mathcal{C} \mathcal{B} \, dx$ . The *elasticity tensor*  $\mathcal{C}$ , which relates the strain to the stress, may in general depend on  $\rho$ . It is called *homogeneous* if  $\mathcal{C}$  is independent of  $\rho$ . The elasticity tensor  $\mathcal{C} \in \mathbb{R}^{d \times d \times d \times d}$  is assumed to be a symmetric positive definite fourth order tensor, i.e.

- $\mathcal{C}_{ijmn} = \mathcal{C}_{ijnm} = \mathcal{C}_{jimn}$  and  $\mathcal{C}_{ijmn} = \mathcal{C}_{mnij}$  for all  $1 \leq i, j, m, n \leq d$ ,
- there is  $c^* > 0$  such that  $\mathcal{A} : \mathcal{C} \mathcal{A} \geq c^* |\mathcal{A}|^2$  for all symmetric matrices  $\mathcal{A} \in \mathbb{R}^{d \times d}$ .

Following [16], the elastic energy can be modeled by

$$\mathcal{W}(\rho, \mathcal{E}(\vec{u})) := \frac{1}{2} (\mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho)) : \mathcal{C} (\mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho)).$$

Since the time scale of mechanical relaxation is much smaller than the scale at which diffusion takes place, we may assume an equilibrium state governed by

$$0 = \text{div}(\mathcal{C}[\mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho)]) \quad \text{in } \Omega.$$

When the natural boundary conditions  $\vec{n} \cdot (\mathcal{C}[\mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho)]) = 0$  are prescribed on all of the domain boundary  $\partial\Omega$ , then  $\vec{u}$  is not uniquely determined by this equation. From the solution space we have to exclude the kernel of  $\mathcal{E}$ , which contains all rigid body motions. This does not affect the elastic energy because  $\mathcal{W}(\rho, \mathcal{E}(\vec{u}))$  depends only on  $\mathcal{E}(\vec{u})$ , not on  $\vec{u}$  itself. The set of all infinitesimal, linearized rigid body motions is given by

$$\text{RM}(\Omega) = \begin{cases} \{(x, y) \mapsto (a, b) + c(y, -x) : a, b, c \in \mathbb{R}, (x, y) \in \Omega \subset \mathbb{R}^2\} & \text{for } d = 2, \\ \{x \mapsto a + c \times x : a, b \in \mathbb{R}^3, x \in \Omega \subset \mathbb{R}^3\} & \text{for } d = 3. \end{cases}$$

We denote  $H_{\text{RM}}^1(\Omega) := (H^1(\Omega))^d / \text{RM}(\Omega)$  and define the solution space of the Cahn–Larché system to be

$$X_{\text{CL}} := X_{\text{CH}} \times L^\infty([0, T], H_{\text{RM}}^1(\Omega)).$$

In the weak form we seek the solution of the following problem:

$$(\text{CL}) \begin{cases} \text{Given } \rho(0, x) = \rho_0(x) \in H^1(\Omega), \text{ find } (\rho, w, \vec{u}) \in X_{\text{CL}} \text{ such that for almost all } t \in (0, T), \\ \mathcal{E}(\vec{\xi}), \mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho)_{\mathcal{C}} = 0 & \text{for all } \vec{\xi} \in H_{\text{RM}}^1(\Omega), \\ \langle \varphi, \partial_t \rho \rangle + (\nabla \varphi, \nabla w) = 0 & \text{for all } \varphi \in H^1(\Omega), \\ (\psi, w) - \gamma (\nabla \psi, \nabla \rho) = \gamma^{-1} (\psi, f(\rho) + \mathcal{W}(\rho, \mathcal{E}(\vec{u}))) & \text{for all } \psi \in H^1(\Omega). \end{cases}$$

If the elasticity tensor  $\mathcal{C}$  is homogeneous, the first equation of (CL) defines the linear solution operator  $\mathcal{D}^{-1} : \rho \mapsto \vec{u}$ , and the contribution of the elastic energy to (CL) is

$$W(\rho, \mathcal{E}(\vec{u})) = -\kappa \mathbb{I} : \mathcal{C}(\mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho)). \quad (32)$$

In the next lemma, we show that the mapping  $\rho \mapsto A\rho = W(\rho, \mathcal{E}(\vec{u}))$  satisfies (A3).

LEMMA 5.1 There is a constant  $C_{\mathcal{C}, K}$  such that for given  $\rho \in L^2(\Omega)$  and  $\vec{u}$  as in (CL) we have

$$\|\nabla \vec{u}\|_{L^2(\Omega)} \leq \kappa C_{\mathcal{C}, K} \|\rho\|_{L^2(\Omega)}. \quad (33)$$

Moreover, the linear operator  $A : \rho \mapsto W(\rho, \mathcal{E}(\vec{u})) = -\kappa \mathbb{I} : \mathcal{C}(\mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho))$  is selfadjoint and there are constants  $\alpha, C_A \geq 0$  such that assumption (A3) is satisfied, i.e.

$$\begin{aligned} -(\rho, A\rho) &\leq \alpha \|\rho\|_{L^2(\Omega)}^2 && \text{for all } \rho \in L^2(\Omega), \\ (\psi, A\rho) &\leq C_A \|\nabla \psi\|_{L^2(\Omega)} \|\nabla \rho\|_{L^2(\Omega)} && \text{for all } \rho \in \dot{H}^1(\Omega). \end{aligned}$$

*Proof.* Since we excluded rigid body motions, we can apply Korn's inequality. On the other hand,  $\mathcal{C}$  is positive definite. Together this yields

$$C_K \|\nabla \vec{u}\|_{L^2(\Omega)}^2 \leq (\mathcal{E}(\vec{u}), \mathcal{E}(\vec{u})) \leq \frac{1}{c^*} (\mathcal{E}(\vec{u}), \mathcal{E}(\vec{u}))_{\mathcal{C}}. \quad (34)$$

By choosing  $\vec{\xi} = \vec{u}$  in (CL) and employing Hölder's inequality we infer

$$\begin{aligned} \|\nabla \vec{u}\|_{L^2(\Omega)}^2 &\leq \frac{1}{c^* C_K} (\bar{\mathcal{E}}(\rho), \mathcal{E}(\vec{u}))_{\mathcal{C}} = \frac{\kappa}{c^* C_K} \int_{\Omega} \rho \mathcal{E}(\vec{u}) : \mathcal{C} \mathbb{I} \, dx \\ &\leq \frac{\kappa}{c^* C_K} \|\rho\|_{L^2(\Omega)} \|\mathcal{E}(\vec{u}) : \mathcal{C} \mathbb{I}\|_{L^2(\Omega)} \leq \kappa C_{\mathcal{C}, K} \|\rho\|_{L^2(\Omega)} \|\nabla \vec{u}\|_{L^2(\Omega)} \end{aligned}$$

with some constant  $C_{\mathcal{C}, K} > 0$ . The symmetry of  $\mathcal{C}$  implies that  $A$  is selfadjoint and

$$\begin{aligned} (\rho, A\rho) &= - \int_{\Omega} \kappa \mathbb{I} \rho : \mathcal{C}(\mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho)) \, dx = -(\bar{\mathcal{E}}(\rho), \mathcal{E}(\vec{u}) - \bar{\mathcal{E}}(\rho))_{\mathcal{C}} \\ &= -(\mathcal{E}(\vec{u}), \mathcal{E}(\vec{u}))_{\mathcal{C}} + (\bar{\mathcal{E}}(\rho), \bar{\mathcal{E}}(\rho))_{\mathcal{C}} \geq -c^* C_K \|\nabla \vec{u}\|_{L^2(\Omega)}^2 + \kappa^2 c^* \|\rho\|_{L^2(\Omega)}^2. \end{aligned}$$

