Universität Regensburg
Mathematik

Non-local Allen-Cahn systems: Analysis and a primal dual active set method

Luise Blank, Harald Garcke, Lavinia Sarbu
and Vanessa Styles

Preprint Nr. 02/2011

Non-local Allen-Cahn systems: Analysis and a primal dual active set method

Luise Blank [†] Harald Garcke[†] Lavinia Sarbu[‡]
Vanessa Styles[‡]

Abstract

We show existence and uniqueness of a solution for the non-local vector-valued Allen-Cahn variational inequality in a formulation involving Lagrange multipliers for local and non-local constraints. Furthermore, we propose and analyze a primal-dual active set method for local and non-local vector-valued Allen-Cahn variational inequalities. Convergence of the primal-dual active set algorithm is shown by interpreting the approach as a semi-smooth Newton method and numerical simulations are presented demonstrating its efficiency.

Key words: Allen-Cahn systems, non-local constraints, variational inequality, vector-valued obstacle problems, primal-dual active set method, semi-smooth Newton method.

AMS subject classification. 35K55, 65K10, 90C33, 90C53, 82C24, 65M60

1 Introduction

The Allen-Cahn equation was introduced by Allen and Cahn [1] and describes the capillarity driven evolution of an interface separating two bulk phases. In the Allen-Cahn model interfaces are modelled to have a thickness of order ε where $0 < \varepsilon \ll 1$, and in the interfacial layer a phase field or order parameter rapidly changes its value. The Allen-Cahn model (or phase field model) has a variety of applications, e.g. in materials science, image processing, biology and geology, see [6, 11, 14, 23, 25, 30, 39]. In many of these applications more than two phases occur. Therefore, the model has been extended to deal with N phases [13, 24]. The phase field takes now the form of a vector-valued function $\mathbf{u} : \Omega \times (0,T) \to \mathbb{R}^N$ which describes the fractions of the phases, i.e. each component of $\mathbf{u}$ describes one phase.
The underlying non-convex energy functional is based on the Ginzburg-Landau energy for the vector-valued phase field $\mathbf{u} \in \mathbb{R}^N$

$$E(\mathbf{u}) := \int_\Omega \left(\tfrac{\gamma\varepsilon}{2} |\nabla \mathbf{u}|^2 + \tfrac{\gamma}{\varepsilon} \psi(\mathbf{u}) \right) dx$$

[†]Fakultät für Mathematik, Universität Regensburg, 93040 Regensburg, Germany
[‡]Department of Mathematics, University of Sussex, Brighton BN1 9RF, UK

where $\Omega \subset \mathbb{R}^d$ is a bounded domain, $\gamma > 0$ is a parameter related to the interfacial energy and ψ is a bulk potential. Since each component of $\mathbf{u}$ stands for the fraction of one phase, the phase space for the order parameter $\mathbf{u}$ is the Gibbs simplex

$$\boldsymbol{G} := \{\boldsymbol{\xi} \in \mathbb{R}^N \mid \boldsymbol{\xi} \geq \mathbf{0}, \boldsymbol{\xi} \cdot \mathbf{1} = 1\}.$$

Here $\boldsymbol{\xi} \geq \mathbf{0}$ means $\xi_i \geq 0$ for all $i \in \{1, ..., N\}$, $\mathbf{1} = (1, ..., 1)^T$ and $\boldsymbol{\xi} \cdot \mathbf{1} = \sum_{i=1}^{N} \xi_i$.
For the bulk potential $\psi : \mathbb{R}^N \to \mathbb{R}_0^+ \cup \{\infty\}$ we consider the multi-obstacle potential

$$\psi(\boldsymbol{\xi}) = \psi_0(\boldsymbol{\xi}) + I_G = \begin{cases} \psi_0(\boldsymbol{\xi}) := -\frac{1}{2}\boldsymbol{\xi} \cdot \mathbf{A}\boldsymbol{\xi} & \text{for } \boldsymbol{\xi} \in \boldsymbol{G}, \\ \infty & \text{otherwise,} \end{cases} \tag{1}$$

where I_G is the indicator function of the Gibbs simplex and $\mathbf{A}$ is a symmetric constant $N \times N$ matrix [19]. Let $\sigma_{max}(\mathbf{A})$ and $\sigma_{min}(\mathbf{A})$ be the largest and lowest eigenvalues and $|||\mathbf{A}|||$ the spectral norm of $\mathbf{A}$. If all eigenvalues of $\mathbf{A}$ are negative ψ would be a convex potential. Different phases which correspond to minima of ψ only occur if $\mathbf{A}$ has at least one positive eigenvalue. We hence assume that $\mathbf{A}$ has at least one positive eigenvalue; the analysis in this paper would simplify if this were not the case.
Given an initial phase distribution $\mathbf{u}(., 0) = \mathbf{u}_0 : \Omega \to \boldsymbol{G}$ at time $t = 0$ the interface motion can be modelled by the steepest descent dynamics of E with respect to the L^2-norm which results, after suitable rescaling of time, in the following vector-valued Allen-Cahn equation

$$\varepsilon \frac{\partial \mathbf{u}}{\partial t} = -\text{grad}_{L^2} E(\mathbf{u}) = \gamma\varepsilon\Delta\mathbf{u} + \frac{\gamma}{\varepsilon}(\mathbf{A}\mathbf{u} - \boldsymbol{\mu}^*) \tag{2}$$

where $\boldsymbol{\mu}^* \in \partial I_G$ and ∂I_G denotes the subdifferential of I_G. As for the scalar case, see e.g. [9, 12], this equation leads to the following variational inequality

$$\varepsilon(\frac{\partial \mathbf{u}}{\partial t}, \boldsymbol{\chi} - \mathbf{u}) + \gamma\varepsilon(\nabla\mathbf{u}, \nabla(\boldsymbol{\chi} - \mathbf{u})) - \frac{\gamma}{\varepsilon}(\mathbf{A}\mathbf{u}, \boldsymbol{\chi} - \mathbf{u}) \geq 0 \tag{3}$$

