

NUMERICAL SIMULATION OF COARSENING IN BINARY SOLDER ALLOYS

CARSTEN GRÄSER, RALF KORNHUBER, AND ULI SACK

ABSTRACT. Coarsening in solder alloys is a widely accepted indicator for possible failure of joints in electronic devices. Based on the well-established Cahn–Larché model with logarithmic chemical energy density [20], we present a numerical environment for the efficient and reliable simulation of coarsening in binary alloys. Main features are adaptive mesh refinement based on hierarchical error estimates, fast and reliable algebraic solution by multigrid and Schur–Newton multigrid methods, and the quantification of the coarsening speed by the temporal growth of mean phase radii. We provide a detailed description and a numerical assessment of the algorithm and its different components, together with a practical application to a eutectic AgCu brazing alloy.

AMS classification: 65M55, 65M60, 65N30, 65N55, 65Z05

Keywords: Cahn–Larché system, phase separation, adaptive finite elements, Schur–Newton multigrid

1. INTRODUCTION

The life span of electronic devices strongly depends on the reliability of solder joints connecting the different components. As voids and cracks typically develop at phase boundaries, a well-known source of failure is thermomechanically induced phase separation in solder alloys, often called coarsening. Though there is good knowledge about the coarsening of classical tin–lead solders, such alloys are intended to be significantly reduced worldwide and are even banned in the European Union since 2006, in order to avoid the distribution of lead by electronic waste. The investigation of environmentally friendly substitutes based both on experiments and numerical simulation is still underway.

About ten years ago, Dreyer and Müller [20] utilized the framework of stress-induced diffusion [15] to derive meanwhile well-established Cahn–Larché models for thermomechanically induced phase separation of binary solder alloys. Such models consist of a Cahn–Hilliard system, accounting for spinodal decomposition and Ostwald ripening, coupled with an elasticity equation that represents thermo-mechanical interaction. In the Cahn–Hilliard system, both theoretical considerations [28, 29] and practical reasoning [9, 10, 50] suggest a chemical energy density of logarithmic type.

The singular behavior of logarithmic chemical energy densities turned out to be one of the major challenges of Cahn–Larché systems both in analysis and numerical approximation. First existence and uniqueness results were obtained by Garcke [22, 23, 24], who also investigated sharp interface limits. Related results for a viscous Cahn–Larché system were obtained by Bonetti et al. [11]. Upper bounds for time-averaged coarsening rates, i.e. for the average speed of demixing of the alloy, have been provided by Novick-Cohen et al. [52, 53] in absence of mechanical effects, i.e.

This work was supported by the DFG Research Center MATHEON.

for the pure Cahn–Hilliard system. These results extend earlier work of Kohn and Otto [43] for a quartic chemical energy density,

Numerical simulations with Cahn–Larché systems are facing both locally small mesh sizes, as required by the spatial resolution of the diffuse interface, and the algebraic solution of corresponding large-scale systems with logarithmic nonlinearity occurring in each time step. First qualitative numerical studies of Garcke et al. [26] utilize local adaptive mesh refinement, accounting for the diffuse interface, and, in order to enable algebraic solution by standard Newton methods, content themselves with a quartic chemical energy density. An accompanying paper contains the convergence analysis of the underlying implicit Euler discretization in time and finite element discretization in space [27]. Later, Merkle [48] aimed at quantitative results based on physical data. However, still lacking for a suitable algebraic solver, he used smooth spline interpolations instead of the logarithmic chemical energy density. Recent numerical simulations of coarsening in a eutectic AgCu alloy by Anders and Weinberg [1] suffer from a severe under-resolution of the diffuse interface.

In this paper, we present a numerical environment for the efficient and reliable simulation of coarsening in binary alloys. A semi-implicit Euler scheme provides the decomposition into a Cahn–Hilliard system and an elasticity equation. We analyze existence and uniqueness of the resulting spatial problems. Spatial discretization by adaptive finite elements based on hierarchical a posteriori error estimation [36] provides an efficient resolution of the diffuse interface. Careful assembling by high-order quadrature rules instead of mass lumping provides mass conservation, even for temporally varying grids [31, p. 42, Section 6.2]. Recent Schur–Newton solvers [30, 31, 34, 35, 32] allow for an efficient and reliable solution of the resulting large-scale algebraic systems with logarithmic nonlinearity without any regularization but with linear multigrid efficiency. In order to provide direct compatibility with experimental results, the coarsening speed is quantified by the temporal growth of the *mean phase radius* [9] rather than the inverse of interfacial energy [26, 43].

In our numerical experiments, we observed optimal convergence rates of the adaptive finite element discretization and local mesh independence both of the Schur–Newton method and of a classical multigrid solver for the elasticity problem. Equilibrium concentrations are reproduced up to 0.04% inside the phases and mass is preserved up to $1.9 \cdot 10^{-9}\%$ after 2000 time steps. Coarsening is enhanced by increasing influence of thermomechanical stress, as expected. In quantitative simulations, the temporal growth of the mean phase radius seems to strongly depend on the selection of the chemical free energy. More precisely, replacing a chemical free energy of logarithmic type by a smooth polynomial interpolate (cf., e.g., Merkle [48]), turned out to slow down the coarsening dynamics significantly (cf. Figure 6 in Subsection 6.3 below).

As a practical application, we present the simulation of coarsening in a eutectic AgCu brazing alloy. In agreement with experimental results, we found only minor influence of mechanical interactions for this alloy.

2. MATHEMATICAL MODELING OF BINARY SOLDER ALLOYS

2.1. Generalized Cahn–Larché equations. We consider the local mass concentrations c_A, c_B of the two constituents A and B of some binary alloy in a bounded domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$. As concentrations satisfy the pointwise constraints $c_A, c_B \geq 0$ and $c_A + c_B = 1$, we can eliminate c_B and reduce our considerations to the single concentration variable $c = c_A \in [0, 1]$. Introducing the displacement field \mathbf{u} and the corresponding linearized strain $\varepsilon(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^T)$ we consider the

generalized Ginzburg–Landau free energy

$$(1) \quad \mathcal{E}(c, \mathbf{u}) = \int_{\Omega} \frac{1}{2} \Gamma(c) \nabla c \cdot \nabla c + \Psi(c) + \mathcal{W}(c, \varepsilon(\mathbf{u})) \, dx - \int_{\partial\Omega} \mathbf{u} \cdot \mathbf{g} \, ds.$$

Here, the interfacial energy term $\Gamma(c) \nabla c \cdot \nabla c$ penalizes large concentration gradients and involves the concentration-dependent symmetric, positive definite matrix $\Gamma(c) \in \mathbb{R}^{d \times d}$. A most simple choice is $\Gamma(c) = \gamma(c) \mathbf{ld}$, where \mathbf{ld} denotes the identity matrix in $\mathbb{R}^{d \times d}$ and

$$(2) \quad \gamma(c) = c\gamma_A + (1-c)\gamma_B$$

relies on linear interpolation of constant parameters γ_A, γ_B associated with the pure constituents A, B , respectively. The double-well potential Ψ represents the chemical energy density and drives the uphill diffusion in the separation process towards the equilibrium concentrations. Chemical energy densities of logarithmic type are suggested both by theoretical and practical considerations [20, 28, 29]. To fix the ideas, we consider the Margules ansatz [10]

$$(3) \quad \begin{aligned} \Psi(c) &= \beta_0 R \theta (c \log(c) + (1-c) \log(1-c)) \\ &\quad + \beta_1 (1-c) + \beta_2 c + c(1-c) (\beta_3 c + \beta_4 (1-c)) \end{aligned}$$

with given temperature $\theta > 0$, universal gas constant R , and material parameters $\beta_i, i = 0, \dots, 4$. Note that the choice $\beta_0 = 1/R, \beta_1 = \beta_2 = 0$, and $\beta_3 = \beta_4 = \frac{\theta c}{2}$ leads to the classical logarithmic potential

$$(4) \quad \Psi(c) = \frac{1}{2} \theta (c \log(c) + (1-c) \log(1-c)) + \frac{1}{2} \theta c (1-c)$$

with critical temperature $\theta_c > \theta$.

The elastic energy density \mathcal{W} takes the form

$$(5) \quad \mathcal{W}(c, \varepsilon(\mathbf{u})) = \frac{1}{2} (\varepsilon(\mathbf{u}) - \bar{\varepsilon}(c)) : \sigma.$$

We assume Hooke's law $\sigma = \mathcal{C}(c) (\varepsilon(\mathbf{u}) - \bar{\varepsilon}(c))$ with a given, positive definite tensor $\mathcal{C}(c)$ that fulfills the usual symmetry conditions of linear elasticity [51] and given eigenstrains $\bar{\varepsilon}(c)$. Both Hooke's tensor $\mathcal{C}(c)$ and the eigenstrains $\bar{\varepsilon}(c)$ are allowed to depend on the concentration c , e.g., linearly as (2). The boundary integral term finally accounts for the prescribed boundary stress \mathbf{g} .

We postulate conservation of mass of the components of the alloy. Hence, the evolution of c is given by

$$(6) \quad \partial_t c = -\operatorname{div} J$$

with some diffusional flux J . We assume that J is defined by

$$J = -M(c) \nabla w,$$

where $M(c)$ denotes a concentration-dependent mobility matrix and

$$(7) \quad w = \frac{\partial \mathcal{E}}{\partial c} = -\operatorname{div} (\Gamma(c) \nabla c) + \frac{1}{2} \nabla c^T \Gamma'(c) \nabla c + \Psi'(c) + \frac{\partial}{\partial c} \mathcal{W}(c, \varepsilon)$$

is the chemical potential. Since mechanical equilibrium is expected to be attained much faster than thermodynamical equilibrium, we assume that

$$(8) \quad \frac{\partial \mathcal{E}}{\partial \mathbf{u}} = 0$$

holds throughout the evolution. Selecting some final time $T > 0$, the above equations constitute the generalized Cahn–Larché system

$$(9a) \quad \partial_t c - \operatorname{div} M(c) \nabla w = 0$$

$$(9b) \quad -\operatorname{div} (\Gamma(c) \nabla c) + \frac{1}{2} \nabla c^T \Gamma'(c) \nabla c + \Psi'(c) + \frac{\partial}{\partial c} \mathcal{W}(c, \varepsilon(\mathbf{u})) - w = 0$$

$$(9c) \quad \operatorname{div} (\mathcal{C}(c) (\varepsilon(\mathbf{u}) - \bar{\varepsilon}(c))) = 0$$

on $\Omega \times [0, T]$ for the unknown concentration c , chemical potential w , and displacement \mathbf{u} . We prescribe the Neumann boundary conditions

$$(10) \quad \Gamma(c)\nabla c \cdot \mathbf{n} = 0, \quad \nabla w \cdot \mathbf{n} = 0, \quad \sigma \cdot \mathbf{n} = \mathbf{g} \quad \text{on } \partial\Omega \times [0, T]$$

with \mathbf{n} denoting the outward unit normal to $\partial\Omega$ and given boundary stress \mathbf{g} . Finally, we impose the initial condition

$$(11) \quad c(\cdot, 0) = c^0 \quad \text{on } \Omega.$$

A thermodynamical derivation of the Cahn–Larché system (9) using a higher gradient theory of mixtures was carried out by Böhme et al. [9, 10].

Observe that the Cahn–Larché system (9) is invariant under infinitesimal rigid body motions

$$\ker(\varepsilon) = \{v : \Omega \rightarrow \mathbb{R}^d \mid v(x) = Ax + b \text{ with a skew symmetric matrix } A\}$$

representing the kernel of the differential operator $\operatorname{div}(\mathcal{C}(c)\varepsilon(\cdot))$. As only the strain ε enters the phase field equations (9a) and (9b), the remaining elasticity equation (9c) can be considered in the corresponding quotient space

$$\mathbf{H} = (H^1(\Omega))^d / \ker(\varepsilon).$$

In preparation of a weak formulation of (9), we assume that there is a splitting

$$(12) \quad \Psi(c) = \Psi_1(c) + \Psi_2(c)$$

of $\Psi : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ into a *convex, piecewise smooth function* $\Psi_1 : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ with domain $\operatorname{dom}(\Psi_1) = [0, 1]$ and a *globally smooth function* $\Psi_2 : \mathbb{R} \rightarrow \mathbb{R}$. For example, the logarithmic potential (4) allows for such a splitting with the definitions

$$(13) \quad \Psi_1(c) = \frac{1}{2}\theta(c \log(c) + (1-c) \log(1-c)), \quad \Psi_2(c) = \frac{1}{2}\theta_c c(1-c).$$

We introduce the corresponding functionals

$$(14) \quad \psi_1(c) = \int_{\Omega} \Psi_1(c) \, dx, \quad \psi_2(c) = \int_{\Omega} \Psi_2(c) \, dx.$$

Now the weak formulation of the Cahn–Larché system (9) reads as follows.