Upon setting  $\alpha = \max(0, \kappa^2(C_K C_{\mathcal{C}, K}^2 - |\Omega|))c^*$  we get the lower bound for  $A$ . For the upper bound, we set  $\vec{\xi} = \mathcal{D}^{-1}\psi$  and  $\vec{v} = \mathcal{D}^{-1}\rho$ . An application of Hölder's inequality, (33) and (34) yield

$$\begin{aligned} (\psi, A\rho) &= -(\bar{\mathcal{E}}(\psi), \mathcal{E}(\vec{v}) - \bar{\mathcal{E}}(\rho))_{\mathcal{C}} = -(\mathcal{E}(\vec{\xi}), \mathcal{E}(\vec{v}))_{\mathcal{C}} + (\bar{\mathcal{E}}(\psi), \bar{\mathcal{E}}(\rho))_{\mathcal{C}} \\ &\leq (\mathcal{E}(\vec{\xi}), \mathcal{E}(\vec{\xi}))_{\mathcal{C}}^{1/2} (\mathcal{E}(\vec{v}), \mathcal{E}(\vec{v}))_{\mathcal{C}}^{1/2} + \kappa^2 \mathbb{I} : \mathcal{C} \mathbb{I} \|\psi\|_{L^2} \|\rho\|_{L^2(\Omega)} \\ &\leq C_K^2 \|\nabla \vec{\xi}\|_{L^2(\Omega)} \|\nabla \vec{v}\|_{L^2(\Omega)} + \kappa^2 \mathbb{I} : \mathcal{C} \mathbb{I} C_P^2 \|\nabla \psi\|_{L^2} \|\nabla \rho\|_{L^2(\Omega)} \\ &\leq (C_K^2 \kappa^2 C_{\mathcal{C}, K}^2 + \kappa^2 \mathbb{I} : \mathcal{C} \mathbb{I} C_P^2) \|\nabla \psi\|_{L^2} \|\nabla \rho\|_{L^2(\Omega)}, \end{aligned}$$

which implies the asserted estimate.  $\square$

## 5.2 Numerical scheme and residual estimates

We define  $\mathcal{S}_{\text{RM}} := \mathcal{S}^d \cap H_{\text{RM}}^1(\Omega)$  and propose the following decoupled semi-implicit numerical method to approximate (CL):

$$(\text{CL}_h) \left\{ \begin{array}{l} \text{Given the approximations } \rho_h^{(j-1)} \text{ at time } t_{j-1}, \text{ first compute } \vec{u}_h^{(j)} \in \mathcal{S}_{\text{RM}}^{(j)} \\ \text{such that for all } \vec{\xi}_h \in \mathcal{S}_{\text{RM}}^{(j)}, \\ \quad 0 = (\mathcal{E}(\vec{\xi}_h), \mathcal{E}(\vec{u}_h^{(j)}) - \bar{\mathcal{E}}(\mathcal{I}^{(j)} \rho_h^{(j-1)}))_{\mathcal{C}}; \\ \text{then compute } (\rho_h^{(j)}, w_h^{(j)}) \in \mathcal{S}^{(j)} \times \mathcal{S}^{(j)} \text{ such that for all } \varphi_h, \psi_h \in \mathcal{S}^{(j)}, \\ (\varphi_h, \rho_h^{(j)}) + \tau_j (\nabla \varphi_h, \nabla w_h^{(j)}) = (\varphi_h, \mathcal{I}^{(j)} \rho_h^{(j-1)}), \\ (\psi_h, w_h^{(j)}) - \gamma (\nabla \psi_h, \nabla \rho_h^{(j)}) - \gamma^{-1} (\psi_h, f'(\mathcal{I}^{(j)} \rho_h^{(j-1)}) \rho_h^{(j)}) \\ = \gamma^{-1} (\psi_h, f(\mathcal{I}^{(j)} \rho_h^{(j-1)}) - f'(\mathcal{I}^{(j)} \rho_h^{(j-1)}) \rho_h^{(j-1)}) \\ \quad + \gamma^{-1} (\psi_h, W(\mathcal{I}^{(j)} \rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)}))). \end{array} \right.$$

Each time step is decoupled into an elasticity part and a Cahn–Hilliard part. Existence of a unique solution can be shown by standard methods. Because the elasticity tensor is positive definite, the elasticity problem is elliptic and a unique solution is guaranteed. Owing to the validity of an inf-sup condition, unique solvability of the discrete system is guaranteed if  $\tau \leq C\gamma^2 / \|f'(\mathcal{I}^{(j)} \rho_h^{(j-1)})\|_{L^\infty}^2$ , which is the same condition as for a fully implicit scheme considered in [24]. The scheme  $(\text{CL}_h)$  is equivalent to performing one Newton iteration step within each time step of an implicit Euler discretization. From  $(\text{CL}_h)$  we extract the approximated linear operator  $A_h$  as

$$A_h \mathcal{I}^{(j)} \rho_h^{(j-1)} = W(\mathcal{I}^{(j)} \rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)})).$$

With this choice of  $A_h$  the discrete eigenvalue problem (19) is well defined, and in the lemma below we show that  $A_h$  satisfies the assumptions of Lemma 4.3 and Theorem 4.5.

**DEFINITION 5.2** For  $\rho_h \in L^2(\Omega)$  given, denote by  $\vec{u}_h$  the finite element solution of  $0 = (\mathcal{E}(\vec{\xi}_h), \mathcal{E}(\vec{u}_h) - \bar{\mathcal{E}}(\rho_h))_{\mathcal{C}}$  for all  $\vec{\xi}_h \in \mathcal{S}_{\text{RM}}$ . The *residual*  $R_3$  is defined as

$$\langle \vec{\xi}, R_3 \rangle := (\mathcal{E}(\vec{\xi}), \mathcal{E}(\vec{u}_h) - \bar{\mathcal{E}}(\rho_h))_{\mathcal{C}} \quad \text{for all } \vec{\xi} \in H_{\text{RM}}^1(\Omega). \quad (35)$$

**LEMMA 5.3** Suppose  $\mathcal{C}$  is homogeneous and let  $W(\rho, \mathcal{E}(\vec{u}))$  be given by (32). Let  $A : \rho \mapsto W(\rho, \mathcal{E}(\vec{u}))$  with  $\vec{u}$  defined by the first identity in (CL). Similarly, let  $A_h : \rho_h \mapsto W(\rho_h, \mathcal{E}(\vec{u}_h))$  with  $\vec{u}_h$  as in Definition 5.2. Set

$$\tilde{\eta}_A := \|h_{\mathcal{T}}(\text{div}_{\mathcal{T}}(\mathcal{C}[\mathcal{E}(\vec{u}_h) - \bar{\mathcal{E}}(\rho_h)]))\|_{L^2(\Omega)} + \|h_{\mathcal{E}}^{1/2} \llbracket \vec{n} \cdot (\mathcal{C}\mathcal{E}(\vec{u}_h) - \bar{\mathcal{E}}(\rho_h)) \rrbracket \|_{L^2(\cup \mathcal{E})}.$$

Then, with  $\eta_A := \kappa C_P C_{\mathcal{C}, K} C_{\text{CI}} \tilde{\eta}_A$ , we have

$$\begin{aligned} \langle \vec{\xi}, R_3 \rangle &\leq \tilde{\eta}_A C_{\text{CI}} \|\nabla \vec{\xi}\|_{L^2(\Omega)} && \text{for all } \vec{\xi} \in H_{\text{RM}}^1(\Omega), \\ (\varphi, (A_h - A)\rho_h) &\leq \eta_A \|\nabla \varphi\|_{L^2(\Omega)} && \text{for all } \varphi \in \dot{H}^1(\Omega). \end{aligned} \quad (36)$$