which has to hold for almost all t and all $\boldsymbol{\chi} \in \mathbf{H}^1(\Omega)$ with $\boldsymbol{\chi} \in \boldsymbol{G}$ *a.e.*. Here, we denote by $\mathbf{L}^2(\Omega)$ and $\mathbf{H}^1(\Omega)$ the spaces of vector-valued functions, $(.,.)$ is the standard L^2 inner product for scalar functions, $(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{N}(v_i, w_i)$ for $\mathbf{v}, \mathbf{w} \in \mathbf{L}^2(\Omega)$ and $(\mathbf{A}, \mathbf{B}) = \sum_{i=1}^{N}\sum_{j=1}^{d}(a_{ij}, b_{ij})$ for matrix-valued functions.
Often one considers systems in which the total spatial amount of the phases are conserved. In this case one studies the steepest descent of E under the constraint $\fint_\Omega \mathbf{u}\, dx = \mathbf{m} = (m^i)$ where $m^i \in (0, 1)$ for $i \in \{1, ..., N\}$ is a fixed number and we use the notation $\fint_\Omega f(x)dx := \frac{1}{|\Omega|}\int_\Omega f(x)dx$ with $|\Omega|$ being the Lebesgue measure of Ω. Here we use the notation $\mathbf{m}$ and m^i in order to avoid confusion with the mass

vector $\boldsymbol{m}$ and its components m_i which is introduced in Section 4. To ensure that all phases are present we require $0 < m^i < 1$ and $\sum_{i=1}^{N} m^i = 1$, where the last condition makes sure that $\sum_{i=1}^{N} u_i = 1$ can be true. We define

$$\mathcal{G} := \{\mathbf{v} \in H^1(\Omega) \mid \mathbf{v} \in \mathbf{G} \quad a.e.\} \quad \text{and} \quad \mathcal{G}^{\mathbf{m}} := \{\mathbf{v} \in \mathcal{G} \mid \fint_{\Omega} \mathbf{v} = \mathbf{m}\}\,.$$

Then the interface evolution with mass conservation can be formulated as:

($\mathbf{P_m}$) *For given* $\mathbf{u}(.,0) = \mathbf{u}_0 \in \mathcal{G}^{\mathbf{m}}$ *find* $\mathbf{u} \in L^2(0,T;\mathcal{G}^{\mathbf{m}}) \cap H^1(0,T;\mathbf{L}^2(\Omega))$ *such that*

$$\varepsilon(\tfrac{\partial \mathbf{u}}{\partial t}, \boldsymbol{\chi} - \mathbf{u}) + \gamma\varepsilon(\nabla\mathbf{u}, \nabla(\boldsymbol{\chi} - \mathbf{u})) - \tfrac{\gamma}{\varepsilon}(\mathbf{A}\mathbf{u}, \boldsymbol{\chi} - \mathbf{u}) \geq 0 \tag{4}$$

which has to hold for almost all $t \in (0,T)$ *and all* $\boldsymbol{\chi} \in \mathcal{G}^{\mathbf{m}}$.

This paper is organized as follows. In Section 2 we reformulate ($\mathbf{P_m}$) with the help of Lagrange multipliers $\boldsymbol{\mu}, \boldsymbol{\lambda}$ and Λ corresponding to the constraints $\mathbf{u} \geq \mathbf{0}$, $\fint_{\Omega} \mathbf{u} = \mathbf{m}$ and $\sum_{i=1}^{N} u_i = 1$ respectively. We show existence and uniqueness of a solution $\mathbf{u}$ of ($\mathbf{P_m}$) and of the Lagrange multipliers, $\boldsymbol{\mu}, \boldsymbol{\lambda}$ and Λ.
In Sections 3 and 4 we introduce the main ideas of a primal-dual active set strategy. We apply the algorithm to a finite element discretization of an implicit Euler-discretization of ($\mathbf{P_m}$). Using that the primal-dual active set method can be reformulated as a semi-smooth Newton method [27] we show local convergence of our algorithm.
Finally, in Section 5 we present numerical simulations for the non-local as well as for the local Allen-Cahn variational inequality with three and more phases. Using two model problems for which the explicit solution is known we show the efficiency and accuracy of the proposed method. We also discuss numerically how the primal-dual active set method depends on mesh parameters as well as on the number of phases.

2 Existence theory

In this section we show existence and uniqueness to the vector-valued Allen-Cahn variational inequality with integral constraints. As a first step we reformulate the problem ($\mathbf{P_m}$) with the help of (scaled) Lagrange multipliers $\boldsymbol{\mu}$ corresponding to the inequality constraint $\mathbf{u} \geq \mathbf{0}$, Λ corresponding to the constraint $\sum_{i=1}^{N} u_i = 1$ and $\boldsymbol{\lambda}$ corresponding to the constraint $\fint_{\Omega} \mathbf{u} = \mathbf{m}$. However, to be more precise, we have to reformulate the integral constraints in order to guarantee that the constraints are linearly independent, which is required to obtain uniqueness of the Lagrange multipliers. Let us therefore consider the sum constraint $\sum_{i=1}^{N} u_i = 1$, which is equivalent

to $(\sum_{i=1}^{N} u_i - 1, v) = 0$ for all $v \in L^2(\Omega)$, and the integral constraint $\fint_\Omega \mathbf{u} = \mathbf{m}$, which is equivalent to $(\mathbf{u} - \mathbf{m}, \mathbf{e}_i) = 0$ for all $i = 1, ..., N$ where $\mathbf{e}_i$ is the function which is identical to 1 in the i-th component and 0 otherwise. Then for $v \equiv 1$ we have $(\sum_{i=1}^{N} u_i - 1, 1) = \sum_{i=1}^{N} (u_i - m^i, 1) = \sum_{i=1}^{N} (\fint_\Omega \mathbf{u} - \mathbf{m}, \mathbf{e}_i)$ which reveals the linear dependence of the constraints. For the reformulation we observe that if $\sum_{i=1}^{N} u_i = 1$ it follows that $\sum_{i=1}^{N} (\fint_\Omega u_i - m^i) = 0$, i.e.

$$\fint_\Omega \mathbf{u} - \mathbf{m} \in S := \{\mathbf{v} \in \mathbb{R}^N \mid \textstyle\sum_{i=1}^{N} v_i = 0\}.$$

Noting the above it follows that if $\sum_{i=1}^{N} u_i = 1$ and $\mathbf{P}_S(\fint_\Omega \mathbf{u} - \mathbf{m}) = \mathbf{0}$ it follows that $\fint_\Omega \mathbf{u} = \mathbf{m}$.