(CL) Find $c \in L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^1(\Omega)')$ with the property $c(\cdot, 0) = c^0$, $w \in L^2(0, T; H^1(\Omega))$, and $\mathbf{u} \in L^2(0, T; \mathbf{H})$ such that

$$(15a) \quad \langle c_t, v \rangle + (M(c)\nabla w, \nabla v) = 0 \quad \forall v \in H^1(\Omega),$$

$$(15b) \quad \begin{cases} (\Gamma(c)\nabla c, \nabla(v-c)) - (w, v-c) \\ + \psi_1(v) - \psi_1(c) \geq (R(c, \mathbf{u}), v-c) \end{cases} \quad \forall v \in H^1(\Omega),$$

$$(15c) \quad (\mathcal{C}(c)(\varepsilon(\mathbf{u}) - \bar{\varepsilon}(c)), \varepsilon(\mathbf{v})) = \int_{\partial\Omega} \mathbf{g} \cdot \mathbf{v} \, ds \quad \forall \mathbf{v} \in \mathbf{H}$$

with

$$R(c, \mathbf{u}) = -\Psi_2'(c) - \frac{1}{2}(\nabla c)^T \Gamma'(c)\nabla c - \frac{\partial}{\partial c} \mathcal{W}(c, \varepsilon(\mathbf{u}))$$

holds a.e. in $(0, T]$. Here, (\cdot, \cdot) stands for the L^2 -inner product in $L^2(\Omega)$, $L^2(\Omega)^d$, and $L^2(\Omega)^{d \times d}$ and $\langle \cdot, \cdot \rangle$ for the duality pairing.

Remark 2.1. If Ψ is smooth on $(0, 1)$, as, e.g., the Margules potential (3), and if the concentration c is uniformly contained in $(0, 1)$, then (15b) can be equivalently rewritten as the variational equality

$$(\Gamma(c)\nabla c, \nabla v) + (\Psi_1'(c), v) - (w, v) = (R(c, \mathbf{u}), v) \quad \forall v \in H^1(\Omega).$$

However, large values of $\Psi'(c)$, as typically occurring in this formulation, often cause severe problems in numerical computations. Anticipating that our numerical

solver (cf. Section 4.1) relies on convexity rather than smoothness, we focus on the more general formulation (15b).

The following existence result is due to Garcke [22].

Theorem 2.1. *Assume that $\Omega \subset \mathbb{R}^d$ is a bounded domain with Lipschitz boundary $\partial\Omega$, the interfacial function takes the form $\Gamma(c) = \gamma I$ with unit matrix $I \in \mathbb{R}^{d \times d}$ and $\gamma > 0$ independent of c , the double-well potential Ψ is given by (4), the Hooke tensor $\mathcal{C}(c) = \mathcal{C}_0$ is independent of c , the eigenstrain takes the form $\bar{\varepsilon}(c) = c\bar{\varepsilon}_1 + \bar{\varepsilon}_0$ with given $\bar{\varepsilon}_1, \bar{\varepsilon}_0 \in \mathbb{R}^{d \times d}$, the boundary stress takes the form $\mathbf{g} = \sigma \mathbf{n}$ with constant stress tensor σ on $\partial\Omega \times (0, T]$, the mobility $M(c) = M_0 > 0$ is independent of c , and $c^0 \in H^1(\Omega)$ satisfies $c \in (0, 1)$ almost everywhere.*

Then the weak Cahn–Larché system (CL) admits a unique solution.

Existence and uniqueness results involving more general, concentration-dependent mobility $M(c)$, Hooke tensor $\mathcal{C}(c)$ and eigenstrains $\bar{\varepsilon}(c)$ have been presented by Merkle [47, 48]. However, these results are valid only for globally smooth approximations of logarithmic-type potentials Ψ which may strongly deteriorate numerical simulation, e.g. of coarsening rates (cf. Section 6.3).

3. DISCRETIZATION IN TIME AND SPACE

In this section we present a discretization of the weak formulation (CL) of the generalized Cahn–Larché system by an Euler-type discretization in time and finite elements in space. Since efficient approximation of the order parameter c requires different, locally refined spatial grids at different time instants, it is convenient to use Rothe’s method [13], i.e., (CL) is first discretized in time and the resulting spatial problems are then discretized in space, independently from each other.

3.1. Implicit time discretization. In order to avoid any time step restrictions, we apply an implicit Euler discretization to the second order term and to the convex part ψ_1 of the double-well potential ψ in the phase field equation (15b). The remaining, often concave part ψ_2 of ψ is taken explicitly (cf., e.g., [8]). For a discussion of such kind of semi-implicit time discretization and its fully implicit counterpart in the case of Allen–Cahn equations, we refer, e.g., to [7, 38]. Assuming moderate variation of the other solution-dependent coefficients, these functions are frozen at the preceding time step. Note that this leads to a decoupling of phase field equation (15b) and mechanics (15c), which will simplify the algebraic solution of the discretized spatial problems later on. Denoting

$$\mathcal{K} = \text{dom } \psi_1 = \{v \in H^1(\Omega) \mid v(x) \in [0, 1] \text{ a.e. in } \Omega\}$$

this approach results in the scheme:

(CL $_{\Delta t}$) For $n = 1, \dots$, find $(c^n, w^n, \mathbf{u}^n) \in \mathcal{K} \times H^1(\Omega) \times \mathbf{H}$ such that

$$(16a) \quad (c^n, v) + \Delta t (M(c^{n-1}) \nabla w^n, \nabla v) = (c^{n-1}, v) \quad \forall v \in H^1(\Omega),$$

$$(16b) \quad \begin{cases} (\Gamma(c^{n-1}) \nabla c^n, \nabla(v - c^n)) - (w^n, v - c^n) \\ + \psi_1(v) - \psi_1(c^n) \geq (R(c^{n-1}, \mathbf{u}^{n-1}), v - c^n) \end{cases} \quad \forall v \in \mathcal{K},$$

$$(16c) \quad (\mathcal{C}(c^n) (\varepsilon(\mathbf{u}^n) - \bar{\varepsilon}(c^n)), \varepsilon(\mathbf{v})) = \int_{\partial\Omega} \mathbf{g} \cdot \mathbf{v} \, ds \quad \forall \mathbf{v} \in \mathbf{H}.$$

with given initial value $c^0 \in \mathcal{K}$, displacement $\mathbf{u}^0 \in \mathbf{H}$ obtained from

$$\mathbf{u}^0 \in \mathbf{H} : \quad (\mathcal{C}(c^0) (\varepsilon(\mathbf{u}^0) - \bar{\varepsilon}(c^0)), \varepsilon(\mathbf{v})) = \int_{\partial\Omega} \mathbf{g} \cdot \mathbf{v} \, ds \quad \forall \mathbf{v} \in \mathbf{H},$$

and suitable time step size $\Delta t > 0$.

To show existence and uniqueness of solutions we impose the following conditions on coefficient functions and the initial value.

(A1) $M(\cdot)$, $\Gamma(\cdot)$, and $\mathcal{C}(\cdot)$ are uniformly bounded from below on $[0, 1]$, i.e., there are constants $\gamma_M, \gamma_\Gamma, \gamma_C > 0$ such that

$$\begin{aligned} \gamma_M |x|^2 &\leq M(c)x \cdot x & \forall c \in [0, 1], x \in \mathbb{R}^d, \\ \gamma_\Gamma |x|^2 &\leq \Gamma(c)x \cdot x & \forall c \in [0, 1], x \in \mathbb{R}^d, \\ \gamma_C |x|^2 &\leq \mathcal{C}(c)x : x & \forall c \in [0, 1], y \in \mathbb{R}^{d \times d}. \end{aligned}$$

(A2) The norms of $M(\cdot), \Gamma(\cdot), \mathcal{C}(\cdot), \bar{\varepsilon}(\cdot), \Gamma'(\cdot), \mathcal{C}'(\cdot), \bar{\varepsilon}'(\cdot)$ are uniformly bounded from above on $[0, 1]$.

(A3) The initial value c^0 is nontrivial in the sense that $0 < (c^0, 1) < |\Omega|$.

Theorem 3.1. *Assume that conditions (A1)–(A3) hold and that for a fixed $n > 0$ $c^{n-1} \in \mathcal{K}$, c^{n-1} is nontrivial in the sense that $0 < (c^{n-1}, 1) < |\Omega|$, and $\mathbf{u}^{n-1} \in \mathbf{H}$. Then there is a solution $(c^n, w^n, \mathbf{u}^n) \in \mathcal{K} \times H^1(\Omega) \times \mathbf{H}$ of (16) and $(c^n, \nabla w^n, \mathbf{u}^n)$ is unique.*

We only sketch the idea of the proof which is carried out in detail in the appendix. While existence and uniqueness of the elasticity equation (16c) is straightforward, the existence proof for the phase field system (16a)–(16b) is slightly more involved. It is based on the equivalence of these equations to a saddle point problem for the Lagrangian

$$\mathcal{L}^n(c, w) = \mathcal{J}^n(c) - (c - c^{n-1}, w) - \frac{\Delta t}{2} (M(c^{n-1})\nabla w, \nabla w)$$

with the strictly convex energy functional

$$\mathcal{J}^n(c) = \int_{\Omega} \frac{1}{2} \Gamma(c^{n-1}) \nabla c \cdot \nabla c \, dx + \frac{\gamma}{2} (c - c^{n-1}, 1)^2 + \psi_1(c) - (R(c^{n-1}, \mathbf{u}^{n-1}), c).$$

Note that (16a) yields $(c^n, 1) = (c^{n-1}, 1)$. Hence, by Theorem 3.1, $c^0 \in \mathcal{K}$ and $0 < (c^0, 1) < |\Omega|$ inductively implies existence for all time steps.

Corollary 3.2. *Assume that conditions (A1)–(A3) hold. Then $(\mathbf{CL}_{\Delta t})$ has a solution $(c^n, w^n, \mathbf{u}^n)_{n=1, \dots}$ and $(c^n, \nabla w^n, \mathbf{u}^n)_{n=1, \dots}$ is unique.*

□

3.2. Adaptive finite element discretization. We will now consider the adaptive finite element discretization of the stationary problems (16), as obtained in each step of $(\mathbf{CL}_{\Delta t})$.

3.2.1. Conforming finite element spaces on locally refined grids. We first introduce some notation concerning finite element spaces on hierarchies of locally refined grids.

Definition 3.1. *Let \mathcal{T}_1 and \mathcal{T}_2 be two simplicial partitions of $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$. Then \mathcal{T}_2 is called a refinement of \mathcal{T}_1 , if, for each $e \in \mathcal{T}_1$, the intersection $\mathcal{F}_e = \{e' \in \mathcal{T}_2 \mid \text{int } e' \cap \text{int } e \neq \emptyset\}$ is a simplicial partition of e . The refinement \mathcal{T}_2 is called regular, if, for each $e \in \mathcal{T}_1$, the partition \mathcal{F}_e is obtained by connecting the midpoints of the edges of e .*

Definition 3.2. *$(\mathcal{T}_0, \dots, \mathcal{T}_j)$ is called a (regular) grid hierarchy on Ω , if \mathcal{T}_0 is a conforming triangulation of Ω and each \mathcal{T}_i , $i = 1, \dots, j$, is a (regular) refinement and a conforming partition of a subset of \mathcal{T}_{i-1} . \mathcal{T}_i is called the i -th level grid of the grid hierarchy $(\mathcal{T}_0, \dots, \mathcal{T}_j)$.*

This notion reflects the implementation of adaptively refined grids in finite element codes, such as DUNE [5, 4]. Obviously, higher levels in the grid hierarchy, in general, only cover subsets of Ω . The corresponding partition of the whole domain Ω is the so-called *leaf grid*.

Definition 3.3. *Let $(\mathcal{T}_0, \dots, \mathcal{T}_j)$ be a (regular) grid hierarchy on Ω . Then the leaf grid denoted by $\mathcal{L}(\mathcal{T}_0, \dots, \mathcal{T}_j)$ is defined by*

$$\mathcal{L}(\mathcal{T}_0, \dots, \mathcal{T}_j) = \mathcal{T}_j \cup \bigcup_{i=0}^{j-1} \{e \in \mathcal{T}_i : \text{int } e \cap \text{int } e' = \emptyset \ \forall e' \in \mathcal{T}_{i+1}\}.$$

Note that the partition $\mathcal{T} = \mathcal{L}(\mathcal{T}_0, \dots, \mathcal{T}_j)$ of Ω consists of all elements of $\mathcal{T}_0, \dots, \mathcal{T}_j$ that are not refined. In general, \mathcal{T} involves so-called *hanging vertices*. A hanging vertex is the vertex of an element $e \in \mathcal{T}$ which is contained in, but is not a vertex of another element $e' \neq e$. The set of nodes, i.e. of vertices of \mathcal{T} which are no hanging vertices, is called $\mathcal{N}(\mathcal{T})$.