Under the assumption of  $H^2$  regularity of the elasticity equations in the sense that  $\|\vec{u}\|_{H^2(\Omega)} \leq C\|\nabla\rho_h\|_{L^2(\Omega)}$ , there is a constant  $C_{Ah} > 0$  such that we have the a priori estimate

$$|(\rho_h, (A_h - A)\rho_h)| \leq C_{Ah}h\|\nabla\rho_h\|_{L^2(\Omega)}^2.$$

*Proof.* Let  $\vec{\xi} \in H_{\text{RM}}^1(\Omega)$ . By Galerkin orthogonality we can insert the Clément interpolant of  $\vec{\xi}$  in (35). Using elementwise integration by parts and applying standard estimates we get

$$\langle \vec{\xi}, R_3 \rangle \leq \tilde{\eta}_A C_{\text{Cl}} \|\nabla \vec{\xi}\|_{L^2(\Omega)}.$$

Define  $\vec{v} := \mathcal{D}^{-1}\rho_h$ . Then  $A\rho_h = -\kappa\mathbb{I} : \mathcal{C}(\mathcal{E}(\vec{v}) - \bar{\mathcal{E}}(\rho_h))$  and for  $\varphi \in \dot{H}^1(\Omega)$  set  $\vec{\xi} := \mathcal{D}^{-1}\varphi$ . Since  $\vec{u}_h - \vec{v} \in H_{\text{RM}}^1(\Omega)$  is an admissible test function, we have

$$\begin{aligned} (\varphi, (A_h - A)\rho_h) &= (\bar{\mathcal{E}}(\varphi), \mathcal{E}(\vec{u}_h) - \mathcal{E}(\vec{v}))_{\mathcal{C}} \\ &= (\mathcal{E}(\vec{\xi}), \mathcal{E}(\vec{u}_h) - \mathcal{E}(\vec{v}))_{\mathcal{C}} = (\mathcal{E}(\vec{\xi}), \mathcal{E}(\vec{u}_h) - \bar{\mathcal{E}}(\rho_h))_{\mathcal{C}} = \langle \vec{\xi}, R_3 \rangle. \end{aligned}$$

With Lemma 5.1 and Poincaré's inequality we infer (36). For the a priori estimate, we choose  $\varphi = \rho_h$  and apply Hölder's inequality to deduce

$$\begin{aligned} |(\rho_h, (A_h - A)\rho_h)| &= |(\mathcal{E}(\vec{v}), \mathcal{E}(\vec{u}_h) - \mathcal{E}(\vec{v}))_{\mathcal{C}}| \leq (\mathcal{E}(\vec{v}), \mathcal{E}(\vec{v}))_{\mathcal{C}}^{1/2} (\mathcal{E}(\vec{u}_h - \vec{v}), \mathcal{E}(\vec{u}_h - \vec{v}))_{\mathcal{C}}^{1/2} \\ &\leq c^* C_K \|\nabla \vec{v}\|_{L^2(\Omega)} \|D^2 \vec{v}\|_{L^2(\Omega)}, \end{aligned}$$

due to (34) and standard a priori estimates for linear elasticity with a constant  $C > 0$ . We finish the proof by using (33) and the  $H^2$  regularity of  $\vec{v}$ .  $\square$

**LEMMA 5.4** Let  $(\rho_h^{(j)}, w_h^{(j)}, u_h^{(j)})_{j=0,\dots,N}$  be the solution of  $(\text{CL}_h)$  and let  $(\rho_h, w_h, u_h)$  denote the piecewise affine in time interpolation. Set  $r^{(j)} := \max(\|\rho_h^{(j-1)}\|_{L^\infty(\Omega)}, \|\rho_h^{(j)}\|_{L^\infty(\Omega)})$  and  $I^{(j)} := [-r^{(j)}, r^{(j)}]$ . Let  $\eta_A^{(j)}$  be as in Lemma 5.3. Define the residual estimators  $\eta_1^{(j)} := C_{\text{Cl}}\eta_{1h}^{(j)} + \eta_{1t}^{(j)} + \eta_{1c}^{(j)}$ , where

$$\begin{aligned} \eta_{1h}^{(j)} &:= \|h_{\mathcal{T}^{(j)}}(\tau_j^{-1}(\rho_h^{(j)} - \mathcal{I}^{(j)}\rho_h^{(j-1)}) - \Delta w_h^{(j)})\|_{L^2(\Omega)} + \|h_{\mathcal{E}^{(j)}}^{1/2} \llbracket \partial_{\vec{n}} w_h^{(j)} \rrbracket\|_{L^2(\cup \mathcal{E}^{(j)})}, \\ \eta_{1t}^{(j)} &:= C_P \|\nabla w_h^{(j)} - \nabla w_h^{(j-1)}\|_{L^2(\Omega)}, \\ \eta_{1c}^{(j)} &:= \tau_j^{-1} C_P \|\mathcal{I}^{(j)}\rho_h^{(j-1)} - \rho_h^{(j-1)}\|_{L^2(\Omega)}, \\ \eta_{21}^{(j)} &:= C_{\text{Cl}}\eta_{21h}^{(j)} + \eta_{21t}^{(j)} + \gamma^{-1}(\eta_A^{(j-1)} + \eta_A^{(j)}), \quad \text{where} \\ \eta_{21h}^{(j)} &:= \|h_{\mathcal{T}^{(j)}}(w_h^{(j)} + \gamma \Delta_{\mathcal{T}^{(j)}}\rho_h^{(j)} - \gamma^{-1}[f(\mathcal{I}^{(j)}\rho_h^{(j-1)}) + W(\mathcal{I}^{(j)}\rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)})) \\ &\quad + f'(\mathcal{I}^{(j)}\rho_h^{(j-1)})(\rho_h^{(j)} - \mathcal{I}^{(j)}\rho_h^{(j-1)})])\|_{L^2(\Omega)} + \gamma \|h_{\mathcal{E}^{(j)}}^{1/2} \llbracket \partial_{\vec{n}} \rho_h^{(j)} \rrbracket\|_{L^2(\cup \mathcal{E}^{(j)})}, \\ \eta_{21t}^{(j)} &:= \gamma \|\nabla(\rho_h^{(j)} - \rho_h^{(j-1)})\|_{L^2(\Omega)}, \\ \eta_{22}^{(j)} &:= \eta_{2t}^{(j)} + \eta_{2c}^{(j)} + \eta_{2\ell}^{(j)}, \\ \eta_{22t}^{(j)} &:= \|w_h^{(j)} - w_h^{(j-1)}\|_{L^2(\Omega)} + \gamma^{-1} \|f'\|_{L^\infty(I^{(j)})} \|\rho_h^{(j)} - \rho_h^{(j-1)}\|_{L^2(\Omega)} \\ &\quad + \gamma^{-1} \|W(\rho_h^{(j)}, \mathcal{E}(\vec{u}_h^{(j)})) - W(\rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)}))\|_{L^2(\Omega)}, \\ \eta_{22c}^{(j)} &:= \gamma^{-1} \|W(\rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)})) - W(\mathcal{I}^{(j)}\rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)}))\|_{L^2(\Omega)} \\ \eta_{22\ell}^{(j)} &:= \gamma^{-1} \|f(\rho_h^{(j)}) - f(\mathcal{I}^{(j)}\rho_h^{(j-1)}) - f'(\mathcal{I}^{(j)}\rho_h^{(j-1)})(\rho_h^{(j)} - \rho_h^{(j-1)})\|_{L^2(\Omega)}. \end{aligned}$$

Then, for almost all  $s \in (t_{j-1}, t_j)$ ,  $j = 1, \dots, N$ , we have

$$\langle \varphi, R_1(s) \rangle \leq \eta_1^{(j)} \|\nabla \varphi\|_{L^2(\Omega)} \quad \text{for all } \varphi \in \dot{H}^1(\Omega), \quad (37a)$$

$$\langle \psi, R_2(s) \rangle \leq \eta_{21}^{(j)} \|\nabla \psi\|_{L^2(\Omega)} + \eta_{22}^{(j)} \|\psi\|_{L^2(\Omega)} \quad \text{for all } \psi \in \dot{H}^1(\Omega). \quad (37b)$$