Since we have $\sum_{i=1}^{N} u_i = 1$ we substitute the integral constraint $\fint_\Omega \mathbf{u} = \mathbf{m}$ by $\mathbf{P}_S(\fint_\Omega \mathbf{u} - \mathbf{m}) = \mathbf{0}$ and we take $\boldsymbol{\lambda}(t) \in S$, for almost every t, to be the corresponding Lagrange multiplier.

Lemma 2.1 *Let T be a positive time and let $\Omega \subset \mathbb{R}^d$ be a bounded domain which is either convex or fulfills $\partial\Omega \in C^{1,1}$. If there exist $\mathbf{u} \in L^2(0,T;\mathbf{H}^2(\Omega)) \cap H^1(0,T;\mathbf{L}^2(\Omega)) \cap L^2(0,T;\mathcal{G}^{\mathbf{m}})$, $\boldsymbol{\mu} \in L^2(0,T;\mathbf{L}^2(\Omega))$, $\boldsymbol{\lambda} \in \mathbf{L}^2(0,T;S)$ and $\Lambda \in L^2(0,T;L^2(\Omega))$ such that*

$$\varepsilon \tfrac{\partial \mathbf{u}}{\partial t} - \gamma\varepsilon\Delta\mathbf{u} - \tfrac{\gamma}{\varepsilon}\mathbf{A}\mathbf{u} - \tfrac{1}{\varepsilon}\boldsymbol{\mu} - \tfrac{1}{\varepsilon}\Lambda\mathbf{1} - \tfrac{1}{\varepsilon}\boldsymbol{\lambda} = \mathbf{0} \qquad \text{a.e. in } \Omega_T := \Omega \times (0,T), \tag{5}$$

$$\mathbf{u}(0) = \mathbf{u}_0, \quad \tfrac{\partial \mathbf{u}}{\partial \nu} = \mathbf{0} \qquad \text{a.e. on } \partial\Omega \times (0,T), \tag{6}$$

$$\sum_{i=1}^{N} u_i = 1, \quad \mathbf{u} \geq \mathbf{0}, \quad \boldsymbol{\mu} \geq \mathbf{0} \qquad \text{a.e. in } \Omega_T, \tag{7}$$

$$\mathbf{P}_S(\fint_\Omega \mathbf{u} - \mathbf{m}) = \mathbf{0}, \quad (\boldsymbol{\mu}, \mathbf{u}) = 0 \qquad \text{for almost all } t \in (0,T), \tag{8}$$

then $\mathbf{u}$ solves $(\mathbf{P}_\mathbf{m})$.

Proof: We have $\mathbf{u}(\cdot, t) \in \mathcal{G}^{\mathbf{m}}$ for almost all t and now choose $\boldsymbol{\chi} \in \mathcal{G}^{\mathbf{m}}$. Multiplying (5) by $(\boldsymbol{\chi} - \mathbf{u})$ and integrating by parts and noting (6) gives

$$\int_\Omega (\varepsilon \tfrac{\partial \mathbf{u}}{\partial t} - \tfrac{\gamma}{\varepsilon}\mathbf{A}\mathbf{u} - \tfrac{1}{\varepsilon}\boldsymbol{\mu} - \tfrac{1}{\varepsilon}\Lambda\mathbf{1} - \tfrac{1}{\varepsilon}\boldsymbol{\lambda}) \cdot (\boldsymbol{\chi} - \mathbf{u}) + \int_\Omega \gamma\varepsilon\nabla\mathbf{u} \cdot \nabla(\boldsymbol{\chi} - \mathbf{u}) = 0$$

for a.e. $t \in (0,T)$. Using the property $\boldsymbol{\chi} \geq \mathbf{0}$ and (7) - (8) gives $\int_\Omega \boldsymbol{\mu} \cdot (\boldsymbol{\chi} - \mathbf{u}) \geq 0$. Using (8) and $\fint_\Omega \boldsymbol{\chi} = \mathbf{m}$ we have $\int_\Omega (\boldsymbol{\chi} - \mathbf{u}) = \mathbf{0}$. The fact that $\boldsymbol{\lambda}$ is independent of

$x \in \Omega$ yields

$$\int_\Omega \boldsymbol{\lambda} \cdot (\boldsymbol{\chi} - \mathbf{u}) = 0.$$

Since $\mathbf{1} \cdot (\boldsymbol{\chi} - \mathbf{u}) = 0$ also $\int_\Omega \Lambda \mathbf{1} \cdot (\boldsymbol{\chi} - \mathbf{u}) = 0$. Hence we obtain for all $\boldsymbol{\chi} \in \mathcal{G}^{\mathbf{m}}$ and almost all $t \in (0,T)$

$$\int_\Omega (\varepsilon \tfrac{\partial \mathbf{u}}{\partial t} - \tfrac{\gamma}{\varepsilon} \mathbf{A}\mathbf{u}) \cdot (\boldsymbol{\chi} - \mathbf{u}) + \int_\Omega \gamma \varepsilon \nabla \mathbf{u} \cdot \nabla (\boldsymbol{\chi} - \mathbf{u}) \geq 0$$

and hence $\mathbf{u}$ solves $(\mathbf{P_m})$. □

To show the existence of $(\mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \Lambda)$ we now regularize ψ, i.e. I_G, and therefore substitute $\boldsymbol{\mu}^* \in \partial I_G$ in (2). First, we introduce the following regularization of $\psi_0(\boldsymbol{\xi}) + I_{\{\boldsymbol{\xi} \geq \mathbf{0}\}}$, where $I_{\{\boldsymbol{\xi} \geq \mathbf{0}\}}$ is the indicator function of the set $\{\boldsymbol{\xi} \in \mathbb{R}^N \mid \boldsymbol{\xi} \geq \mathbf{0}\}$,

$$\psi_\delta(\boldsymbol{\xi}) = \psi_0(\boldsymbol{\xi}) + \tfrac{1}{\delta} \hat{\psi}(\boldsymbol{\xi}) \tag{9}$$

where

$$\hat{\psi}(\boldsymbol{\xi}) = \sum_{i=1}^{N} (\min(\xi_i, 0))^2. \tag{10}$$

Similar regularizations were used in [3, 4, 19].