Now we introduce the space

$$\mathcal{S}(\mathcal{T}) = \left\{ v \in C(\overline{\Omega}) \mid v|_e \text{ is affine } \forall e \in \mathcal{T} \right\} \subset H^1(\Omega)$$

of piecewise linear conforming finite elements on the partition \mathcal{T} of Ω .

Lemma 3.3. *Let $(\mathcal{T}_0, \dots, \mathcal{T}_j)$ be a regular simplicial grid hierarchy on Ω and let $\mathcal{T} = \mathcal{L}(\mathcal{T}_0, \dots, \mathcal{T}_j)$. Then, for each $p \in \mathcal{N}(\mathcal{T})$ there is a uniquely determined function $\lambda_p^{\mathcal{T}} \in \mathcal{S}(\mathcal{T})$ such that $\lambda_p^{\mathcal{T}}(q) = \delta_{pq}$ holds for all $q \in \mathcal{N}(\mathcal{T})$. The set $\{\lambda_p^{\mathcal{T}} \mid p \in \mathcal{N}(\mathcal{T})\}$ is a basis of $\mathcal{S}(\mathcal{T})$.*

Proof. See [31, Theorem 3.1]. □

The basis given by Lemma 3.3 is called the *conforming nodal basis* of $\mathcal{S}(\mathcal{T})$. It reduces to the classical nodal basis, if no hanging vertices occur. Otherwise the values at hanging vertices are obtained by linear interpolation of the nodal values (see [31, Section 3.1] for details). We emphasize that a (regular) grid hierarchy $(\mathcal{T}_0, \dots, \mathcal{T}_j)$ gives rise to an associated hierarchy

$$(17) \quad \mathcal{S}_0 \subset \dots \subset \mathcal{S}_j = \mathcal{S}(\mathcal{T}) \subset H^1(\Omega)$$

of finite element spaces $\mathcal{S}_k = \mathcal{S}(\mathcal{L}(\mathcal{T}_0, \dots, \mathcal{T}_k))$ with $k = 0, \dots, j$.

3.2.2. Finite element discretization. In the following, we assume that \mathcal{T} is the leaf grid of an underlying simplicial grid hierarchy $(\mathcal{T}_0, \dots, \mathcal{T}_j)$ with an intentionally coarse initial grid \mathcal{T}_0 . We will discretize the spatial problems occurring in $(\mathbf{CL}_{\Delta t})$ with respect to the finite element space $\mathcal{S}_j = \mathcal{S}(\mathcal{T})$. The induced subspace hierarchy will be denoted according to (17). In particular, the discretization of (16c) is based on the discrete quotient space $\mathbf{H}_j = \mathcal{S}_j^d / \ker(\varepsilon) = \mathcal{S}_j^d \cap \mathbf{H}$ which is well defined since $\ker(\varepsilon) \subset \mathcal{S}_j^d$. Furthermore, we introduce the approximate nonsmooth nonlinear functional

$$\psi_1^{\mathcal{T}}(v) = \sum_{p \in \mathcal{N}(\mathcal{T})} \Psi_1(v(p)) \int_{\Omega} \lambda_p^{\mathcal{T}}(x) dx,$$

as obtained by replacing exact integration by a quadrature rule based on nodal interpolation in \mathcal{S}_j .

Assuming that $c^{\text{old}} \in \mathcal{K}$ and $\mathbf{u}^{\text{old}} \in \mathbf{H}$ are approximations of c^{n-1} and \mathbf{u}^{n-1} and denoting

$$\mathcal{K}_j = \mathcal{S}_j \cap \mathcal{K},$$

the discretized spatial problem in the n -th time step is given by

($\mathbf{CL}_{\Delta t, \mathcal{T}}^n$) Find $(c_{\mathcal{T}}^n, w_{\mathcal{T}}^n, \mathbf{u}_{\mathcal{T}}^n) \in \mathcal{K}_j \times \mathcal{S}_j \times \mathbf{H}_j$ such that

$$(18a) \quad (c_{\mathcal{T}}^n, v) + \Delta t (M(c^{\text{old}}) \nabla w_{\mathcal{T}}^n, \nabla v) = (c^{\text{old}}, v) \quad \forall v \in \mathcal{S}_j,$$

$$(18b) \quad \begin{cases} (\Gamma(c^{\text{old}}) \nabla c_{\mathcal{T}}^n, \nabla(v - c_{\mathcal{T}}^n)) - (w_{\mathcal{T}}^n, v - c_{\mathcal{T}}^n) \\ + \psi_1^{\mathcal{T}}(v) - \psi_1^{\mathcal{T}}(c_{\mathcal{T}}^n) \geq (R(c^{\text{old}}, \mathbf{u}^{\text{old}}), v - c_{\mathcal{T}}^n) \end{cases} \quad \forall v \in \mathcal{K}_j,$$

$$(18c) \quad (\mathcal{C}(c_{\mathcal{T}}^n) (\varepsilon(\mathbf{u}_{\mathcal{T}}^n) - \bar{\varepsilon}(c_{\mathcal{T}}^n)), \varepsilon(\mathbf{v})) = \int_{\partial\Omega} \mathbf{g} \cdot \mathbf{v} \, ds \quad \forall \mathbf{v} \in \mathbf{H}_j.$$

The algebraic solution of ($\mathbf{CL}_{\Delta t, \mathcal{T}}^n$) will be considered in Section 4.

Theorem 3.4. *Assume that conditions (A1)–(A3) hold and that for a fixed $n > 0$ $c^{\text{old}} \in \mathcal{K}$, c^{old} is nontrivial in the sense that $0 < (c^{\text{old}}, 1) < |\Omega|$, and $\mathbf{u}^{\text{old}} \in \mathbf{H}$. Then there is a solution $(c_{\mathcal{T}}^n, w_{\mathcal{T}}^n, \mathbf{u}_{\mathcal{T}}^n) \in \mathcal{K}_j \times \mathcal{S}_j \times \mathbf{H}_j$ of (18) and $(c_{\mathcal{T}}^n, \nabla w_{\mathcal{T}}^n, \mathbf{u}_{\mathcal{T}}^n)$ is unique. If there is furthermore a vertex $p \in \mathcal{N}(\mathcal{T})$ such that Ψ_1 is differentiable at $c_{\mathcal{T}}^n(p)$, then $w_{\mathcal{T}}^n$ is unique.*

Proof. The existence of $(c_{\mathcal{T}}^n, w_{\mathcal{T}}^n, \mathbf{u}_{\mathcal{T}}^n)$ and uniqueness of $(c_{\mathcal{T}}^n, \nabla w_{\mathcal{T}}^n, \mathbf{u}_{\mathcal{T}}^n)$ follows by the same arguments as in the proof of Theorem 3.1.

To show uniqueness of $w_{\mathcal{T}}^n$ let $p \in \mathcal{N}(\mathcal{T})$ such that Ψ_1 is differentiable at $\xi = c_{\mathcal{T}}^n(p)$. Then $\xi \in (0, 1)$ and we can use $v_{\pm} = c_{\mathcal{T}}^n \pm \delta \lambda_p^{\mathcal{T}}$ for sufficiently small $\delta > 0$ in (18b). Testing (18b) with v_+ and v_- for two solutions $w_{\mathcal{T},1}^n$ and $w_{\mathcal{T},2}^n$, respectively, and adding both inequalities yields

$$(w_{\mathcal{T},2}^n - w_{\mathcal{T},1}^n, \lambda_p^{\mathcal{T}}) \geq \left(\frac{\Psi_1(\xi) - \Psi_1(\xi + \delta)}{\delta} - \frac{\Psi_1(\xi - \delta) - \Psi_1(\xi)}{\delta} \right) \int_{\Omega} \lambda_p^{\mathcal{T}}(x) \, dx.$$

Taking the limit $\delta \rightarrow 0$ and switching the role of $w_{\mathcal{T},1}^n$ and $w_{\mathcal{T},2}^n$ we get

$$0 = (w_{\mathcal{T},2}^n - w_{\mathcal{T},1}^n, \lambda_p^{\mathcal{T}}) = (w_{\mathcal{T},2}^n - w_{\mathcal{T},1}^n, 1) \frac{(\lambda_p, 1)}{|\Omega|}.$$

The last equation holds because $w_{\mathcal{T},2}^n - w_{\mathcal{T},1}^n$ is constant and implies that this constant is zero. \square

Note that the condition for uniqueness of $w_{\mathcal{T}}^n$ is always satisfied for logarithmic potentials of the form (3) because $0 < c_{\mathcal{T}}^n < 1$. For the obstacle potential it is satisfied if there is at least on vertex $p \in \mathcal{N}(\mathcal{T})$ in the discrete interfacial region.

Usually, c^{old} is a finite element function on a grid \mathcal{T}^{old} . In case of adaptive refinement, \mathcal{T}^{old} is usually different from \mathcal{T} . Desired properties as, e.g., mass conservation $(c_{\mathcal{T}}^n, 1) = (c^{\text{old}}, 1)$ then impose the following restrictions on the choice of possible approximations of the occurring L^2 -inner products.

- (i) To guarantee mass conservation, the approximate inner products used on the left and right hand side of (18a) should both be exact for $v = 1$.
- (ii) To guarantee that (18a) is equivalent to $c_{\mathcal{T}}^n = c^{\text{old}}$ for $\Delta t \rightarrow 0$, the approximate inner products on the left and right hand side of (18a) should be the same.
- (iii) To preserve the symmetric saddle point structure (see proof of Theorem 3.4), the approximate inner products used on the left hand side of (18a) and (18b) should be the same.

As a consequence of (i), lumping should be carried out with respect to a fine grid that contains both \mathcal{T} and \mathcal{T}^{old} . In general, lumping then no longer provides a diagonal matrix and thus its main advantage is lost. Hence, lumping is avoided here.

We emphasize that for affine parameter functions Γ , \mathcal{C} , $\bar{\varepsilon}$, and M , e.g., of the form (2), and finite element functions c^{old} and \mathbf{u}^{old} , all integrals involved in ($\mathbf{CL}_{\Delta t, \mathcal{T}}^n$)

can be calculated exactly using suitable quadrature rules. Numerical computations indicate that this is particularly important for the leading order terms. As expected, such quadrature rules usually involve a fine grid that contains both \mathcal{T} and \mathcal{T}^{old} [40].

3.2.3. Hierarchical a posteriori error estimation. As the concentration c is expected to strongly vary across phase boundaries, spatial adaptivity based on a posteriori error estimates is mandatory. Hierarchical error estimates rely on the solution of local defect problems. While originally introduced for linear elliptic problems [14, 17, 41, 60] this technique was successfully extended to nonlinear problems [3], constrained minimization [42, 44, 45, 56, 61] and nonsmooth saddle point problems [36, 31].

Thermomechanical stress is caused by different thermal expansion coefficients and the mismatch of the different constituents [19]. Hence, we assume that the accuracy of the finite element approximation (18c) is controlled by the resolution of the diffuse interface and thus concentrate on hierarchical error estimation of the phase field variables $c_{\mathcal{T}}^n$ and $w_{\mathcal{T}}^n$.

To this end, we note that the discrete spatial Cahn–Hilliard system (18a), (18b) is equivalent to a saddle point problem in $\mathcal{S}_j \times \mathcal{S}_j$ for a Lagrangian functional $\mathcal{L}_{\mathcal{T}}^n$ similar to \mathcal{L}^n . In fact, selecting $c^{n-1} = c^{\text{old}}$ and $\mathbf{u}^{n-1} = \mathbf{u}^{\text{old}}$, the Lagrangian $\mathcal{L}_{\mathcal{T}}^n$ is obtained from \mathcal{L}^n by replacing ψ_1 by the approximation $\psi_1^{\mathcal{T}}$. Following [31, 36], we now derive an a posteriori error estimate by suitable approximation of the defect problem associated with the defect Lagrangian

$$\mathcal{D}(e_c, e_w) = \mathcal{L}^n(c_{\mathcal{T}}^n + e_c, w_{\mathcal{T}}^n + e_w) - \psi_1(c_{\mathcal{T}}^n + e_c) + \psi_1^{\mathcal{T}}(c_{\mathcal{T}}^n + e_c).$$

In the first step the defect problem is discretized with respect to a larger finite element space $\mathcal{Q} \times \mathcal{Q}$, where $\mathcal{Q} = \mathcal{S}(\mathcal{T}')$ and \mathcal{T}' is obtained by uniform refinement of \mathcal{T} . Note that we have $\mathcal{Q} = \mathcal{S}_j \oplus \mathcal{V}$ with \mathcal{V} denoting the incremental space

$$\mathcal{V} = \text{span}\{\lambda_p^{\mathcal{T}'} \mid p \in \mathcal{E}'\}.$$

Here, $\mathcal{E}' = \mathcal{N}(\mathcal{T}') \setminus \mathcal{N}(\mathcal{T})$ is the set of non-hanging edge mid points in \mathcal{T} .