*Proof.* We add and subtract terms to the residuals defined in Definition 2.2, so that we split the residuals into  $R_1(s) = R_{1h}^{(j)} + R_{1t}(s) + R_{1c}^{(j)}$  and  $R_2(s) = R_{2h}^{(j)} + R_{2t}(s) + R_{2c}^{(j)} + R_{2\ell}^{(j)} + \gamma^{-1}(A - A_h)\rho_h$ , where the subscripts  $c$ ,  $\ell$ ,  $h$  and  $t$  refer to coarsening, linearization, space discretization, and time discretization residuals, respectively. The space discretization residuals are given by the discrete operator applied to test functions in the whole space  $\dot{H}^1(\Omega)$ ,

$$\langle \varphi, R_{1h}^{(j)} \rangle := \tau_j^{-1}(\varphi, \rho_h^{(j)} - \mathcal{I}^{(j)}\rho_h^{(j-1)}) + (\nabla \varphi, \nabla w_h^{(j)}), \quad (38a)$$

$$\begin{aligned} \langle \psi, R_{2h}^{(j)} \rangle &:= \gamma(\nabla \psi, \nabla \rho_h^{(j)}) + \gamma^{-1}(\psi, f'(\mathcal{I}^{(j)}\rho_h^{(j-1)})(\rho_h^{(j)} - \mathcal{I}^{(j)}\rho_h^{(j-1)})) - (\psi, w_h^{(j)}) \\ &\quad + \gamma^{-1}(\psi, f(\mathcal{I}^{(j)}\rho_h^{(j-1)}) + W(\mathcal{I}^{(j)}\rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)}))). \end{aligned} \quad (38b)$$

The linearization residual is

$$\langle \psi, R_{2\ell}^{(j)} \rangle := \gamma^{-1}(\psi_h, f(\rho_h^{(j)}) - f(\mathcal{I}^{(j)}\rho_h^{(j-1)}) - f'(\mathcal{I}^{(j)}\rho_h^{(j-1)})[\rho_h^{(j)} - \mathcal{I}^{(j)}\rho_h^{(j-1)}]) \quad (39)$$

and the coarsening residuals contain the remaining explicit terms in the scheme, i.e.

$$\langle \varphi, R_{1c}^{(j)} \rangle := \tau_j^{-1}(\varphi, \mathcal{I}^{(j)}\rho_h^{(j-1)} - \rho_h^{(j-1)}), \quad (40a)$$

$$\langle \varphi, R_{2c}^{(j)} \rangle := \gamma^{-1}(\varphi, W(\rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)})) - W(\mathcal{I}^{(j)}\rho_h^{(j-1)}, \mathcal{E}(\vec{u}_h^{(j)}))). \quad (40b)$$

Finally, the time discretization residuals are given by

$$\langle \varphi, R_{1t}(s) \rangle := (\nabla \varphi, \nabla w_h(s) - \nabla w_h^{(j)}), \quad (41a)$$

$$\begin{aligned} \langle \psi, R_{2t}(s) \rangle &:= \gamma(\nabla \psi, \nabla \rho_h(s) - \nabla \rho_h^{(j)}) - (\psi, w_h(s) - w_h^{(j)}) \\ &\quad + \gamma^{-1}(\psi, f(\rho_h(s)) - f(\rho_h^{(j)}) + W(\rho_h(s), \mathcal{E}(\vec{u}_h(s))) - W(\rho_h^{(j)}, \mathcal{E}(\vec{u}_h^{(j)}))). \end{aligned} \quad (41b)$$

With Lemma 5.3 we can directly estimate  $\langle \psi, (A - A_h)\rho_h(s) \rangle \leq (\eta_A^{(j)} + \eta_A^{(j-1)})\|\nabla \psi\|_{L^2(\Omega)}$  because  $\rho_h$  is affine in  $(t_{j-1}, t_j)$ . By Galerkin orthogonality we can insert the Clément interpolant  $\Pi_h \varphi$  into the space discretization residual  $R_{1h}^{(j)}$ . Then, by elementwise integration by parts and standard estimates, we conclude

$$\begin{aligned} \langle \varphi, R_{1h}^{(j)} \rangle &= \sum_{K \in \mathcal{T}^{(j)}} (\varphi - \Pi_h \varphi, \tau_j^{-1}(\rho_h^{(j)} - \mathcal{I}^{(j)}\rho_h^{(j-1)}) - \Delta_{\mathcal{T}^{(j)}} w_h) + \sum_{E \in \mathcal{E}^{(j)}} (\varphi - \Pi_h \varphi, [[\partial_{\vec{n}} w_h]]) \\ &\leq C_{Cl} \eta_{1h} \|\nabla \varphi\|_{L^2(\Omega)}. \end{aligned}$$

Analogously, we infer  $\langle \psi, R_{2h}^{(j)} \rangle \leq C_{Cl} \eta_{21h} \|\psi\|_{L^2(\Omega)}$ . Since  $\rho_h$  and  $w_h$  are affine in  $(t_{j-1}, t_j)$ , by Hölder's and Poincaré's inequality we conclude

$$\langle \varphi, R_{1c} \rangle \leq \tau_j^{-1} \|\mathcal{I}^{(j)}\rho_h^{(j-1)} - \rho_h^{(j-1)}\|_{L^2(\Omega)} \|\varphi\|_{L^2(\Omega)} \leq \eta_{1c} \|\nabla \varphi\|_{L^2(\Omega)}$$

and in analogous way  $\langle \varphi, R_{1t}^{(j)} \rangle \leq \eta_{1t}^{(j)} \|\nabla \varphi\|_{L^2(\Omega)}$ . Further application of Hölder's inequality proves  $\langle \psi, R_{2\ell}^{(j)} \rangle \leq \eta_{22\ell}^{(j)} \|\psi\|_{L^2(\Omega)}$  and  $\langle \psi, R_{2c}^{(j)} \rangle \leq \eta_{22c}^{(j)} \|\psi\|_{L^2(\Omega)}$ . The first term of the time discretization residual  $R_{2t}(s)$  is estimated with  $\eta_{21t}$ . Moreover,  $W$  is linear with respect to  $\mathcal{E}(\vec{u})$  and for each  $x \in \Omega$  there is some  $r_x$  between  $\rho_h(s, x)$  and  $\rho_h^{(j)}(x)$  such that

$$f(\rho_h(s, x)) - f(\rho_h^{(j)}(x)) = f'(r_x)(\rho_h(s, x) - \rho_h^{(j)}(x)) \leq \|f'\|_{L^\infty(I^{(j)})} |\rho_h^{(j)}(x) - \rho_h^{(j-1)}(x)|.$$

Then, by Hölder's inequality, we deduce  $\langle \psi, R_{2t}^{(j)} \rangle \leq \eta_{21t}^{(j)} \|\nabla \psi\|_{L^2(\Omega)} + \eta_{22t}^{(j)} \|\psi\|_{L^2(\Omega)}$ .  $\square$

## 6. Numerical experiments

In this section we illustrate the theoretical results of the previous sections and analyze the asymptotic behavior of the estimators as well as the range of applicability of the a posteriori error estimate. In particular, we are interested in the endpoints of the time interval in which a uniform bound for the principal eigenvalue holds. We analyze two prototypical examples of topological changes. The relevant quantities are the approximate eigenvalue  $-\Lambda(t)$  as well as the Ginzburg–Landau energy  $E(t)$  and the error estimators.