In order to deal with the constraints $\sum_{i=1}^{N} u_i = 1$ and $\fint_\Omega \mathbf{u} = \mathbf{m}$ we project $D\psi_\delta$ orthogonally onto the corresponding tangent space. This projection $\mathbf{P}$ can be realized by successive orthogonal projections $\mathbf{P}_\fint$ and $\mathbf{P}_\sum$ where

$$\mathbf{P}_\fint \mathbf{v} := \mathbf{v} - \fint_\Omega \mathbf{v},\ \mathbf{P}_\sum \mathbf{v} := \mathbf{v} - (\textstyle\sum \mathbf{v})\mathbf{1} \text{ and } \textstyle\sum \mathbf{v} := \frac{1}{N} \sum_{i=1}^{N} v_i \text{ for all } \mathbf{v} \in \mathbb{R}^N .$$

We note that $\mathbf{P} = \mathbf{P}_\fint \mathbf{P}_\sum = \mathbf{P}_\sum \mathbf{P}_\fint$. This results in the following regularized version of the Allen-Cahn equation

$$\varepsilon \tfrac{\partial \mathbf{u}_\delta}{\partial t} - \gamma \varepsilon \Delta \mathbf{u}_\delta + \tfrac{\gamma}{\varepsilon} \mathbf{P}_\fint \mathbf{P}_\sum D\psi_\delta(\mathbf{u}_\delta) = \mathbf{0}\,. \tag{11}$$

Equivalently we have to solve the following problem:

$(\mathbf{P}^\delta_\mathbf{m})$ *Given* $\mathbf{u}_0 \in \mathcal{G}^{\mathbf{m}}$ *find* $\mathbf{u}_\delta \in H^1(0,T;\mathbf{L}^2(\Omega)) \cap L^2(0,T;\mathbf{H}^1(\Omega))$ *such that* $\mathbf{u}_\delta(.,0) = \mathbf{u}_0$ *and*

$$\varepsilon(\tfrac{\partial \mathbf{u}_\delta}{\partial t}, \boldsymbol{\chi}) + \gamma\varepsilon(\nabla \mathbf{u}_\delta, \nabla \boldsymbol{\chi}) + \tfrac{\gamma}{\varepsilon}(\mathbf{P}_\sum(D\psi_\delta(\mathbf{u}_\delta)), \mathbf{P}_\fint \boldsymbol{\chi}) = 0 \tag{12}$$

for all $\boldsymbol{\chi} \in \mathbf{H}^1(\Omega)$ *and almost all* $t \in (0,T)$.

Choosing $\boldsymbol{\chi} = \chi \mathbf{1}$, for any $\chi \in H^1(\Omega)$, in (12) gives that a solution $\mathbf{u}_\delta$ of $(\mathbf{P}^\delta_\mathbf{m})$ fulfills $\frac{\partial}{\partial t}(\sum_{i=1}^{N} (u_\delta)_i) - \Delta(\sum_{i=1}^{N} (u_\delta)_i) = 0$ in a weak sense. Using the facts that $\mathbf{u}_0 \in \mathcal{G}^{\mathbf{m}}$ and

that the initial value problem to the parabolic problem $\partial_t v - \Delta v = 0$ with Neumann boundary conditions is uniquely solvable now gives

$$\sum_{i=1}^{N} (u_\delta(x,t))_i = 1 \quad \text{for almost all} \quad (x,t) \in \Omega_T. \tag{13}$$

Hence $\mathbf{P}_\sum(\mathbf{u}_\delta) = \mathbf{u}_\delta$. Choosing constant test functions in (12) gives

$$\tfrac{d}{dt} \fint_\Omega \mathbf{u}_\delta = \mathbf{0} \tag{14}$$

and hence the total masses of the components of $\mathbf{u}_\delta$ are preserved, i.e. $\fint_\Omega \mathbf{u}_\delta = \mathbf{m}$.
Before we show that the regularized problem $(\mathbf{P}^\delta_\mathbf{m})$ has a unique solution, let us state some properties of ψ_δ and $D\psi_\delta$.
We have $D\psi_\delta(\boldsymbol{\xi}) = \frac{1}{\delta}\hat{\boldsymbol{\phi}}(\boldsymbol{\xi}) - \mathbf{A}\boldsymbol{\xi}$ with $\hat{\boldsymbol{\phi}}(\boldsymbol{\xi}) := D\psi(\boldsymbol{\xi}) = (\hat{\phi}(\xi_i))_{i=1}^N$ where $\hat{\phi}(r) = 2\min(r,0) = 2[r]_-$ and $[\cdot]_- := \min(\cdot, 0)$.
Furthermore:

- For all $r, s \in \mathbb{R}$
$$0 \le (\hat{\phi}(r) - \hat{\phi}(s))(r-s). \tag{15}$$

- For all $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^N$
$$\begin{aligned} (\boldsymbol{\xi} - \boldsymbol{\eta}) \cdot D\psi_\delta(\boldsymbol{\eta}) &\le \tfrac{1}{\delta}\hat{\psi}(\boldsymbol{\xi}) - \tfrac{1}{\delta}\hat{\psi}(\boldsymbol{\eta}) - (\boldsymbol{\xi} - \boldsymbol{\eta}) \cdot \mathbf{A}\boldsymbol{\eta} \\ &= \psi_\delta(\boldsymbol{\xi}) - \psi_\delta(\boldsymbol{\eta}) + \tfrac{1}{2}(\boldsymbol{\xi} - \boldsymbol{\eta}) \cdot \mathbf{A}(\boldsymbol{\xi} - \boldsymbol{\eta}) \\ &\le \psi_\delta(\boldsymbol{\xi}) - \psi_\delta(\boldsymbol{\eta}) + \tfrac{1}{2}\sigma_{max}(\mathbf{A})\|\boldsymbol{\xi} - \boldsymbol{\eta}\|^2, \end{aligned} \tag{16}$$

  where we have used that $\hat{\psi}$ is convex and the identity

$$-2(\boldsymbol{\xi} - \boldsymbol{\eta}) \cdot \mathbf{A}\boldsymbol{\eta} = \boldsymbol{\eta} \cdot \mathbf{A}\boldsymbol{\eta} - \boldsymbol{\xi} \cdot \mathbf{A}\boldsymbol{\xi} + (\boldsymbol{\xi} - \boldsymbol{\eta}) \cdot \mathbf{A}(\boldsymbol{\xi} - \boldsymbol{\eta}).$$