In the second step, the discrete defect problem is localized by ignoring the coupling between \mathcal{S}_j and \mathcal{V} and also the coupling between $\lambda_p^{\mathcal{T}'}$ for all $p \in \mathcal{E}'$. Denoting $\mathcal{D}_p(r, s) = \mathcal{D}(r\lambda_p^{\mathcal{T}'}, s\lambda_p^{\mathcal{T}'})$, this results in the local saddle point problems

$$(e_{c,p}, e_{w,p}) \in \mathbb{R}^2 : \quad \mathcal{D}_p(e_{c,p}, s) \leq \mathcal{D}_p(e_{c,p}, e_{w,p}) \leq \mathcal{D}_p(r, e_{w,p}) \quad \forall r, s \in \mathbb{R}$$

for all $p \in \mathcal{E}'$ that give rise to the hierarchical a posteriori error estimate

$$(19) \quad \eta = \left(\sum_{p \in \mathcal{E}'} \eta_p^2 \right)^{\frac{1}{2}}, \quad \eta_p^2 = \left\| e_{c,p} \lambda_p^{\mathcal{T}'} \right\|_c^2 + \left\| e_{w,p} \lambda_p^{\mathcal{T}'} \right\|_{w, \mathcal{T}'}^2, \quad p \in \mathcal{E}'$$

for the norms

$$\begin{aligned} \|c\|_c^2 &= (\Gamma(c^{\text{old}}) \nabla c, \nabla c) + \gamma_0 (c, 1)^2, \\ \|w\|_{w, \mathcal{T}}^2 &= \Delta t \left((M(c^{\text{old}}) \nabla(w), \nabla w) + \|w\|_{0, \mathcal{T}}^2 \right) \end{aligned}$$

and an averaged surface tension coefficient $\gamma_0 = \frac{1}{d} \sum_{i=1}^d \Gamma_{ii}(0)$.

After elimination of $e_{w,p}$, the local saddle point problems can be expressed in terms of scalar convex minimization problems which can be easily solved, e.g., by bisection. Numerical computations indicate efficiency and reliability of this error estimate [36], but theoretical justification is still open.

3.2.4. *Adaptive mesh refinement.* Adaptive mesh refinement based on a posteriori error estimation is carried out in two steps. In each time step, we first select a hierarchical coarse mesh. This mesh is intended to be coarse enough to allow for adaptive grids that are strongly varying in time and fine enough to make sure that relevant features of the solution enter the a posteriori error estimation.

To this end, we apply successive derefinement to the grid \mathcal{T}^{old} from the preceding time step. In the first time step, we chose a uniformly refined mesh \mathcal{T}^{old} which is sufficiently fine to resolve all relevant features of the initial value c^0 . In each of the m derefinement steps, we derefine all simplices e of \mathcal{T}^{old} that were obtained by more than minLevel refinements and satisfy either the condition (i) $|\nabla(I_e c^{\text{old}})|_e| < \text{Tol}_{\text{derefine}}$ or the condition (ii) $|\nabla(I_{e'} c^{\text{old}})|_{e'}| \geq \text{Tol}_{\text{derefine}}$ with e' chosen such that e is obtained by refinement of e' . Here, I_e and $I_{e'}$ denote the linear interpolation on e and e' , respectively. Note that in the case $|\nabla(I_e c^{\text{old}})|_e| > \text{Tol}_{\text{derefine}}$ the condition (ii) indicates that some strong variation is present and not overlooked after derefinement.

Adaptive mesh refinement of the resulting grid \mathcal{T} is based on the local error indicators η_p defined in (19). In each step, the indicators η_{p_i} , $i = 1, \dots, |\mathcal{E}'|$, are arranged with decreasing order, to determine the minimal number i_0 of indicators such that

$$(20) \quad \sum_{i=1}^{i_0} \eta_{p_i}^2 \geq \kappa \eta^2$$

holds with a given parameter $\kappa \in [0, 1]$. Then all simplices $e \in \mathcal{T}$ with the property $p_i \in e$ for some p_i with $i \leq i_0$ are marked for refinement [18]. Each marked simplex is partitioned by (red) refinement [6, 12]. Possible additional refinement is used to uniformly bound the ratio of diameters of neighboring simplices. The refinement process is terminated, if the estimated relative error is less than a given tolerance $\text{Tol}_{\text{adapt}} > 0$, i.e.,

$$(21) \quad \eta < \text{Tol}_{\text{adapt}} \cdot \left(\|c\|_c^2 + \|w\|_{w, \mathcal{T}}^2 \right)^{\frac{1}{2}}.$$

In all numerical experiments to be reported below, we selected the derefinement parameters $m = 2$, $\text{minLevel} = 6$, and $\text{Tol}_{\text{derefine}} = 2.0$, and the error tolerance $\text{Tol}_{\text{adapt}} = 0.1$ if not explicitly stated otherwise. In general we used $\kappa = 0.8$ for the refinement criterion (20). In order to avoid severe 'overrefinement' the parameter κ was reduced heuristically if η was already close to the desired tolerance in (21).

4. ALGEBRAIC SOLVERS

In this section we will discuss the efficient algebraic solution of the discrete problems $(\mathbf{CL}_{\Delta t, \mathcal{T}}^n)$ by iterative methods. In each time step this amounts in the solution of the nonsmooth nonlinear saddle point problem (18a)–(18b) and the linear equation (18c) in the quotient space \mathbf{H}_j .

4.1. **Nonsmooth Schur–Newton methods.** Assuming an ordering $\lambda_{p_1}^{\mathcal{T}}, \dots, \lambda_{p_N}^{\mathcal{T}}$ with $N = |\mathcal{N}(\mathcal{T})|$ of the basis of $\mathcal{S}(\mathcal{T})$ we can represent $c_{\mathcal{T}}^n$ and $w_{\mathcal{T}}^n$ by coefficient vectors $\underline{c}, \underline{w} \in \mathbb{R}^N$. Following the proof of Theorem 3.1 we will use an equivalent reformulation of (18a)–(18b) with $0 = \gamma_0 (c_{\mathcal{T}}^n - c^{\text{old}}, 1) (1, v - c_{\mathcal{T}}^n)$ added to (18b) resulting in the discrete variational inequality

$$(A\underline{c}) \cdot (\underline{v} - \underline{c}) - (M\underline{w}) \cdot (\underline{v} - \underline{c}) + \varphi(\underline{v}) - \varphi(\underline{c}) \geq \underline{f} \cdot (\underline{v} - \underline{c}) \quad \forall \underline{v} \in \mathbb{R}^N$$

with $A, M \in \mathbb{R}^{N \times N}$, $\underline{f} \in \mathbb{R}^N$, and the functional

$$\varphi(\underline{v}) = \sum_{i=1}^N \Psi_1(\underline{v}_i) \int_{\Omega} \lambda_{p_i}^{\mathcal{T}}(x) dx.$$

Utilizing the subdifferential $\partial\varphi : \mathbb{R}^N \rightarrow 2^{\mathbb{R}^N}$ of φ the whole system can be written as inclusion

$$(22) \quad \begin{pmatrix} A + \partial\varphi - M & \\ -M & -C \end{pmatrix} \begin{pmatrix} \underline{c} \\ \underline{w} \end{pmatrix} \ni \begin{pmatrix} \underline{f} \\ \underline{g} \end{pmatrix}$$

with symmetric positive semi-definite $C \in \mathbb{R}^{N \times N}$ and symmetric positive definite matrix $A \in \mathbb{R}^{N \times N}$. Notice that A is the sum of a sparse matrix of rank $(N - 1)$ and a dense matrix of rank 1. The inclusion (22) is called a saddle point problem since its solutions are saddle points of

$$L(\underline{c}, \underline{w}) = \frac{1}{2} A \underline{c} \cdot \underline{c} - \underline{f} \cdot \underline{c} + \varphi(\underline{c}) - M \underline{c} \cdot \underline{w} - \underline{g} \cdot \underline{w} - \frac{1}{2} C \underline{w} \cdot \underline{w}.$$

In case of a logarithmic potential (3), the inclusion (22) can be as well written as an equation, involving the derivative of Ψ . However, in light of its singularities at $c = 0, 1$ and desired robustness of iterative solution with respect to temperature θ , we concentrate on the more general formulation (22).

We now describe the iterative solution of the saddle point system (22) by so-called nonsmooth Schur–Newton multigrid methods, NSNMG methods in short. Here, it does not matter whether $\partial\varphi$ is set- or single-valued, because NSNMG relies on convexity rather than smoothness. While originally introduced for saddle point problems with obstacles [34, 35], NSNMG has been meanwhile extended to more general nonlinearities with nonsmooth convex energies [31, 32, 37]. The NSNMG approach relies on the equivalent minimization problem

$$\underline{w} \in \mathbb{R}^N : \quad h(\underline{w}) \leq h(\underline{v}) \quad \forall \underline{v} \in \mathbb{R}^N$$

with the dual energy functional

$$(23) \quad h(\underline{w}) = - \inf_{\underline{v} \in \mathbb{R}^N} L(\underline{v}, \underline{w}).$$

This equivalence was already used in the proofs of Theorems 3.1 and 3.4.

Now the main observation is that $h : \mathbb{R}^N \rightarrow \mathbb{R}$ is convex and differentiable with Lipschitz continuous derivative ∇h . Hence, the saddle point system (22) is equivalent to the equation

$$(24) \quad \nabla h(\underline{w}) = 0,$$

where the derivative ∇h is given by the nonlinear Schur-complement

$$\nabla h(\underline{w}) = M \left((A + \partial\varphi)^{-1} (\underline{f} + M \underline{w}) \right) + C \underline{w} + \underline{g}.$$

Lipschitz continuity of ∇h allows to apply Newton-like gradient related descent methods

$$(25) \quad \underline{w}^{\nu+1} = \underline{w}^{\nu} - \rho_{\nu} \left(\partial^2 h(\underline{w}^{\nu}) \right)^{-1} \nabla h(\underline{w}^{\nu})$$

with $\partial^2 h(\underline{w}^{\nu})$ being a generalized linearization of ∇h at \underline{w}^{ν} and ρ_{ν} a suitable damping parameter. Exploiting Lipschitz continuity and enforcing a chain rule, we obtain the linear Schur-complement

$$\partial^2 h(\underline{w}^{\nu}) = M A_{\nu}^{+} M + C$$

with A_ν^+ denoting the Moore–Penrose pseudoinverse of the truncated matrix

$$(A_\nu)_{ij} = \begin{cases} A_{ij} & \text{if } i \neq j \text{ and } i, j \in \mathcal{I}_\nu, \\ A_{ii} + \Psi_1''(\underline{c}_i^\nu) & \text{if } i \in \mathcal{I}_\nu, \\ 0 & \text{else.} \end{cases}$$

The inactive set $\mathcal{I}_\nu = \{i \mid \Psi_1 \text{ is twice differentiable at } \underline{c}_i^\nu \text{ and } \Psi_1''(\underline{c}_i^\nu) < c_{\max}\}$ is determined by the associated primal iterate \underline{c}^ν ,

$$(26) \quad \underline{c}^\nu = (A + \partial\varphi)^{-1}(f + M\underline{w}^\nu).$$

Here $c_{\max} > 0$ is a large constant meant to avoid unbounded diagonal elements of A_ν . Assuming existence and uniqueness of the solution $w_\mathcal{T}^\nu$, it was shown in [31, 32] that the resulting algorithm is globally convergent, if the damping parameters ρ_ν are properly chosen, e.g., by bisection or by the Armijo rule. Global convergence is preserved by inexact evaluation of the directions $(\partial^2 h(\underline{w}^\nu))^{-1} \nabla h(\underline{w})$ with increasing accuracy.

Each NSNMG iteration is stopped, if the norm of the actual correction of the dual iterate falls below a given threshold, i.e.,

$$(27) \quad \|\underline{w}^{\nu+1} - \underline{w}^\nu\|_{\underline{w}, \mathcal{T}} \leq \text{Tol}_{\text{NSNMG}}.$$

In the numerical experiments reported below $\text{Tol}_{\text{NSNMG}} = 10^{-12}$ was chosen.