We used the numerical method (CL<sub>h</sub>) to approximate the nonlinear evolution. A rearrangement of the estimators in Lemma 5.4 leads in a natural way to local indicators for grid refinement and coarsening,

$$\eta_h^2 := \frac{1}{4}\eta_{1h}^2 + \frac{2}{\gamma^4} \left( \eta_{21h} + \frac{1}{\gamma}\eta_A \right)^2, \quad \eta_c^2 := \frac{1}{4}\eta_{1c}^2 + \frac{4}{\gamma^2}\eta_{22c}^2.$$

The discrete eigenvalue problem in the  $j$ th time step reads as follows:

$$(EV_h) \left\{ \begin{array}{l} \text{Given } \rho_h^{(j)}, \text{ find } (q_h, z_h, \vec{v}_h, \Lambda) \in \mathcal{S}^{(j)} \times \mathcal{S}^{(j)} \times \mathcal{S}_{\text{RM}}^{(j)} \times \mathbb{R} \text{ such that} \\ \text{for all } \vec{\xi}_h \in \mathcal{S}_{\text{RM}}^{(j)} \text{ and for all } \varphi_h, \psi_h \in \mathring{\mathcal{S}}^{(j)}, \\ (\mathcal{E}(\vec{\xi}_h), \mathcal{E}(\vec{v}_h) - \bar{\mathcal{E}}(q_h))_{\mathcal{C}} = 0, \\ \gamma(\varphi_h, q_h) - \gamma(\nabla \varphi_h, \nabla z_h) = 0, \\ \gamma(\nabla \psi_h, \nabla q_h) + \gamma^{-1}(\psi_h, f'(\rho_h^{(j)})q_h + W(q_h, \mathcal{E}(\vec{v}_h))) = -\Lambda(\psi_h, z_h). \end{array} \right.$$

We solve this problem by a shifted inverse vector iteration. Therefore, we introduce a term  $\varepsilon_0(\psi_h, z_h)$  to each side of the last equation of (EV<sub>h</sub>). We note that the eigenvalue problem cannot be decoupled into an elasticity part and a Cahn–Hilliard part, so the size of the linear system to solve is twice as large compared to each subproblem in (CL<sub>h</sub>). Due to the constraints related to the spaces  $\mathring{\mathcal{S}}^{(j)}$  and  $\mathcal{S}_{\text{RM}}^{(j)}$  this matrix has a large bandwidth, thus it is not well suited for direct solvers.

### 6.1 Sharp interface limit and detection of topological changes

To check the accuracy of the numerical method we compare the numerical solution to a reference solution. We therefore neglect elastic effects and consider the pure Cahn–Hilliard problem, i.e.  $A = 0$ . In the sharp interface limit  $\gamma \rightarrow 0$ , the Ginzburg–Landau energy  $E$  from (1) converges to the interface length in the corresponding Mullins–Sekerka problem.

EXPERIMENT 1 Let  $r_1 = 0.55$  and  $r_2 = 0.2$  and choose

$$\rho_0(x) = \min\left(-\tanh\left(\frac{|x| - r_1}{\gamma\sqrt{2}}\right), \tanh\left(\frac{|x| - r_2}{\gamma\sqrt{2}}\right)\right).$$

Initially the interface consists of two concentric circles. Then the smaller circle shrinks until it vanishes completely and the solution reaches a stable state with only one circular interface (see Figure 2). Given this initial data the Mullins–Sekerka problem can be reduced to a system of ordinary differential equations for the radii  $r_1$  and  $r_2$  (cf. [4, 24]). For  $\gamma \rightarrow 0$  the Ginzburg–Landau energy  $E$  converges to the interface length  $2\pi|r_1(t) + r_2(t)|$  (see Figure 3). In the numerical computations we chose a uniform time step size of  $\tau = \gamma^2/10$ . We adaptively refined elements  $K \in \mathcal{T}^{(j)}$  which the estimator  $\eta_h(K) \geq \frac{1}{2} \max_{K' \in \mathcal{T}} \eta_h(K')$ , but limited the mesh size to  $h \leq \gamma/16$ . Further reduction of step sizes did not lead to a significant change in  $E(t)$  or  $\Lambda(t)$ .

When the inner interface vanishes, the system undergoes a topological change and the numerically computed principal eigenvalue shows the expected peak in  $\Lambda$  that grows proportionally

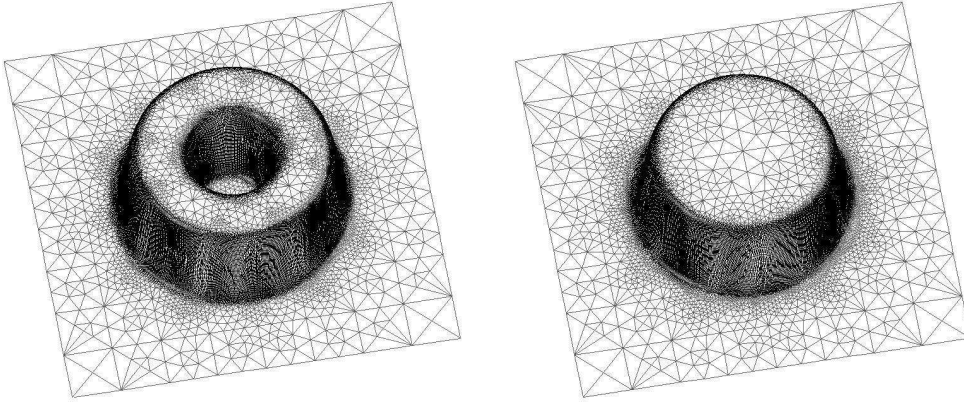


FIG. 2. Experiment 1. Left: initial values; right: final state.

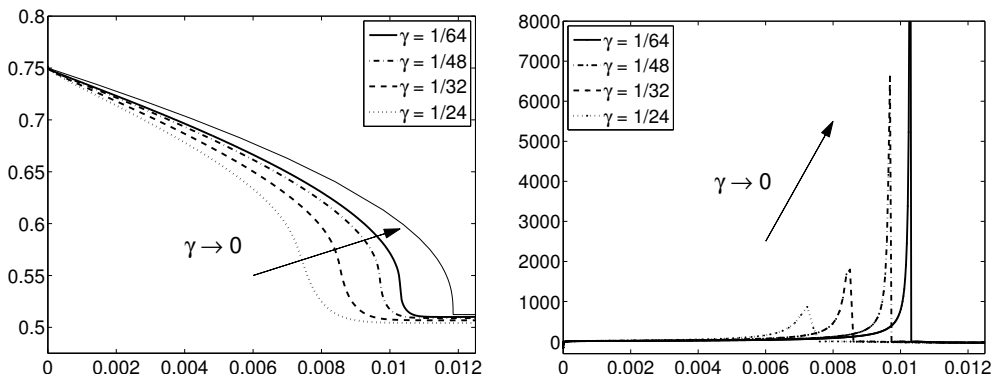


FIG. 3. Left: Ginzburg–Landau energy  $E(t)/(2\pi)$  and interface length in the Mullins–Sekerka problem (thin solid line); right: numerically computed negative principal eigenvalue  $\Lambda(t)$ .

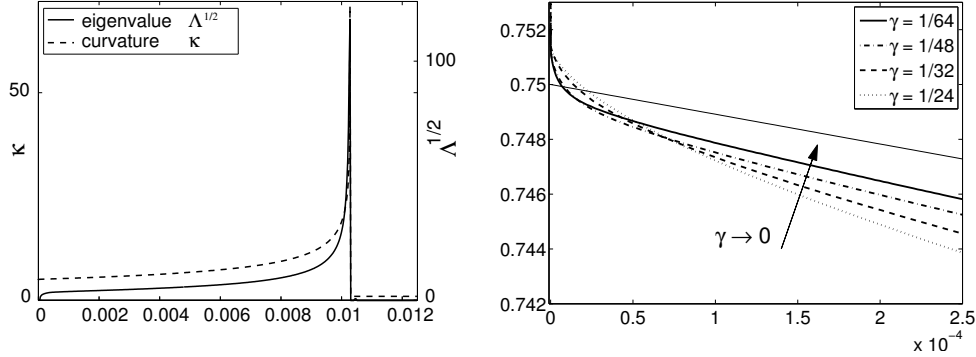


FIG. 4. Experiment 1. Left: curvature of the inner interface in relation to the principal eigenvalue; right: the energy  $E(t)$  indicates a rapid decay of perturbations in the initial values.

to  $\gamma^{-1}$  (see Figure 3). The relation of the principal eigenvalue to the maximal interface curvature in the solution is illustrated in Figure 4. With decreasing parameter size  $\gamma$ , the interface thickness reduces and larger curvatures can be resolved.