- For all $\boldsymbol{\xi} \in \mathcal{M} := \{\boldsymbol{\xi} \in \mathbb{R}^N : \sum_{i=1}^N \xi_i = 1\}$ and $\delta \le \delta_0 := \frac{1}{4N(N-1)^2\sigma_{max}(\mathbf{A})}$ we have that
$$\psi_\delta(\boldsymbol{\xi}) \ge \tfrac{1}{2\delta} \sum_{i=1}^N [\xi_i]_-^2 - C(N, \sigma_{max}(\mathbf{A}))\,. \tag{17}$$

  This follows from

$$\begin{aligned} \psi_\delta(\boldsymbol{\xi}) &\ge \tfrac{1}{\delta} \sum_{i=1}^N [\xi_i]_-^2 - \sum_{i=1}^N \sigma_{max}(\mathbf{A})\xi_i^2 \\ &\ge \tfrac{1}{2\delta} \sum_{i=1}^N [\xi_i]_-^2 + \tfrac{1}{2\delta}[\xi_m]_-^2 - N\sigma_{max}(\mathbf{A})\xi_M^2 \end{aligned} \tag{18}$$

where $\xi_m := \min_{i=1,\dots,N} \xi_i$ and ξ_M is chosen such that $|\xi_M| = \max_{i=1,\dots,N} |\xi_i|$.
Since $\boldsymbol{\xi} \in \mathcal{M}$ it follows that

$$[\xi_m]_- \leq \xi_M \leq 1 - (N-1)[\xi_m]_-$$

and

$$\xi_M^2 \leq 2(1 + (N-1)^2[\xi_m]_-^2)$$

and hence (17) follows from (18) for all $\delta \leq \delta_0$.

Theorem 2.1 *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain and assume that either Ω is convex or $\partial\Omega \in C^{1,1}$. Let $\mathbf{u}_0(x) \in \mathbf{H}^1(\Omega)$ with $\mathbf{u}_0 \geq \mathbf{0}$, $\fint \mathbf{u}_0 = \mathbf{m} > \mathbf{0}$ and $\sum_{i=1}^{N}(u_0)_i = 1$ a.e. in Ω. Then there exists a unique solution $\mathbf{u}_\delta$ to $(\mathbf{P}_\mathbf{m}^\delta)$ for all $\delta \in (0, \delta_0]$ and a constant $C > 0$ which does not depend on δ such that*

$$\|\mathbf{u}_\delta\|_{L^\infty(0,T;\mathbf{H}^1(\Omega))} + \|\mathbf{u}_\delta\|_{H^1(0,T;\mathbf{L}^2(\Omega))} \leq C, \tag{19}$$

$$\|[\mathbf{u}_\delta]_-\|_{L^\infty(0,T;\mathbf{L}^2(\Omega))} \leq C\delta^{1/2}, \tag{20}$$

$$\tfrac{1}{\delta}\|\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq C, \tag{21}$$

and

$$\|\mathbf{u}_\delta\|_{L^2(0,T;\mathbf{H}^2(\Omega))} \leq C. \tag{22}$$

Proof: The existence of a solution to $(\mathbf{P}_\mathbf{m}^\delta)$ follows by using a Galerkin approximation, a priori estimates and compactness arguments. Using the assumptions on Ω and the growth property of $D\psi_\delta$ regularity theory gives $\mathbf{u}_\delta \in L^2(0,T;\mathbf{H}^2(\Omega))$, using methods from [21] and [26].
Now we show that the solution is unique. Therefore assume that $(\mathbf{P}_\mathbf{m}^\delta)$ has two solutions $\mathbf{u}_\delta^1, \mathbf{u}_\delta^2$, subtracting them and choosing $\boldsymbol{\chi} \equiv \mathbf{d} := \mathbf{u}_\delta^1 - \mathbf{u}_\delta^2$ gives

$$\varepsilon(\tfrac{\partial \mathbf{d}}{\partial t}, \mathbf{d}) + \gamma\varepsilon\|\nabla \mathbf{d}\|^2 + \tfrac{\gamma}{\varepsilon}(\mathbf{P}_\sum(\tfrac{1}{\delta}\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta^1) - \tfrac{1}{\delta}\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta^2) - \mathbf{A}\mathbf{d}), \mathbf{P}_\int \mathbf{d}) = 0.$$

Since $\sum \mathbf{d} = 0$ and $\fint_\Omega \mathbf{d} = \mathbf{0}$ it follows that

$$\tfrac{\varepsilon}{2}\tfrac{d}{dt}\|\mathbf{d}\|^2 + \gamma\varepsilon\|\nabla \mathbf{d}\|^2 + \tfrac{\gamma}{\delta\varepsilon}(\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta^1) - \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta^2), \mathbf{d}) = \tfrac{\gamma}{\varepsilon}(\mathbf{A}\mathbf{d}, \mathbf{d}) \leq \tfrac{\gamma}{\varepsilon}\sigma_{max}(\mathbf{A})\|\mathbf{d}\|^2$$

where $\|\cdot\|$ denotes the $\mathbf{L}^2$ -norm. With (15) we obtain

$$\tfrac{\varepsilon}{2}\tfrac{d}{dt}\|\mathbf{d}\|^2 \leq \tfrac{\gamma}{\varepsilon}\sigma_{max}(\mathbf{A})\|\mathbf{d}\|^2.$$

Now Grönwall's inequality gives that $\|\mathbf{d}\|^2 = 0$ and thus $\mathbf{u}_\delta^1 \equiv \mathbf{u}_\delta^2$.
Choosing $\boldsymbol{\chi} \equiv \partial\mathbf{u}_\delta/\partial t$ in (12) and using $\mathbf{P}_\sum \mathbf{P}_\int \frac{\partial \mathbf{u}_\delta}{\partial t} = \frac{\partial \mathbf{u}_\delta}{\partial t}$ which is due to (13) and (14) we obtain

$$\varepsilon\|\tfrac{\partial \mathbf{u}_\delta}{\partial t}\|^2 + \tfrac{\gamma\varepsilon}{2}\tfrac{d}{dt}\|\nabla \mathbf{u}_\delta\|^2 + \tfrac{\gamma}{\varepsilon}(D\psi_\delta(\mathbf{u}_\delta), \tfrac{\partial \mathbf{u}_\delta}{\partial t}) = 0.$$