Each iteration step of NSNMG requires the (approximate) solution of the nonlinear Allen–Cahn-type problem (26) and the linear system (25) for the Schur complement $MA_\nu^+M + C$ in order to obtain $\nabla h(\underline{w}^\nu)$ and the new iterate $\underline{w}^{\nu+1}$, respectively. In our numerical computations to be reported below, the nonlinear Allen–Cahn-type problem (26) is solved in an efficient and robust way by V(3,3) cycles of *truncated nonsmooth Newton multigrid* (TNNMG), combining nonlinear relaxation methods with nonsmooth Newton techniques [31, 33, 39]. The linear system (25) is equivalent to a linear saddle point problem and can be solved by a multitude of direct or iterative solvers. We used the GMRES method [54] preconditioned with a multigrid method with block Gauß–Seidel smoother [55, 58, 62] in our numerical computations.

4.2. A multigrid method for singular elasticity problems. In order to describe the iterative solution of the elasticity problem (18c), we first consider the related problem

$$(28) \quad \tilde{\mathbf{u}}_\mathcal{T}^n \in \mathcal{S}_j^d : \quad a(\tilde{\mathbf{u}}_\mathcal{T}^n, \mathbf{v}) = l(\mathbf{v}) \quad \forall \mathbf{v} \in \mathcal{S}_j^d$$

with symmetric positive semi-definite bilinear form and linear functional given by

$$a(\cdot, \cdot) = (\mathcal{C}(c_\mathcal{T}^n)\varepsilon(\cdot), \varepsilon(\cdot)), \quad l(\cdot) = \int_{\partial\Omega} \mathbf{g} \cdot (\cdot) ds + (\mathcal{C}(c_\mathcal{T}^n)\bar{\varepsilon}(c_\mathcal{T}^n), \varepsilon(\cdot)).$$

Without loss of generality we assume that \mathbf{g} satisfies the compatibility condition

$$\int_{\partial\Omega} \mathbf{g} \cdot \mathbf{v} ds = 0 \quad \forall \mathbf{v} \in \ker(\varepsilon).$$

Then the solutions space of (28) is given by $\mathbf{u}_\mathcal{T}^n + \ker(\varepsilon)$.

For the solution of (28) we can now use a classical linear multigrid method with linear Gauß–Seidel smoother with respect to the hierarchy $\mathcal{S}_0^d \subset \dots \subset \mathcal{S}_j^d$ of subspaces, as introduced in (17). We emphasize that the Gauß–Seidel smoother is well-defined for all levels k because the vector valued nodal basis functions of \mathcal{S}_k^d are not contained in the kernel $\ker(\varepsilon)$ of the bilinear form a .

It is easy to see that this multigrid method converges to the solution $\mathbf{u}_\mathcal{T}^n$ with respect to the half-norm $a(\cdot, \cdot)^{1/2}$. For the Poisson problem with Neumann boundary conditions discretized with respect to a hierarchy of quasi-uniform grids, mesh

independence of the convergence rates was shown in [46]. Mesh independence for classical multigrid applied to the present singular elasticity problem (28) can be shown using the same arguments.

Notice that only the projections of the iterates to \mathbf{H} converge to \mathbf{u}_T^n with respect to the $H^1(\Omega)^d$ -norm. This is due to the non-uniqueness of rigid body motions of solutions to (28).

5. QUANTIFICATION OF COARSENING

In order to assess the coarsening of microstructure in binary alloys we need to introduce a characteristic length scale of a given grain distribution. It is well-known that for the Cahn–Hilliard equation (without elasticity) such a length scale – in a time averaged sense – cannot grow faster than $t^{1/3}$ (cf. [43, 16, 52, 53]). Obviously, there are no corresponding global lower bounds, as there are stable states that do not coarsen at all. While the inverse of the Ginzburg–Landau energy has emerged as a convenient length scale in analysis and numerical computations (cf. [16, 25, 43, 52, 53, 59]), it suffers from inaccessibility in physical experiments. Hence, we choose the so-called *mean phase radius* \bar{r} (i.e. mean radius of precipitates) as characteristic length scale to maintain quantitative comparability with material science experiments as for example in [9]. In order to determine \bar{r} we follow standard procedure of quantitative stereology performing a lineal analysis (cf. [57]). From this we can approximate the *mean intercept length* \bar{L} which in turn can be translated into the mean phase radius assuming spherical phase shape. The details of the applied procedure are given below.

For the following let's call the matrix and particle (or precipitate) phases β - resp. α -phase and assume that the phase field value c_α of the α -phase is greater than that of β -phase, c_β . In this context let Θ denote the set of secants of α -precipitates and $L(\omega)$, $\omega \in \Theta$ their lengths. The mean intercept length is then

$$\bar{L} = E(L)$$

the expected value of secant length assuming uniform distribution on Θ . In order to approximate \bar{L} we lay an equidistant cube mesh with mesh size h over the computational domain and count the number of mesh edges $N_\mathcal{E}$ and the number of line segments N_L intersecting an α -precipitate. Partially intersecting edges are counted as one half. Thus we have as an approximation

$$\bar{L} \approx N_\mathcal{E} \cdot h / N_L.$$

If no particles of diameter less than $\sqrt{2}h$ are present, counting the intersecting edges amounts to counting grid vertices inside α -regions and multiplying by the dimension of the computational domain (note that smaller particles might be overlooked this way, cf. fig. 1a). A vertex is counted as inside α -phase iff the value of the phase field at its coordinates is above a given threshold c_0 . As intersecting line segments we count nonempty α -vertex sets which are discretely connected along mesh lines (cf. fig. 1b). Under the assumption that the minimal distance of two neighboring precipitates is larger than h this is exactly the number N_L .

The mean intercept length of a single spherical particle of radius r is $\nu_d \cdot r$ where $\nu_d = \frac{\pi}{2}$ in 2D and $\nu_d = \frac{4}{3}$ in 3D. Hence the mean phase radius of conglomerates of particles – under the assumption of spherical phase shape – is given by

$$\bar{r} = \frac{\bar{L}}{\nu_d}.$$

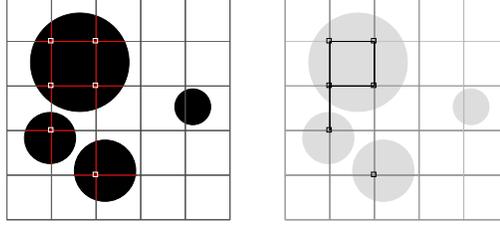


FIGURE 1. Illustration of our lineal analysis. In this example $N_{\mathcal{E}} = 12$; $N_L = 7$

6. COARSENING OF MICROSTRUCTURE IN A EUTECTIC AgCu BRAZING ALLOY

In this section we simulate the microstructure evolution in a eutectic silver/copper (Ag71Cu29) brazing alloy utilizing the Cahn–Larché model (9) and its discretization (18).

6.1. Problem setting. The computations are carried out in the square $\Omega = [-L, L]^2$ with edge length $2L = 0.1 \mu\text{m}$ during the time interval from zero to $T = 375 \text{ s} = 6.25 \text{ min}$. We choose c to be the copper concentration so that $c = 0$ corresponds to pure silver and $c = 1$ corresponds to pure copper. Concerning the setting and material parameters, we closely follow [10], i.e., for the alloy in question the eigenstrains are assumed to result only from thermal expansion: $\bar{\varepsilon}(c) = \Delta\theta\mathcal{A}(c)$. In our computations we chose $\Delta\theta = 1000 \text{ K}$. The surface tension, mobility, and thermal expansion then reduce to scalar functions

$$\Gamma(c) = \gamma(c)\text{Id}, \quad M(c) = m(c)\text{Id}, \quad \mathcal{A}(c) = a(c)\text{Id}.$$

The corresponding quantities for the pure constituents, i.e. for $c = 0, 1$, are given in Table 1 (note that due to a different scaling in our equations, we need to rescale γ by a factor of 2 as compared to [10]). The entries of the Hooke tensors \mathcal{C}_{Ag} and \mathcal{C}_{Cu} for pure silver and copper, respectively, are given in Table 2. As in (2), the values of the functions γ , m , a , and \mathcal{C} at $c \in (0, 1)$ are obtained by linear interpolation.

$\gamma_{\text{Ag}}[\text{N}]$	$\gamma_{\text{Cu}}[\text{N}]$	$m_{\text{Ag}}[\frac{\text{m}^5}{\text{Js}}]$	$m_{\text{Cu}}[\frac{\text{m}^5}{\text{Js}}]$	$a_{\text{Ag}}[\frac{10^6}{\text{K}}]$	$a_{\text{Cu}}[\frac{10^6}{\text{K}}]$
$3.06 \cdot 10^{-10}$	$3.808 \cdot 10^{-10}$	$7.25 \cdot 10^{-25}$	$3.65 \cdot 10^{-25}$	18.9	16.5

TABLE 1. Material parameters (taken from [10])

$\mathcal{C}_{\text{Ag}}^{11}[\text{GPa}]$	$\mathcal{C}_{\text{Ag}}^{12}[\text{GPa}]$	$\mathcal{C}_{\text{Ag}}^{44}[\text{GPa}]$	$\mathcal{C}_{\text{Cu}}^{11}[\text{GPa}]$	$\mathcal{C}_{\text{Cu}}^{12}[\text{GPa}]$	$\mathcal{C}_{\text{Cu}}^{44}[\text{GPa}]$
168	121	75	124	94	46

TABLE 2. Entries of the Hooke tensors for pure silver and copper in Voigt notation (taken from [10])

The chemical energy density Ψ takes the form (3) with parameters β_i given in Table 3. The splitting (12) is chosen according to

$$\begin{aligned} \Psi_1(c) &= \beta_0 R\theta (c \log(c) + (1-c) \log(1-c)) \\ &\quad + (\beta_4 - \beta_3)c^3 + (\beta_3 - \beta_4)c^2 + (\beta_2 - \beta_1 + \beta_4)c + \beta_1 \\ \Psi_2(c) &= -\beta_4 c^2. \end{aligned}$$

$\beta_0[\frac{\text{mol}}{\text{m}^3}]$	$\beta_1[\frac{\text{GJ}}{\text{m}^3}]$	$\beta_2[\frac{\text{GJ}}{\text{m}^3}]$	$\beta_3[\frac{\text{GJ}}{\text{m}^3}]$	$\beta_4[\frac{\text{GJ}}{\text{m}^3}]$
$1.11248134 \cdot 10^5$	-5.20027	-7.2738	2.96683	3.01417

TABLE 3. Fitting parameters for Margules ansatz at $\theta = 1000$ K and given material data (taken from [10])

In our computations we chose L as unit length, $\Psi_0 = 0.1 \frac{\text{GJ}}{\text{m}^3}$ as unit energy density, and $10^{-25} \frac{\text{m}^5}{\text{Js}}$ as unit mobility resulting in a unit time of $t_0 = 250$ s. Numerical tests suggest the choice $\Delta t = 7.5 \cdot 10^{-4} t_0 = 0.210^{-3} T$ of the time step. The coarsest grid \mathcal{T}_0 consists of a partition of Ω into two congruent subtriangles. In the first time step, we start the derefinement process described in Section 3.2.4 from \mathcal{T}^{old} obtained by 8 uniform refinements of \mathcal{T}_0 .

6.2. Evolution of concentration. In our first simulation, we apply no boundary stress and select the initial condition c^0 as shown in the upper left picture of Figure 2. We chose these default data in the sequel, if not otherwise stated. The colors blue and red indicate high concentrations of silver and copper, respectively. The remaining pictures in Figure 2 illustrate the evolution of the approximate concentration $c_{\mathcal{T}}^n$ together with the corresponding final grids of the adaptive procedure over various time steps $n = 50, \dots, 2000$. Observe that the coarsening significantly slows down during the evolution (see Subsection 6.3 for details).

Mass conservation of the constituents is a key feature of the physical process which should be preserved in numerical simulations. In our computation of over 2000 time steps, we found the maximal relative deviation from the initial mass of copper of

$$\max_{n=1, \dots, 2000} \frac{|\int_{\Omega} c_{\mathcal{T}}^n - \int_{\Omega} c^0|}{\int_{\Omega} c^0} \approx 1.9 \cdot 10^{-11}$$

The equilibrium concentrations

$$c_{\alpha} = 0.05096976816135458 \quad \text{and} \quad c_{\beta} = 0.9460270077128279$$

of the Ag-rich α - and the Cu-rich β -phase, respectively, are determined by the Maxwell-tangent construction (see e.g. [49]). In our computations the phase equilibria are recovered up to about $4 \cdot 10^{-4}$. This is illustrated by Figure 3 showing the cross section of the initial condition (black) and of the approximate solution $c_{\mathcal{T}}^{2000}$ along the y -axis.