In the given experiment, the moving fronts of the solution show a stable profile across the interface. The chosen initial data  $\rho_0$  does not perfectly match such a profile. Thus, the Ginzburg–Landau energy  $E$  is initially larger than the corresponding energy in the Mullins–Sekerka problem, but shows a fast relaxation during a time proportional to  $\gamma$  (see Figure 4).

## 6.2 Smooth transition layers after $T \sim \gamma$

In [22] it is stated that perturbations of a smooth transition layer between the bulk phases vanish within a short time frame. The maximal interface curvature can also be expected to be uniformly bounded after short times, as long as there are no topological changes. The numerical experiments below confirm a uniform lower bound for the principal eigenvalue  $-\Lambda$  after a time period of order  $O(\gamma)$ . This is sufficient for robust error control on the whole time interval  $[0, T]$  (cf. Remark 3.5). In the following experiments we used  $\tau = \gamma^2/400$  and  $h \leq \gamma/16$ .

EXPERIMENT 2 The initial configuration is given by a single circular interface

$$\rho_0(x) = -\tanh\left(\frac{|x| - r_1}{2\gamma\sqrt{2}}\right),$$

where the profile is flat compared to the final state in Experiment 1. During the time evolution, there are no other changes in the solution than a steepening of the interface profile to  $-\tanh\left(\frac{|x| - r_1}{\gamma\sqrt{2}}\right)$ .

From Figure 5 we see that within a time proportional to  $\gamma$  the initially large Ginzburg–Landau energy  $E$  reaches a lower level related to the final interface profile. As required by the error estimate, the negative principal eigenvalue  $\Lambda$  reduces to order  $O(1)$  during a period  $t \sim \gamma$ .

When we prescribe initial values  $\rho_0$  with a steep profile close to a jump, the principal eigenvalue initially takes large negative values of  $\Lambda$ , which are uncritical with respect to the error estimate. Again,  $\Lambda$  relaxes to order  $O(1)$  within a time frame proportional to  $\gamma$ .



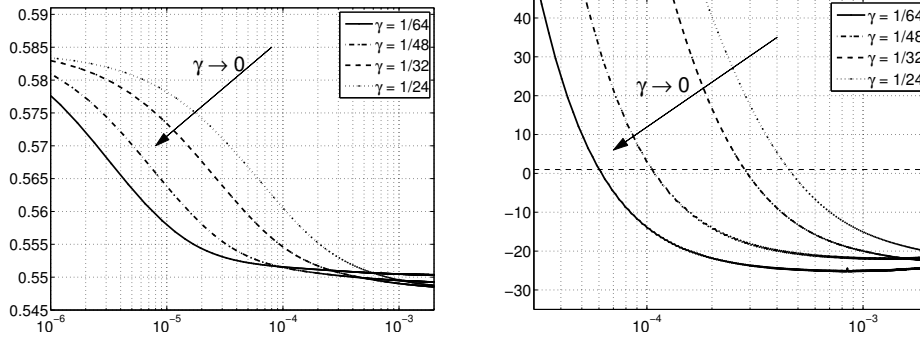


FIG. 5. Experiment 2. Left: reduction of the Ginzburg–Landau energy  $E(t)$ ; right: negative principal eigenvalue  $\Lambda(t)$ .

EXPERIMENT 3 The initial values take the form

$$\rho_0(x) = -\tanh\left(\frac{|x| - r_1}{\gamma\sqrt{2}}\right) + \text{noise}.$$

Here, the circular interface is perturbed by noise on a length scale between the mesh width  $h$  and the interface thickness  $\gamma$  (see Figure 6). As in Experiment 2, the numerical results confirm that the negative principal eigenvalue drops below an upper bound of order  $O(1)$  within a time proportional to  $\gamma$ .

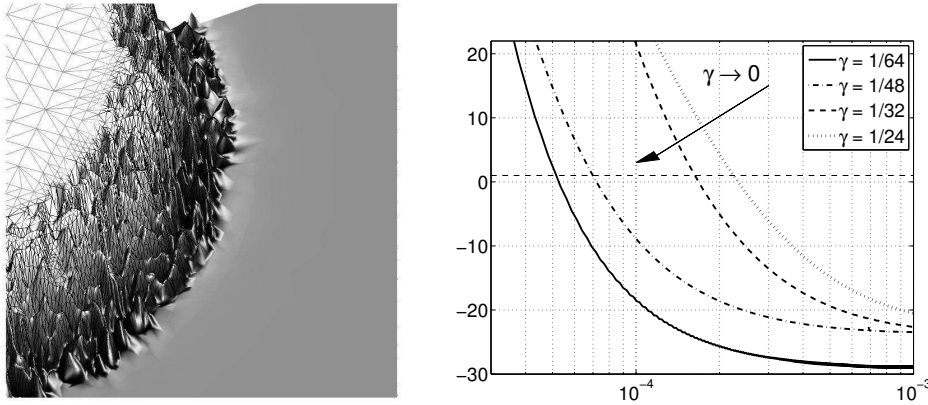


FIG. 6. Experiment 3. Left: initial values; right: principal eigenvalue  $-\Lambda(t)$ .

### 6.3 Application to Cahn–Larché equations

We set  $\kappa = 0.1$  and choose the elasticity tensor to be of cubic symmetry by setting  $\mathcal{C}_{1111} = \mathcal{C}_{2222} = 2$ ,  $\mathcal{C}_{1122} = 1$  and  $\mathcal{C}_{1212} = 20$ .

EXPERIMENT 4 We set  $\gamma = 1/64$  and on  $\Omega = (-2, 2)^2$  we prescribe pseudo-random data representing a mixture with a mass fraction of 48% for one of the phases. Until  $t \approx 0.0002$ , spinodal decomposition takes place. Then regions of pure phases have developed but large parts of the domain

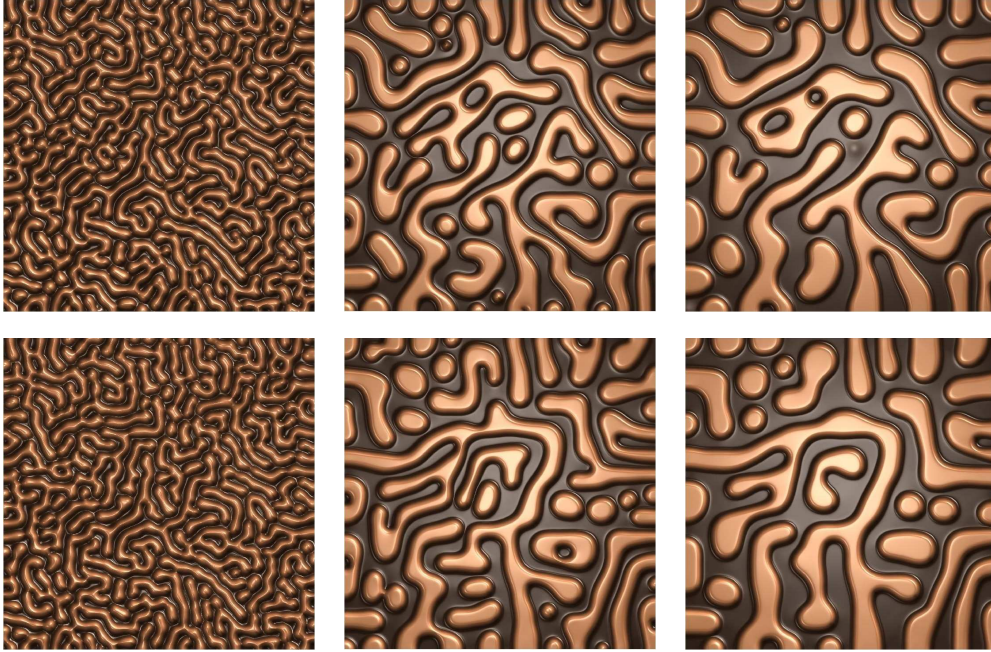


FIG. 7. Experiment 4. Coarsening after spinodal decomposition: snapshots of the solution  $\rho_h$  at  $t = 0.0002$ ,  $t = 0.0024$  and  $t = 0.006$ . Top row: Cahn-Hilliard evolution; bottom row: Cahn-Larché evolution with the same initial data and cubic symmetry of the elasticity tensor.