Integrating this equation over $(0,t)$ and rearranging gives for almost all t

$$\int_0^t \varepsilon \|\tfrac{\partial \mathbf{u}_\delta(s)}{\partial t}\|^2 ds + \tfrac{\gamma\varepsilon}{2}\|\nabla \mathbf{u}_\delta(t)\|^2 + \tfrac{\gamma}{\varepsilon}(\psi_\delta(\mathbf{u}_\delta(t)),1) = \tfrac{\gamma\varepsilon}{2}\|\nabla \mathbf{u}_0\|^2 + \tfrac{\gamma}{\varepsilon}(\psi_\delta(\mathbf{u}_0),1) \le C \tag{23}$$

where C does not depend on δ which follows since $\mathbf{u}_0 \in \mathcal{G}^{\mathbf{m}}$ implies $\psi_\delta(\mathbf{u}_0) = \psi(\mathbf{u}_0)$. In particular, it follows from (17) and (23) that

$$|(\psi_\delta(\mathbf{u}_\delta)(t),1)| \le C. \tag{24}$$

Using (17) gives

$$\sum_{n=1}^N \|[(\mathbf{u}_\delta(t))_n]_-\|^2 = (\sum_{n=1}^N [(\mathbf{u}_\delta(t))_n]_-^2, 1) \le C\delta$$

for almost all t. In conclusion we have that $\|[\mathbf{u}_\delta(t)]_-\| \le C\delta^{1/2}$ for almost all $t \in (0,T)$ and thus (20) follows.
Furthermore, from (23) and (24) it follows that $\|\nabla \mathbf{u}_\delta(t)\| \le C$ for almost all t and using the Poincaré inequality

$$\|\eta\| \le C_P(\|\nabla \eta\| + |(\eta,1)|) \qquad \text{for all } \eta \in H^1(\Omega) \tag{25}$$

gives that $\|\mathbf{u}_\delta(t)\|^2_{\mathbf{H}^1(\Omega)} \le C$ for almost all $t \in (0,T)$ and thus $\mathbf{u}_\delta$ is uniformly bounded in $L^\infty(0,T;\mathbf{H}^1(\Omega))$. From (23) and (24) it also follows that $\left(\frac{\partial \mathbf{u}_\delta}{\partial t}\right)_{\delta>0}$ is uniformly bounded in $L^2(0,T;\mathbf{L}^2(\Omega))$. Hence (19) is proven.
Now we bound $\frac{1}{\delta}\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)$ in $\mathbf{L}^2$.
Setting $\boldsymbol{\chi} \equiv \frac{1}{\delta}\mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)$ in (12) gives

$$\tfrac{\varepsilon}{\delta}(\tfrac{\partial \mathbf{u}_\delta}{\partial t}, \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)) + \tfrac{\gamma\varepsilon}{\delta}(\nabla \mathbf{u}_\delta, \nabla \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)) + \tfrac{\gamma}{\varepsilon\delta}(\mathbf{P}_\sum(D\psi_\delta(\mathbf{u}_\delta)), \mathbf{P}_\int \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)) = 0.$$

Using $\sum_{i=1}^N \nabla (u_\delta)_i = 0$ a.e. which follows from (13) and $\mathbf{P}_\sum(D\psi_\delta(\mathbf{u}_\delta)) \equiv \mathbf{P}_\sum(\frac{1}{\delta}\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)) - \mathbf{P}_\sum(\mathbf{A}\mathbf{u}_\delta)$ we compute

$$\begin{aligned}\tfrac{\varepsilon}{\delta}(\tfrac{\partial \mathbf{u}_\delta}{\partial t}, \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)) + \tfrac{\gamma\varepsilon}{\delta}(\nabla \mathbf{u}_\delta, \nabla \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)) + \tfrac{\gamma}{\delta^2\varepsilon}(\mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta), \mathbf{P}_\int \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)) \\ = \tfrac{\gamma}{\varepsilon\delta}(\mathbf{P}_\sum \mathbf{A}\mathbf{u}_\delta, \mathbf{P}_\int \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)).\end{aligned}$$

Noting that $(\mathbf{P}_\int v, \fint_\Omega v) = 0$ for any $v \in L^2(\Omega)$ and using (14) we obtain

$$\begin{aligned}\tfrac{\gamma\varepsilon}{\delta}(\nabla \mathbf{u}_\delta, \nabla \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)) + \tfrac{\gamma}{\delta^2\varepsilon}\|\mathbf{P}_\int \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|^2 \\ \le \tfrac{\varepsilon}{\delta}|(\tfrac{\partial \mathbf{u}_\delta}{\partial t}, \mathbf{P}_\int \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta))| + \tfrac{\gamma}{\varepsilon\delta}|(\mathbf{P}_\sum \mathbf{A}\mathbf{u}_\delta, \mathbf{P}_\int \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta))|.\end{aligned}$$

Since $\hat{\phi}$ is non-decreasing we have that $\int_0^T (\nabla \mathbf{u}_\delta, \nabla \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta))dt \geq 0$ and hence Young's inequality and the uniform estimates on $\mathbf{u}_\delta \in L^\infty(0,T;\mathbf{H}^1(\Omega))$ and $\partial_t \mathbf{u}_\delta \in L^2(0,T;\mathbf{L}^2(\Omega))$ yield

$$\tfrac{1}{\delta^2}\|\mathbf{P}_\int \mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq C. \tag{26}$$

Choosing $\boldsymbol{\chi} = \mathbf{u}_\delta$ in (12) and using (13) we obtain

$$\begin{array}{rl} 0 &= \varepsilon(\frac{\partial \mathbf{u}_\delta}{\partial t}, \mathbf{u}_\delta) + \gamma\varepsilon(\nabla \mathbf{u}_\delta, \nabla \mathbf{u}_\delta) + \frac{\gamma}{\varepsilon}(\mathbf{P}_\sum D\psi_\delta(\mathbf{u}_\delta), \mathbf{P}_\int \mathbf{u}_\delta) \\ &= \varepsilon(\frac{\partial \mathbf{u}_\delta}{\partial t}, \mathbf{u}_\delta) + \gamma\varepsilon(\nabla \mathbf{u}_\delta, \nabla \mathbf{u}_\delta) + \frac{\gamma}{\varepsilon}(D\psi_\delta(\mathbf{u}_\delta), \mathbf{P}_\int \mathbf{u}_\delta). \end{array} \tag{27}$$