In order to study the influence of thermomechanical stress on the evolution of the phase field we applied boundary stress of the form

$$(29) \quad \begin{aligned} & \mathbf{g} = -g\mathbf{n} \quad \text{on } \partial\Omega \\ & g(x) = g_0 \text{ GPa, if } x = (x_1, \pm L) \quad \text{and} \quad g(x) = 0 \text{ GPa, otherwise} \end{aligned}$$

with the different values $g_0 = 0, \dots, 20$ GPa, but observed only minor changes in the evolution (see also Subsection 6.3). In this and the following experiment the error tolerance $\text{Tol}_{\text{adapt}} = 0.05$ was chosen.

In order to (unphysically!) enhance the influence of thermomechanical stress, the elastic energy density $\mathcal{W}(c, \varepsilon(\mathbf{u}))$ is multiplied by a factor of $\omega = 1, 10, 100, 1000$ while zero boundary stress is prescribed. This leads to a significantly faster dynamics and oblong phase shapes oriented along the coordinate directions. This is illustrated in Figure 4 that shows the initial condition c^0 and the approximate concentration $c_{\mathcal{T}}^n$ together with the underlying grids for $\omega = 1000$ and the time steps $n = 40, 60, 80, 100, 150$.

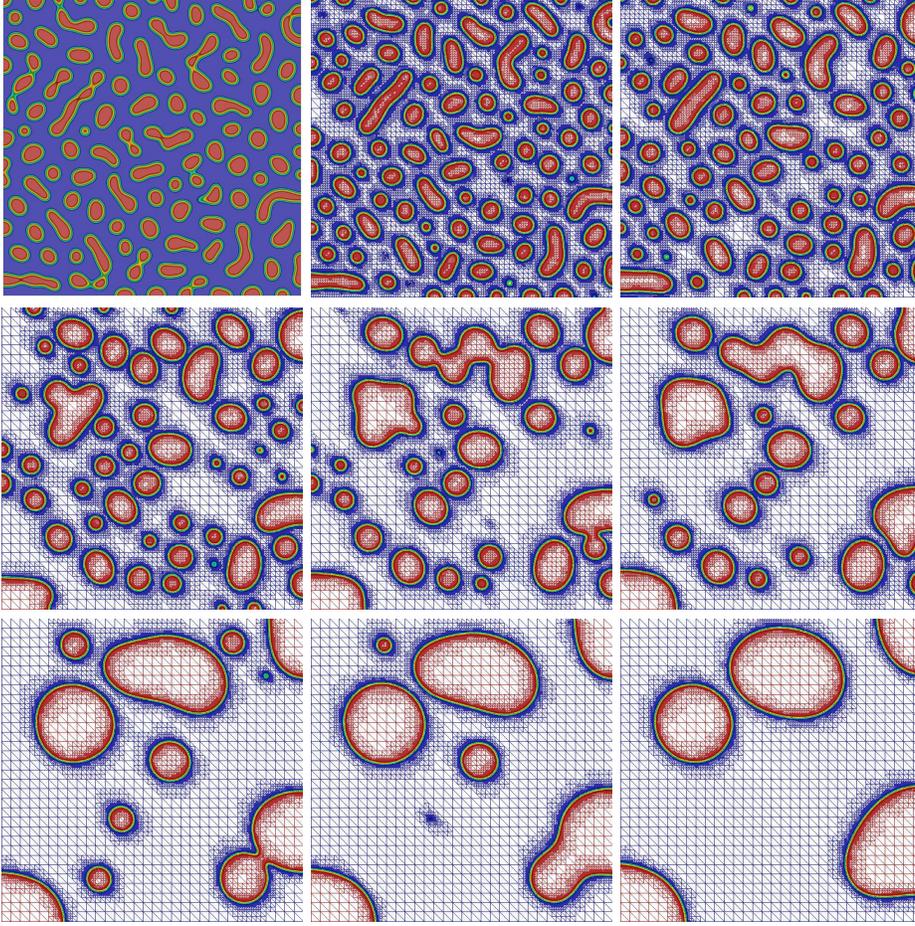


FIGURE 2. Initial value c^0 and approximation c_T^n at time steps $n = 50, 100, 200, 300, 400, 800, 1000, 2000$.

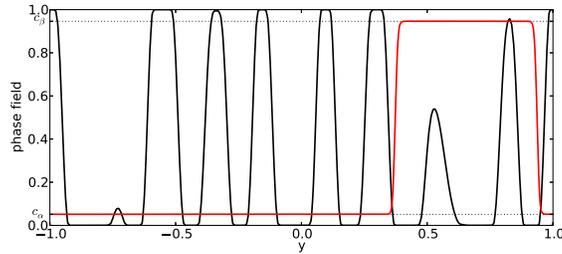


FIGURE 3. Cross section of the initial condition (black) and of the approximate solution c_T^{2000} along the y -axis. The dotted lines represent the equilibrium concentrations c_α and c_β .

6.3. Evolution of mean phase radii. In this subsection, we investigate the dynamics of coarsening in terms of the evolution of the mean phase radii in more detail.

In our first experiment we investigate the sensitivity of the evolution of mean phase radii with respect to a smooth approximation of the logarithmic-type Margules potential Ψ as described in (3) with parameters β_i given in Table 3. To this

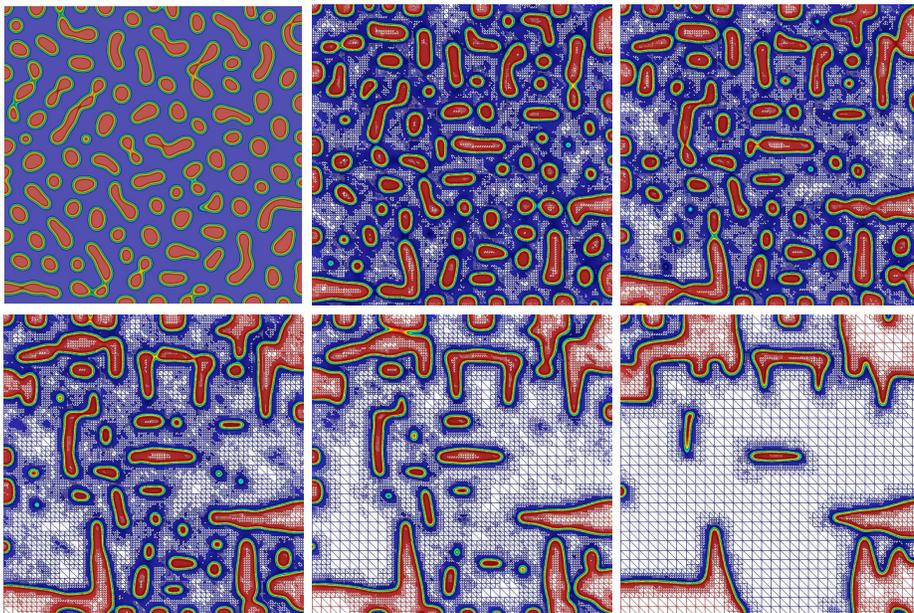


FIGURE 4. Initial value c^0 and approximation c^n at time steps $n = 40, 60, 80, 100, 150$ for (unphysically!) increased elastic energy density $\mathcal{W}(c, \varepsilon(\mathbf{u}))$.

end, we consider the quartic Hermite interpolation $P^\Psi(c) = \sum_{i=0}^4 \alpha_i c^i$ of Ψ at the equilibrium concentrations c_α, c_β , and at the eutectic point c_{eut} , characterized by

$$\begin{aligned} P^\Psi(c_\alpha) &= \Psi(c_\alpha), & P^\Psi(c_\beta) &= \Psi(c_\beta), & P^\Psi(c_{\text{eut}}) &= \Psi(c_{\text{eut}}) \\ (P^\Psi)'(c_\alpha) &= \Psi'(c_\alpha), & (P^\Psi)'(c_\beta) &= \Psi'(c_\beta). \end{aligned}$$

The left and the right picture in Figure 5 show Ψ (solid), P^Ψ (dashed), and

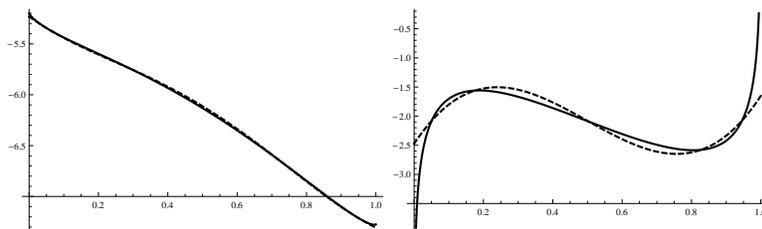


FIGURE 5. Left: Margules potential Ψ (solid), quartic approximation P^Ψ (dashed). Right: First derivatives.

their derivatives, respectively. The splitting (12) of P^Ψ is selected according to $P_1^\Psi = \sum_{\substack{i=0 \\ i \neq 3}}^4 \alpha_i c^i$ and $P_2^\Psi = \alpha_3 c^3$.

We prescribe zero boundary stress (see Figure 2 for the corresponding evolution of the approximate concentrations). Now Figure 6 shows the evolution of the mean phase radii for the Margules potential Ψ (black) and its quartic approximation P^Ψ (red), respectively. It turns out that quartic approximation strongly perturbs the coarsening behavior which is in agreement with the engineering literature [19, 20].

In our next experiment, we investigate the influence of strongly varying boundary pressure $g_0 = 0, 1, 2, 3, 4, 5, 10, 20$ GPa in the boundary condition (29). Figure 7

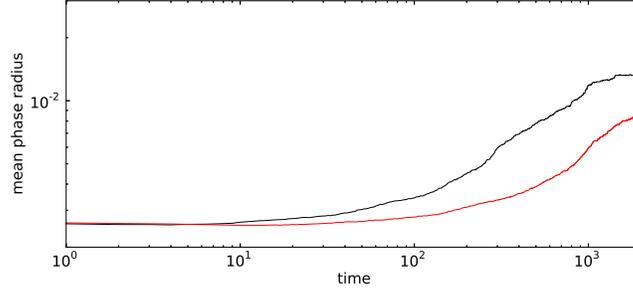


FIGURE 6. Approximate evolution of mean phase radii for Margules potential Ψ (black) and its quartic approximation P^Ψ (red).

shows that even large mechanical stress has only minor influence on the coarsening behavior. This effect has also been observed in previous simulations [20, Section 4.3].

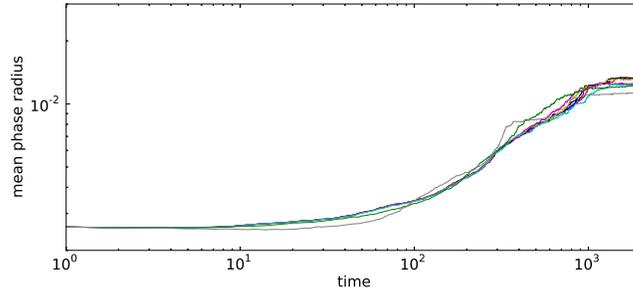


FIGURE 7. Approximate evolution of mean phase radii for boundary pressure $g_0 = 0$ GPa (red), 1 GPa (blue), 2 GPa (black), 3 GPa (yellow), 4 GPa (magenta), 5 GPa (cyan), 10 GPa (green), and 20 GPa (gray).

For a qualitative investigation of mechanically induced coarsening, the elastic energy density $\mathcal{W}(c, \varepsilon(\mathbf{u}))$ is replaced by $\omega \mathcal{W}(c, \varepsilon(\mathbf{u}))$ with (unphysical!) amplification factors $\omega = 1, 10, 100, 1000$. Figure 8 shows that the coarsening speed increases with increasing ω , as expected (recall the evolution of concentrations for $\omega = 1000$ depicted in Figure 4).

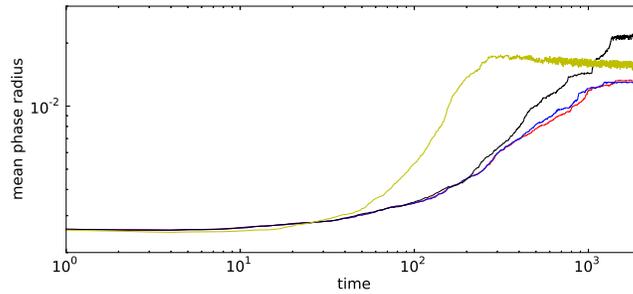


FIGURE 8. Approximate evolution of mean phase radii for (unphysically!) amplified elastic energy by the factor $\omega = 1$ (red), 10 (blue), 100 (black), and 1000 (yellow).