are occupied by interfaces. The subsequent coarsening of the microstructure is known as Oswald ripening. It allows adaptive coarsening of the mesh while still tracking the moving interfaces with locally refined elements. In the numerical calculations we set  $\tau = \gamma^2/64$  and limited the mesh width to  $h \leq \gamma/2$ . While we used a uniform grid with  $h = \gamma/2$  until  $t = 0.001$ , we afterwards adaptively refined and coarsened the mesh so that  $\eta_h \leq \gamma/(2\tau)$  and  $\eta_c \leq \gamma/(256\tau)$ , resulting in a significant reduction in the degrees of freedom (see Figure 8). The anisotropy in the elasticity tensor leads to a preferred orientation of the phase boundaries in the directions of the coordinate axes (see Figure 7).

**EXPERIMENT 5** We choose the same initial data as in Experiment 1, but now we are considering Cahn-Larché equations with material parameters as in Experiment 4. First, we set  $\gamma = 1/32$ . On a uniform grid with  $h = 1/256$  and using  $\tau = \gamma^2/32$ , the indicator  $\eta_h(t)$  is bounded by  $\theta := \gamma/(4\tau)$ . Figure 9 shows that the required amount of degrees of freedom to guarantee  $\eta_h(t) \leq \theta = \gamma/(4\tau)$  and even  $\eta_h(t) \leq \theta/2$  is far less on nonuniform grids that were adapted according to the indicator  $\eta_h$ . Next, we simulated the same problem on several uniform grids, where we varied  $h$  between  $1/128$  and  $1/512$  and set  $\tau = 16\gamma^2 h$ . For each simulation, we computed the quantity  $\|\eta\|_{L^2([0,T])}$  that enters Theorem 3.4. The numerical results in Figure 9 show  $\|\eta\|_{L^2([0,0.012])} \sim h$ , which is known to be the optimal scaling when approximating a linear parabolic problem with P1 finite elements and an implicit Euler time discretization with  $\tau \sim h$ . We also analyzed the dependence of  $\|\eta\|_{L^2([0,T])}$  on  $\gamma$ . On a uniform mesh with  $h = 1/256$  and  $\tau = h/128$  we performed several simulations where  $\gamma$  was varied between  $1/16$  and  $1/64$ . From the results in Figure 9 we conclude  $\|\eta\|_{L^2([0,0.012])} \sim \gamma^{-7/2}$ .

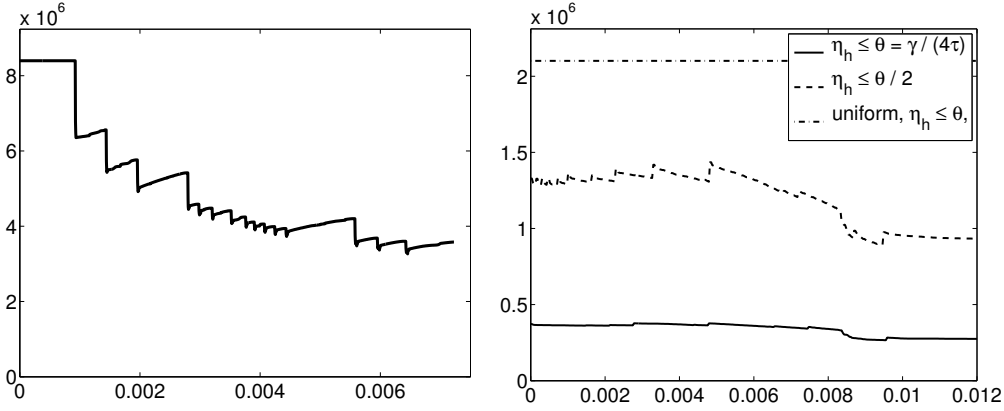


FIG. 8. Degrees of freedom over time. Left: in Experiment 4, coarsening of the microstructure allows an overall reduction of degrees of freedom over time. Right: Degrees of freedom required in Experiment 5 to ensure  $n_h \leq \theta = \gamma/(4\tau)$  on a uniform and on adaptively refined meshes.

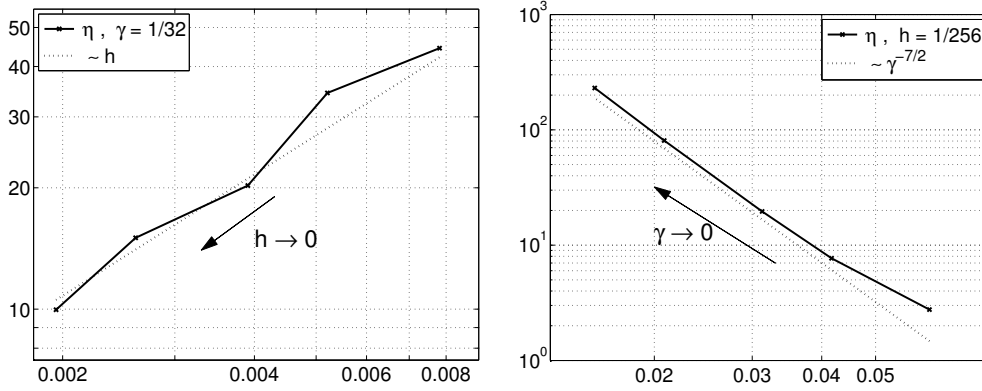


FIG. 9. Asymptotic behavior of  $\|\eta\|_{L^2((0,T))}$ . From Experiment 5 we see  $\|\eta\|_{L^2((0,0.012))} \sim h$  (left) and  $\|\eta\|_{L^2((0,0.012))} \sim \gamma^{-7/2}$  (right).

#### 6.4 Merging of particles

**EXPERIMENT 6** We choose initial values  $\rho_0$  that describe two circular particles, where one is slightly larger than the other. The radius of the larger particle is  $1/6$  whereas the radius of the other one is about 5.5% smaller. During the time evolution, the larger particle grows whereas the smaller one shrinks and is finally absorbed by the larger one by diffusion (see Figure 10).