From (27) and (16) it follows for any constant $\boldsymbol{\xi} \in \mathbb{R}^N$ and for a.e. $t \in (0,T)$ that

$$\begin{array}{rcl} \frac{\gamma}{\varepsilon}(D\psi_\delta(\mathbf{u}_\delta), \boldsymbol{\xi} - \fint_\Omega \mathbf{u}_\delta) &=& \frac{\gamma}{\varepsilon}(D\psi_\delta(\mathbf{u}_\delta), \boldsymbol{\xi} - \mathbf{u}_\delta) - \gamma\varepsilon\|\nabla \mathbf{u}_\delta\|^2 - \varepsilon(\frac{\partial \mathbf{u}_\delta}{\partial t}, \mathbf{u}_\delta) \\ &\leq& \frac{\gamma}{\varepsilon}(\psi_\delta(\boldsymbol{\xi}) - \psi_\delta(\mathbf{u}_\delta), 1) + \frac{\gamma\sigma_{max}(\mathbf{A})}{2\varepsilon}\|\boldsymbol{\xi} - \mathbf{u}_\delta\|^2 + \varepsilon\|\frac{\partial \mathbf{u}_\delta}{\partial t}\|\|\mathbf{u}_\delta\|. \end{array}$$

Setting now $\boldsymbol{\xi} \equiv (\fint_\Omega \mathbf{u}_\delta) \pm \beta \mathbf{e}^n = \mathbf{m} \pm \beta \mathbf{e}^n$, where $\mathbf{e}^n$ is the n-th unit vector, $n = 1, ..., N$, and $\beta \in (0,1)$ such that $\beta \mathbf{1} < \fint_\Omega \mathbf{u}_0 = \mathbf{m} < \mathbf{1}$ we have $\hat{\psi}(\boldsymbol{\xi}) = 0$ and obtain

$$\begin{array}{rcl} \frac{\gamma}{\delta\varepsilon}(\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta), \pm\beta\mathbf{e}^n) &\leq& \frac{\gamma}{\varepsilon}(\mathbf{A}\mathbf{u}_\delta, \pm\beta\mathbf{e}^n) - \frac{\gamma}{2\varepsilon}(\mathbf{A}(\mathbf{m} \pm \beta\mathbf{e}^n), \mathbf{m} \pm \beta\mathbf{e}^n) + \frac{\gamma}{2\varepsilon}(\mathbf{A}\mathbf{u}_\delta, \mathbf{u}_\delta) \\ && -\frac{\gamma}{\delta\varepsilon}(\hat{\psi}(\mathbf{u}_\delta), 1) + \frac{\gamma|||\mathbf{A}|||}{2\varepsilon}\|\mathbf{m} \pm \beta\mathbf{e}^n - \mathbf{u}_\delta\|^2 + \varepsilon\|\frac{\partial \mathbf{u}_\delta}{\partial t}\|\|\mathbf{u}_\delta\| \\ &\leq& \frac{\beta\gamma|||\mathbf{A}|||\sqrt{|\Omega|}}{\varepsilon}\|\mathbf{u}_\delta\| + \frac{3\gamma|||\mathbf{A}|||}{2\varepsilon}\|\mathbf{m} \pm \beta\mathbf{e}^n\|^2 + \frac{3\gamma|||\mathbf{A}|||}{2\varepsilon}\|\mathbf{u}_\delta\|^2 \\ && +\varepsilon\|\frac{\partial \mathbf{u}_\delta}{\partial t}\|\|\mathbf{u}_\delta\| \end{array}$$

where we used that $\hat{\psi}(\mathbf{u}_\delta) \geq 0$. The above estimate gives the existence of a constant C which does not depend on δ such that for almost all $(t \in (0,T)$

$$\tfrac{1}{\delta}|\fint_\Omega \{(\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\}_n| \leq C(1 + \|\mathbf{u}_\delta\|^2 + \|\tfrac{\partial \mathbf{u}_\delta}{\partial t}\|\|\mathbf{u}_\delta\|)$$

for all $n = 1, ..., N$ and hence taking the $L^2(\Omega)$-norm of the constanct vector $\fint_\Omega \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)$ we obtain

$$\tfrac{1}{\delta}\|\fint_\Omega \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\| \leq C(1 + \|\mathbf{u}_\delta\|^2 + \|\tfrac{\partial \mathbf{u}_\delta}{\partial t}\|\|\mathbf{u}_\delta\|) \tag{28}$$

for almost all t, where C depends on $\varepsilon, \gamma, N, \mathbf{A}, |\Omega|$ and β but not on δ. Squaring (28) gives after integration over $t \in (0,T)$ and noting (19) that

$$\|\tfrac{1}{\delta}\fint_\Omega \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq C. \tag{29}$$

Combining (26) and (29) gives that

$$\tfrac{1}{\delta}\|\mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq C. \tag{30}$$

From (13) it follows that for almost all $(x,t) \in \Omega_T$ there exists an $n(x,t) \in \{1,...,N\}$ such that $\{\mathbf{u}_\delta(x,t)\}_{n(x,t)} \geq 0$ and hence $\{\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta(x,t))\}_{n(x,t)} = 0$. This implies

$$(\mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta(x,t)))_{n(x,t)} = -\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta(x,t))$$

and noting (30) we obtain that

$$\tfrac{1}{\delta}\|\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|_{L^2(0,T;L^2(\Omega))} \leq \tfrac{1}{\delta}\|\mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq C.$$

Together with (30) we obtain

$$\tfrac{1}{\delta}\|\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq \tfrac{1}{\delta}\|\mathbf{P}_\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + \tfrac{1}{\delta}\|\sum \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta)\mathbf{1}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq C.$$

We refer to [3] where similar arguments have been used in the context of a Cahn-Hilliard system with a logarithmic free energy. Finally, the fact that the $L^2(0,T;\mathbf{H}^2(\Omega))$-norm of $\mathbf{u}_\delta$ is uniformly bounded in δ, see (22), follows from (12), (19), (21) and by applying elliptic regularity theory, see [26], on time slices.