6.4. Numerical aspects. We now briefly illustrate the performance of the main building blocks of our numerical solution algorithm. For more detailed numerical experiments, we refer to [31, 32, 34, 36].

We first consider a posteriori error estimation and adaptive mesh refinement as described in Subsections 3.2.3 and 3.2.4, respectively. The corresponding adaptive algorithm is applied to the spatial problem arising in the $n = 1^{\text{st}}$ time step of the discretized Cahn–Larché system (16) with material data and discretization parameters given in Subsection 6.1. The grid \mathcal{T} satisfying the stopping criterion (21)

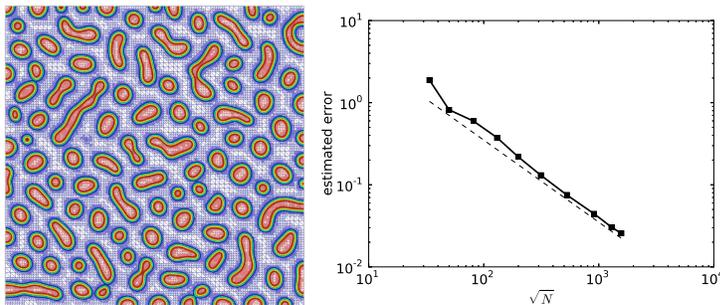


FIGURE 9. Left: Adaptively refined grid \mathcal{T} . Right: Estimated error η over $N^{1/2}$ (solid) in comparison with $\mathcal{O}(N^{-1/2})$ (dashed).

with $\text{Tol}_{\text{adapt}} = 0.03$ after 9 adaptive refinement steps is shown in the left picture of Figure 9. Observe how the initial, 8 times uniformly refined grid \mathcal{T}^{old} has been adaptively coarsened and then refined according to the new approximation in the new time step. The right picture shows the corresponding estimated error η , as introduced in (19), over \sqrt{N} , N denoting the corresponding number of unknowns. Note that $h = N^{-1/2}$ is the mesh size in case of uniform refinement. The dashed line indicates $\mathcal{O}(N^{-1/2})$. A comparison suggests that our adaptive refinement algorithm provides approximations with optimal order.

Using the same problem as above, we now illustrate the iterative solution of the discretized phase field system (18a), (18b). On each computational grid the overall iteration is stopped once the termination criterion (27) with $\text{Tol}_{\text{NSNMG}} = 10^{-12}$ is matched. The initial iterate is selected as the final iterate from the preceding refinement step (nested iteration). The resulting number of *Nonsmooth Schur–Newton Multigrid* (NSNMG) iterations required to reach this tolerance ranges from 5 on the coarser grids to 3 on the finer grids.

Recall that each step of NSNMG is quite expensive (cf. Subsection 4.1): It involves the approximate solution of the discrete Allen–Cahn-type system (26) by V(3,3) cycles of *truncated nonsmooth Newton multigrid* (TNNMG) and of the linear saddle point problem (25) by a GMRES method with a multigrid preconditioner with block Gauß–Seidel smoother. In order to illustrate the share of the numerical solution of each of these two subproblems in the computational effort of an NSNMG iteration step, we consider the discrete spatial problems occurring in the first time step after $j = 0, \dots, 9$ adaptive refinement cycles. This leads to a minimal mesh size $\sqrt{2} \cdot 2^{-11}$ and 2739799 unknowns on the final level.

For a fair comparison, the required computational work is measured in work units rather than iteration steps. One work unit is chosen to be the cpu time for one V(3,3) cycle of TNNMG on the corresponding grid. While the sum of work units for all subproblems on each refinement level is ranging from 12 to 20 for the discrete Allen–Cahn-type system (TNNMG), it reaches values from 647 to 2784 for the linear saddle point problem (preconditioned GMRES). Similarly, for TNNMG

the average error reduction over all subproblems occurring on each refinement level $j = 1, \dots, 9$ never exceeds 0.05, but even reaches values of 0.98 for preconditioned GMRES on finer grids.

Thus, the overall computational work is obviously dominated by the *linear* saddle point solver. A first, simple, reason is that the linear saddle point problem (25) is larger: It involves twice the number of unknowns of the discrete Allen–Cahn-type system (26). Another reason is that an equivalent reformulation of the discrete Allen–Cahn-type system in terms of convex minimization could be directly exploited in the algebraic solution process. This is not the case for linear saddle point problems.

We finally consider the (indefinite) linear elasticity problem (18c). Figure 10 shows the average error reduction ρ^k per iteration step of the *Quotient Space Multi-grid method* (QMG) described in Subsection 4.2 for a Dirichlet problem of linear elasticity (triangles) and the corresponding Neumann problem (circles) in 2D (solid) and 3D (dashed). The iteration is stopped once the estimated relative accuracy of 10^{-10} is reached and the initial iterates are obtained by nested iteration. The average error reduction ρ^k seems to saturate with increasing refinement level. This is in perfect agreement with theoretical considerations (cf., e.g., [46]).

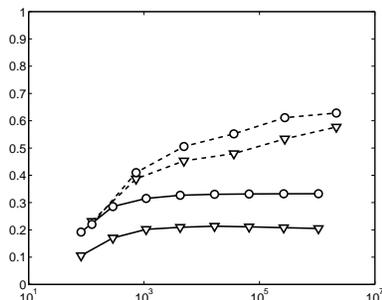


FIGURE 10. Averaged error reduction per iteration step of QMG over number of unknowns for a Dirichlet problem of linear elasticity (triangles) and the corresponding Neumann problem (circles) in 2D (solid) and 3D (dashed)

7. APPENDIX

7.1. Existence and uniqueness of time-discrete solutions. In this section we prove Theorem 3.1. The following continuity result will be helpful.

Lemma 7.1. *Let $z \in L^1(\Omega)$. Then the functional $g(v) = (z, v)$ is continuous on each $L^\infty(\Omega)$ -bounded subset of $L^p(\Omega)$, $1 \leq p \leq \infty$.*

Proof. For $p = \infty$ the assertion follows from Hölders inequality. Let $1 \leq p < \infty$. Consider some $U \subset L^p(\Omega)$ such that there is $r > 0$ with $|v(x)| \leq r$ a.e. in Ω for all $v \in U$. We define the function $f : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ according to

$$f(x, v) = \begin{cases} z(x)v, & \text{if } |v| \leq r, \\ z(x)r, & \text{if } v > r, \\ -z(x)r, & \text{if } v < -r. \end{cases}$$

Then the corresponding superposition operator F , given by $(F(v))(x) = f(x, v(x))$, satisfies $F(v) = zv$ for all $v \in U$. Moreover, $|F(v)| \leq r|z|$ holds for all $v \in L^p(\Omega)$ and therefore $F : L^p(\Omega) \rightarrow L^1(\Omega)$. As $f(x, \cdot)$ is continuous on \mathbb{R} for all $x \in \Omega$ and $f(\cdot, v)$

is measurable on Ω for all $v \in \mathbb{R}$ [2, Theorem 3.7] implies that $F : L^p(\Omega) \rightarrow L^1(\Omega)$ is even continuous. Hence $U \ni v \mapsto \int_{\Omega} F(v) dx = (z, v)$ is continuous from U to \mathbb{R} with respect to $\|\cdot\|_{L^p(\Omega)}$. \square

Note that the linear map is Gâteaux differentiable on bounded functions but its Gâteaux derivative $g'(v) = g$ is in general not continuous on this space.

Lemma 7.2. *Let $c^{n-1} \in \mathcal{K} = \{v \in H^1(\Omega) \mid v(x) \in [0, 1] \text{ a.e.}\}$ and $\mathbf{u}^{n-1} \in X$. Then the functional $\mathcal{J}^n : H^1(\Omega) \rightarrow \mathbb{R} \cup \{\infty\}$ given by*

$$\mathcal{J}^n(c) = \int_{\Omega} \frac{1}{2} \Gamma(c^{n-1}) \nabla c \cdot \nabla c dx + \frac{\gamma}{2} (c - c^{n-1}, 1)^2 + \psi_1(c) - (R^{n-1}, c)$$

with $R^{n-1} = R(c^{n-1}, \mathbf{u}^{n-1})$ is proper, strongly convex, and lower semi-continuous on $H^1(\Omega)$.

Proof. Utilizing the assumptions (A1), (A2) on Γ , the Poincaré inequality implies that the two quadratic terms in \mathcal{J}^n are strongly convex and continuous on $H^1(\Omega)$. Furthermore ψ_1 is convex, proper, and lower semi-continuous on $H^1(\Omega)$ (see e.g. [31, Lemma 3.5]).¹

It remains to show that the linear functional (R^{n-1}, \cdot) is lower semi-continuous. To this end, first note that $c^{n-1}(x) \in [0, 1]$ a.e. in Ω together with smoothness of Ψ_2 implies $\Psi_2'(c^{n-1}) \in L^\infty(\Omega)$. Utilizing the boundedness of the coefficient functions occurring in R^{n-1} , $c^{n-1} \in H^1(\Omega)$, and $\mathbf{u}^{n-1} \in H^1(\Omega)^d$, we get

$$z = \frac{1}{2} (\nabla c^{n-1})^T \Gamma'(c^{n-1}) \nabla c^{n-1} + \frac{1}{2} \varepsilon(\mathbf{u}^{n-1}) : \mathcal{C}'(c^{n-1}) \varepsilon(\mathbf{u}^{n-1}) \in L^1(\Omega)$$

and all other terms are in $L^2(\Omega)$. Lemma 7.1 implies that $v \mapsto g(v) := (z, v)$ is continuous on $\mathcal{K} = \text{dom}(\mathcal{J}^n)$ with respect to $\|\cdot\|_{L^2(\Omega)}$, and all the more with respect to $\|\cdot\|_{H^1(\Omega)}$. Hence, the extension of g by infinity is lower semi-continuous on $H^1(\Omega)$. Thus \mathcal{J}^n is lower semi-continuous on $H^1(\Omega)$. \square

Note that strong convexity implies strict convexity and coercivity.

Proof of Theorem 3.1. To show existence of (16a)–(16b) we can proceed as in [31, Theorem 3.8]: First we note that these equations are equivalent to a saddle point problem for the associated Lagrangian functional

$$\mathcal{L}^n(c, w) = \mathcal{J}^n(c) - (c - c^{n-1}, w)^2 - \frac{\Delta t}{2} (M(c^{n-1}) \nabla w, \nabla w).$$

Note that the additional integral term $(c - c^{n-1}, 1)$ in \mathcal{J}^n vanishes if (16a) is satisfied. While $\mathcal{L}^n(c, \cdot)$ is trivially concave and upper semi-continuous, Lemma 7.2 provides convexity, coercivity, and lower semi-continuity of $\mathcal{L}^n(\cdot, w)$. Now existence follows from [21, Chapter VI, Proposition 2.4], if the dual functional

$$h(w) = - \inf_{v \in \mathcal{K}} \mathcal{L}^n(v, w)$$

is coercive on $H^1(\Omega)$. This can be shown as in [31, Theorem 3.8] by proving that

$$h(w) \geq -\mathcal{L}^n(c(w), w) \geq C\|w\|_{H^1(\Omega)} - C$$

holds with $c(w) = (1 + \text{sgn}(w, 1))/2 = \text{const} \in \{0, 0.5, 1\}$.

In order to prove uniqueness, assume that (c_1^n, w_1^n) and (c_2^n, w_2^n) are two solutions. Then testing (16a) for (c_i^n, w_i^n) with $w_i^n - w_j^n$, $j \neq i$, and adding the equations yields

$$(30) \quad (w_2^n - w_1^n, c_1^n - c_2^n) = \Delta t (M(c^{n-1}) \nabla(w_1^n - w_2^n), \nabla(w_1^n - w_2^n)).$$

¹While the proof in [31] is given for a special case of Ψ_1 it directly carries over to the class of convex functions considered here.

Similarly testing (16b) for (c_i^n, w_i^n) with $c_j^n, j \neq i$, yields

$$(31) \quad (\Gamma(c_1^{n-1} \nabla(c_1^n - c_2^n), \nabla(c_1^n - c_2^n)) + (w_2^n - w_1^n, c_1^n - c_2^n) \leq 0.$$

Inserting (30) into (31) provides uniqueness on ∇c^n and ∇w^n . Testing (16a) with $v = 1 = \text{const}$ finally provides uniqueness of $(c^n, 1)$ and therefore of c^n .