The numerical experiment underlines the importance of tracking the approximated principal eigenvalue. As required for the error control,  $\Lambda(t)$  stays uniformly bounded with respect to  $\gamma^{-1}$  as long as there is no topological change in the solution. The critical point in time, when the smaller particle vanishes, is a priori unknown but can be detected by a peak in  $\Lambda(t)$  that is proportional to  $\gamma^{-1}$ . To illustrate the influence of the elasticity we compared the numerical solution with the results of a simulation where elasticity was neglected but all other parameters have been kept fixed

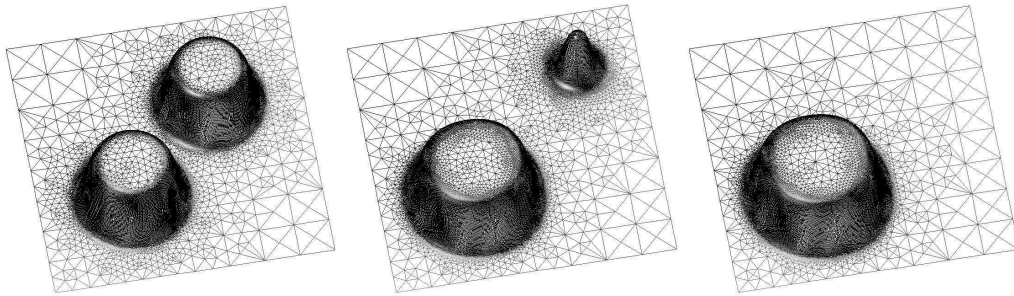


FIG. 10. Experiment 6. Merging of two particles; snapshots of the solution  $\rho$  of the Cahn–Larché equation with homogeneous elasticity and  $\gamma = 1/32$  are shown for  $t = 0$ ,  $t = 0.282$  and  $t = 0.3$ .

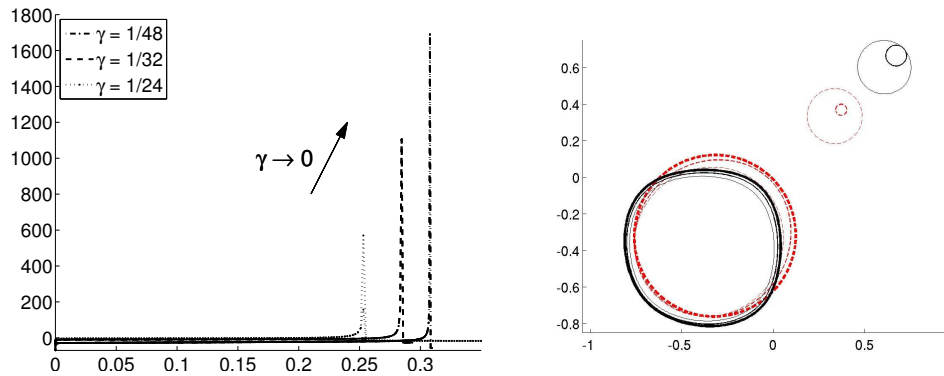


FIG. 11. Experiment 6. Left: Numerically computed eigenvalues. The singularity reflects the topological change when the smaller particle vanishes; right: comparison of the interface shape when elastic effects are neglected (dashed lines). Isolines  $\rho^{(j)}(x) = 0$  are shown at times short before and after the vanishing of the smaller particle.

and the same initial data was used. Due to the anisotropy of the elasticity tensor the interface shows a more square like shape, compared to the pure Cahn–Hilliard case, where particles always develop a spherical shape. Moreover with elasticity included, the particles stay at a larger distance from each other (see Figure 11).

### Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft through the DFG Research Center MATHEON ‘Mathematics for key technologies’.

### REFERENCES

1. ALIKAKOS, N., BATES, P., & CHEN, X. Convergence of the Cahn–Hilliard equation to the Hele–Shaw model. *Arch. Ration. Mech. Anal.* **128** (1994), 165–205. Zbl 0828.35105 MR 1308851
2. BARRETT, J. W., & BLOWEY, J. F. 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. Zbl 0882.65129 MR 1464653

3. BARTELS, S. A posteriori error analysis for time-dependent Ginzburg–Landau type equations. *Numer. Math.* **99** (2005), 557–583. Zbl 1073.65089 MR 2121069
4. BATES, P. W., CHEN, X., & DENG, X. A numerical scheme for the two phase Mullins–Sekerka problem. *Electron. J. Differential Equations* **1995**, no. 11, 27 pp. Zbl 0823.65138 MR 1345250
5. CHEN, X. Spectrum for the Allen–Cahn, Cahn–Hilliard, and phase-field equations for generic interfaces. *Comm. Partial Differential Equations* **19** (1994), 1371–1395. Zbl 0811.35098 MR 1284813
6. CIARLET, P. G. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam (1978), Zbl 0383.65058 MR 0520174
7. COPETTI, M. I. M. & ELLIOTT, C. M. Numerical analysis of the Cahn–Hilliard equation with a logarithmic free energy. *Numer. Math.* **63** (1992), 39–65. Zbl 0762.65074 MR 1182511
8. DE MOTTONI, P., & SCHATZMAN, M. Geometrical evolution of developed interfaces. *Trans. Amer. Math. Soc.* **347** (1995), 1533–1589. Zbl 0840.35010 MR 1672406
9. DREYER, W., & MÜLLER, W. A study of the coarsening in tin/lead solders. *Int. J. Solids Struct.* **37** (2000), 3841–3871. Zbl 0973.74059
10. ELLIOTT, C. M. The Cahn–Hilliard model for the kinetics of phase separation. In: *Mathematical Models for Phase Change Problems* (Óbidos, 1988), J.-F. Rodrigues (ed.), Int. Ser. Numer. Math. 88, Birkhäuser (1989), 35–73. Zbl 0692.73003 MR 1038064
11. ELLIOTT, C. M., FRENCH, D. A., & MILNER, F. A. A second order splitting method for the Cahn–Hilliard equation. *Numer. Math.* **54** (1989), 575–590. Zbl 0668.65097 MR 0978609
12. ELLIOTT, C. M., & ZHENG, S. M. On the Cahn–Hilliard equation. *Arch. Ration. Mech. Anal.* **96** (1986), 339–357. Zbl 0624.35048 MR 0855754
13. FENG, X., & PROHL, A. Numerical analysis of the Allen–Cahn equation and approximation for mean curvature flows. *Numer. Math.* **94** (2003), 33–65. Zbl 1029.65093 MR 1971212
14. FENG, X., & PROHL, A. Error analysis of a mixed finite element method for the Cahn–Hilliard equation. *Numer. Math.* **99** (2004), 47–84. Zbl 1071.65128 MR 2101784
15. FENG, X., & WU, H. 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. Zbl 1174.65035 MR 2464735
16. FRATZL, P., PENROSE, O., & LEBOWITZ, J. L. Modeling of phase separation in alloys with coherent elastic misfit. *J. Statist. Phys.* **95** (1999), 1429–1503. Zbl 0952.74052 MR 1712452
17. GARCKE, H. On mathematical models for phase separation in elastically stressed solids. Habilitation thesis, Friedrich-Wilhelms-Universität, Bonn (2000).
18. GARCKE, H., & WEIKARD, U. Numerical approximation of the Cahn–Larché equation. *Numer. Math.* **100** (2005), 639–662. Zbl 1099.74063 MR 2194588
19. KESSLER, D., NOCHETTO, R. H., & SCHMIDT, A. A posteriori error control for the Allen–Cahn problem: circumventing Gronwall’s inequality. *Math. Model. Numer. Anal.* **38** (2004), 129–142. Zbl 1075.65117 MR 2073933
20. LADYZHENSKAYA, O. A., & URAL’TSEVA, N. N. *Linear and Quasilinear Elliptic Equations*. Academic Press, New York (1968). Zbl 0164.13002 MR 0244627
21. LARSON, M. G. A posteriori and a priori error analysis for finite element approximations of self-adjoint elliptic eigenvalue problems. *SIAM J. Numer. Anal.* **38** (2000), 608–625. Zbl 0974.65100 MR 1770064
22. PEGO, R. Front migration in the nonlinear Cahn–Hilliard equation. *Proc. Roy. Soc. London A* **422** (1989), 261–278. Zbl 0701.35159 MR 0997638
23. VERFÜRTH, R. *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Teubner, Stuttgart (1996). Zbl 0853.65108
24. WEIKARD, U. Numerische Lösungen der Cahn–Hilliard-Gleichung und der Cahn–Larché-Gleichung. Dissertation, Friedrich-Wilhelms-Universität, Bonn (2002).