For the remaining problem (16c) the assumptions (A1)–(A3) on the coefficient functions ensure that the right hand side is in $H^1(\Omega)'$ and therefore the bilinear form $(\mathcal{C}(c^n)\varepsilon(\cdot), \varepsilon(\cdot))$ is $(\varepsilon(\cdot), \varepsilon(\cdot))$ -elliptic. Now Korn's inequality (see, e.g., [51, Theorem 3.5]) provides $H^1(\Omega)^d$ -ellipticity on the quotient space \mathbf{H} and thus the existence of a unique solution $\mathbf{u}^n \in \mathbf{H}$. \square

REFERENCES

- [1] D. Anders and K. Weinberg. Numerical simulation of diffusion induced phase separation and coarsening in binary alloys. *Comp. Mat. Sci.*, pages 1359–1364, 2011.
- [2] J. Appell and P. P. Zabrejko. *Nonlinear Superposition Operators*. Number 95 in Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1990.
- [3] R. E. Bank and R. K. Smith. A posteriori error estimates based on hierarchical bases. *SIAM J. Numer. Anal.*, 30:921–935, 1993.
- [4] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöforn, R. Kornhuber, M. Ohlberger, and O. Sander. A generic interface for adaptive and parallel scientific computing. Part II: Implementation and tests in DUNE. *Computing*, 82(2–3):121–138, 2008.
- [5] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöforn, M. Ohlberger, and O. Sander. A generic interface for adaptive and parallel scientific computing. Part I: Abstract framework. *Computing*, 82(2–3):103–119, 2008.
- [6] J. Bey. Simplicial grid refinement: On Freudenthal's algorithm and the optimal number of congruence classes. *Numer. Math.*, 85:1–29, 2000.
- [7] L. Blank, H. Garcke, L. Sarbu, and V. Styles. Primal–dual active set methods for Allen–Cahn variational inequalities with nonlocal constraints. *Numer. Methods Partial Differential Equations*, 29:99–1030, 2013.
- [8] J. F. Blowey and C. M. Elliott. The Cahn–Hilliard gradient theory for phase separation with non-smooth free energy II: Numerical analysis. *Euro. J. Appl. Math.*, 3:147–179, 1992.
- [9] T. Böhme. *Investigations of Microstructural Changes in Lead-Free Solder Alloys by Means of Phase Field Theories*. PhD thesis, Technische Universität Berlin, 2008.
- [10] T. Böhme, W. Dreyer, F. Duderstadt, and W. H. Müller. A higher gradient theory of mixtures for multi-component materials with numerical examples for binary alloys. Preprint 1286, WIAS Berlin, 2007.
- [11] E. Bonetti, P. Colli, W. Dreyer, G. Gilardi, G. Schimperna, and J. Sprekels. On mathematical models for phase separation in elastically stressed solids. *Physica D*, pages 48–65, 2002.
- [12] F. Bornemann, B. Erdmann, and R. Kornhuber. Adaptive multilevel methods in three space dimensions. *Int. J. Numer. Meth. Engrg.*, 36:3187–3203, 1993.
- [13] F. A. Bornemann. *An Adaptive Multilevel Approach for Parabolic Equations in Two Space Dimensions*. PhD thesis, Freie Universität Berlin, 1991.
- [14] F. A. Bornemann, B. Erdmann, and R. Kornhuber. A posteriori error estimates for elliptic problems in two and three space dimensions. *SIAM J. Numer. Anal.*, 33:1188–1204, 1996.
- [15] J. W. Cahn and F. C. Larché. Thermochemical equilibrium of multiphase solids under stress. *Acta Metallurgica*, 26:1579–1589, 1978.
- [16] S. Conti, B. Niethammer, and F. Otto. Coarsening in off-critical mixtures. *SIAM J. Math. Anal.*, 37(6):1732–1741, 2006.
- [17] P. Deuffhard, P. Leinen, and H. Yserentant. Concepts of an adaptive hierarchical finite element code. *IMPACT Comput. Sci. Engrg.*, 1:3–35, 1989.
- [18] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.
- [19] W. Dreyer and W. H. Müller. A study of the coarsening in tin/lead solders. *Int. J. Solid Struct.*, 37:3841–3871, 2000.
- [20] W. Dreyer and W. H. Müller. Modeling diffusional coarsening in eutectic tin/lead solders: a quantitative approach. *Int. J. Solid Struct.*, 38:1433–1458, 2001.
- [21] I. Ekeland and R. Temam. *Convex Analysis*. Studies in Mathematics and its Applications. North-Holland, Amsterdam New York Oxford, 1976.
- [22] H. Garcke. *On mathematical models for phase separation in elastically stressed solids*. Habilitation thesis, Universität Bonn, 2000.

- [23] H. Garcke. On Cahn–Hilliard systems with elasticity. *Proc. Roy. Soc. Edinburgh*, 133 A:307–331, 2003.
- [24] H. Garcke. Mechanical effects in the Cahn–Hilliard model: A review on mathematical results. In A. Miranville, editor, *Mathematical Methods and Models in phase transitions*, pages 43–77. Nova Science Publ., 2005.
- [25] H. Garcke, B. Niethammer, M. Rumpf, and U. Weikard. Transient coarsening behaviour in the Cahn–Hilliard model. *Acta Mat.*, 51:2823–2830, 2003.
- [26] H. Garcke, M. Rumpf, and U. Weikard. The Cahn–Hilliard equation with elasticity: Finite element approximation and qualitative studies. *Interfaces Free Bound.*, 3:101–118, 2001.
- [27] H. Garcke and U. Weikard. Numerical approximation of the Cahn–Larch equation. *Numer. Math.*, 100:639–662, 2005.
- [28] G. Giacomini and J. L. Lebowitz. Phase segregation dynamics in particle systems with long range interactions I: macroscopic limits. *J. Statist. Phys.*, 87:37–61, 1998.
- [29] G. Giacomini and J. L. Lebowitz. Phase segregation dynamics in particle systems with long range interactions II: interface motion. *SIAM J. Appl. Math.*, 58:1707–1729, 1998.
- [30] C. Gräser. Globalization of nonsmooth Newton methods for optimal control problems. In K. Kunisch, G. Of, and O. Steinbach, editors, *Numerical Mathematics and Advanced Applications, Proceedings of ENUMATH 2007*, pages 605–612. Springer, 2008.
- [31] C. Gräser. *Convex Minimization and Phase Field Models*. PhD thesis, Freie Universität Berlin, 2011.
- [32] C. Gräser. Nonsmooth Schur–Newton methods for nonsmooth saddle point problems. Preprint 1004, Matheon Berlin, 2013.
- [33] C. Gräser. Truncated nonsmooth Newton multigrid methods for anisotropic convex minimization problems. Technical report, 2013. in preparation.
- [34] C. Gräser and R. Kornhuber. On preconditioned Uzawa-type iterations for a saddle point problem with inequality constraints. In O. B. Widlund and D. E. Keyes, editors, *Domain Decomposition Methods in Science and Engineering XVI*, pages 91–102. Springer, 2006.
- [35] C. Gräser and R. Kornhuber. Nonsmooth Newton methods for set-valued saddle point problems. *SIAM J. Numer. Anal.*, 47(2):1251–1273, 2009.
- [36] C. Gräser, R. Kornhuber, and U. Sack. On hierarchical error estimators for time-discretized phase field models. In G. Kreiss, P. Lötstedt, A. Malqvist, and M. Neytcheva, editors, *Proceedings of ENUMATH 2009*, pages 397–406. Springer, 2010.
- [37] C. Gräser, R. Kornhuber, and U. Sack. Nonsmooth Schur–Newton methods for vector-valued Cahn–Hilliard equations. Preprint, Serie A 01/2013, Freie Universität Berlin, 2013.
- [38] C. Gräser, R. Kornhuber, and U. Sack. Time discretizations of anisotropic Allen–Cahn equations. *IMA J. Numer. Anal.*, 33(4):1226–1244, 2013.
- [39] C. Gräser, U. Sack, and O. Sander. Truncated nonsmooth Newton multigrid methods for convex minimization problems. In M. Bercovier, M. Gander, R. Kornhuber, and O. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XVIII*, LNCSE, pages 129–136. Springer, 2009.
- [40] C. Gräser and O. Sander. The dune-subgrid module and some applications. *Computing*, 8(4):269–290, 2009.
- [41] M. Holst, J. Owall, and R. Szypowski. An efficient, reliable and robust error estimator for elliptic problems in \mathbf{R}^3 . *Appl. Num. Math.*, 61:675–695, 2011.
- [42] R. H. W. Hoppe and R. Kornhuber. Adaptive multilevel-methods for obstacle problems. *SIAM J. Numer. Anal.*, 31(2):301–323, 1994.
- [43] R. Kohn and F. Otto. Upper bounds for coarsening rates. *Comm. Math. Phys.*, 229:375–395, 2002.
- [44] R. Kornhuber. A posteriori error estimates for elliptic variational inequalities. *Comput. Math. Appl.*, 31:49–60, 1996.
- [45] R. Kornhuber and Q. Zou. Efficient and reliable hierarchical error estimates for the discretization error of elliptic obstacle problems. *Math. Comp.*, 80:69–88, 2011.
- [46] Y.-L. Lee, J. Wu, J. Xu, and L. Zikatanov. A sharp convergence estimate for the method of subspace corrections for singular systems of equations. *Math. Comp.*, 77(262):831–850, 2008.
- [47] T. Merkle. An energy method for the strongly nonlinear Cahn–Larché equation system. Preprint 16, IANS Stuttgart, 2003.
- [48] T. Merkle. *The Cahn–Larché system: A model for spinodal decomposition in eutectic solder*. PhD thesis, Universität Stuttgart, 2005.
- [49] I. Müller and W. H. Müller. *Fundamentals of Thermodynamics and Applications*. Springer, Berlin, Heidelberg, 2009.
- [50] W. H. Müller. Morphology changes in solder joints – experimental evidence and physical understanding. *Microelectronics Reliability*, 44:1901–1914, 2004.

- [51] J. Nečas and I. Hlaváček. *Mathematical Theory of Elastic and Elastico-Plastic Bodies: An Introduction*. Elsevier, Amsterdam, 1981.
- [52] A. Novick-Cohen. Upper bounds for coarsening for the deep quench obstacle problem. *J. Stat. Phys.*, 141(1):142–157, 2010.
- [53] A. Novick-Cohen and A. Shishkov. Upper bounds for coarsening for the degenerate Cahn–Hilliard equation. *Discrete Contin. Dyn. Syst. A*, 25(1):251–272, 2009.
- [54] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7(3):856–869, 1986.
- [55] J. Schöberl and W. Zulehner. On additive Schwarz-type smoothers for saddle point problems. *Numer. Math.*, 95(2):377–399, 2003.
- [56] K. G. Siebert and A. Veese. A unilaterally constrained quadratic minimization with adaptive finite elements. *SIAM J. Optim.*, 18:260–289, 2007.
- [57] E. Underwood. *Quantitative Stereology*. Metallurgy and Materials. Addison-Wesley, Reading, Mass., 1970.
- [58] S. P. Vanka. Block-implicit multigrid solution of Navier–Stokes equations in primitive variables. *J. Comput. Phys.*, 65:138–158, 1986.
- [59] U. Weikard. *Numerische Lösungen der Cahn–Hilliard Gleichung und der Cahn–Larché Gleichung*. PhD thesis, Rheinische Friedrich–Wilhelms-Universität Bonn, 2002.
- [60] O. C. Zienkiewicz, J. P. De S. R. Gago, and D. W. Kelly. The hierarchical concept in finite element analysis. *Computers & Structures*, 16:53–65, 1983.
- [61] Q. Zou, A. Veese, R. Kornhuber, and C. Gräser. Hierarchical error estimates for the energy functional in obstacle problems. *Numer. Math.*, 117(4):653–677, 2011.
- [62] W. Zulehner. A class of smoothers for saddle point problems. *Computing*, 65(3):227–246, 2000.

CARSTEN GRÄSER, FREIE UNIVERSITÄT BERLIN, INSTITUT FÜR MATHEMATIK, ARNIMALLEE 6, 14195 BERLIN, GERMANY

E-mail address: `graeser@math.fu-berlin.de`

RALF KORNUBER, FREIE UNIVERSITÄT BERLIN, INSTITUT FÜR MATHEMATIK, ARNIMALLEE 6, 14195 BERLIN, GERMANY

E-mail address: `kornhuber@math.fu-berlin.de`

ULI SACK, FREIE UNIVERSITÄT BERLIN, INSTITUT FÜR MATHEMATIK, ARNIMALLEE 6, 14195 BERLIN, GERMANY

E-mail address: `usack@math.fu-berlin.de`