□

We are now in a position to show an existence and uniqueness theorem for the original problem.

Theorem 2.2 *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain and assume that either Ω is convex or fulfills $\partial\Omega \in C^{1,1}$. Let $\mathbf{u}_0 \in \mathbf{H}^1(\Omega)$ with $\mathbf{u}_0 \geq \mathbf{0}, \mathbf{0} < \fint_\Omega \mathbf{u}_0 = \mathbf{m} < \mathbf{1}$ and $\sum_{i=1}^N (u_0)_i = 1$ a.e. in Ω. Then there exists a unique solution $(\mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \Lambda)$ to (5) - (8) with the following properties:*

$$\mathbf{u}(x,0) = \mathbf{u}_0(x) \quad \text{for almost all } x \in \Omega, \tag{31}$$

$$\mathbf{u} \in L^\infty(0,T;\mathbf{H}^1(\Omega)) \cap H^1(0,T;\mathbf{L}^2(\Omega)) \cap L^2(0,T;\mathbf{H}^2(\Omega)), \tag{32}$$

$$\boldsymbol{\mu} \in L^2(0,T;\mathbf{L}^2(\Omega)), \tag{33}$$

$$\boldsymbol{\lambda} \in \mathbf{L}^2(0,T) \text{ and } \sum_{i=1}^N \lambda_i = 0 \quad \text{for almost all } t \in (0,T), \tag{34}$$

$$\Lambda \in L^2(0,T;L^2(\Omega)). \tag{35}$$

Proof: As the bounds (19), (21) and (22) are independent of δ, it follows that there exists a $\mathbf{u} \in L^\infty(0,T;\mathbf{H}^1(\Omega)) \cap H^1(0,T;\mathbf{L}^2(\Omega)) \cap L^2(0,T;\mathbf{H}^2(\Omega))$, and a subsequence $\{\mathbf{u}_{\delta'}\}$ which converges to $\mathbf{u}$ as $\delta' \to 0$

$$\begin{aligned}
&\text{a) in} \quad L^\infty(0,T;\mathbf{H}^1(\Omega)) \text{ weak-star,}\\
&\text{b) in} \quad H^1(0,T;\mathbf{L}^2(\Omega)) \cap L^2(0,T;\mathbf{H}^2(\Omega)) \text{ weakly,}\\
&\text{c) in} \quad L^2(0,T;\mathbf{L}^2(\Omega)) \text{ strongly,}\\
&\text{d) almost everywhere in} \quad \Omega \times (0,T)
\end{aligned} \tag{36}$$

where c) follows from a) and b), see [38].
Since $\mathbf{u}_\delta \in L^2(0,T;\mathbf{H}^2(\Omega))$ we can use the strong formulation of (12) and obtain

$$\varepsilon\tfrac{\partial \mathbf{u}_\delta}{\partial t} - \gamma\varepsilon\Delta\mathbf{u}_\delta - \tfrac{\gamma}{\varepsilon}\mathbf{A}\mathbf{u}_\delta - \tfrac{1}{\varepsilon}\boldsymbol{\mu}_\delta - \tfrac{1}{\varepsilon}\Lambda_\delta\mathbf{1} - \tfrac{1}{\varepsilon}\boldsymbol{\lambda}_\delta = \mathbf{0} \tag{37}$$

where $\Lambda_\delta := \gamma D\psi_\delta(\mathbf{u}_\delta) = \frac{\gamma}{\delta}\sum\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta) - \gamma\sum\mathbf{A}\mathbf{u}_\delta$, $\boldsymbol{\lambda}_\delta := \gamma\int_\Omega \mathbf{P}_\sum D\psi_\delta(\mathbf{u}_\delta) = \frac{\gamma}{\delta}\fint_\Omega \hat{\boldsymbol{\phi}}(\mathbf{u}_\delta) - \gamma\fint_\Omega \mathbf{A}\mathbf{u}_\delta - \fint_\Omega \Lambda_\delta\mathbf{1}$ and $\boldsymbol{\mu}_\delta := -\frac{\gamma}{\delta}\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta) \geq \mathbf{0}$.
Note that $\Lambda_\delta \in L^2(0,T;L^2(\Omega))$, $\boldsymbol{\lambda}_\delta \in \mathbf{L}^2(0,T)$ and $\boldsymbol{\mu}_\delta \in L^2(0,T;\mathbf{L}^2(\Omega))$ are uniformly bounded, see (19) and (21). Hence there exist $\Lambda \in L^2(0,T;L^2(\Omega)), \boldsymbol{\lambda} \in \mathbf{L}^2(0,T)$ and $\boldsymbol{\mu} \in L^2(0,T;\mathbf{L}^2(\Omega))$ such that for a subsequence

$$\begin{aligned}
\Lambda_\delta &\rightharpoonup \Lambda \quad && \text{in } L^2(0,T;L^2(\Omega)) \quad && \text{as } \delta' \to 0\\
\boldsymbol{\lambda}_\delta &\rightharpoonup \boldsymbol{\lambda} \quad && \text{in } \mathbf{L}^2(0,T;L^2(\Omega)) \quad && \text{as } \delta' \to 0\\
\boldsymbol{\mu}_{\delta'} &\rightharpoonup \boldsymbol{\mu} \quad && \text{in } L^2(0,T;\mathbf{L}^2(\Omega)) \quad && \text{as } \delta' \to 0\,.
\end{aligned}$$

Consequently, equation (5) holds for the limit. Furthermore, from (36), (13) and (14) it follows that $\sum_{i=1}^N u_i = 1$ and $\frac{d}{dt}\int_\Omega \mathbf{u} = \mathbf{0}$ and hence $\fint_\Omega \mathbf{u} = \mathbf{m}$. Taking the limit $\delta' \to 0$ in (20) gives that $[\mathbf{u}]_- = \mathbf{0}$ and thus $\mathbf{u}(x,t) \geq \mathbf{0}$ for *a.e.* $(x,t) \in \Omega_T$.
Since $\boldsymbol{\mu}$ is the weak limit of functions which are componentwise non-negative we obtain $\boldsymbol{\mu} \geq \mathbf{0}$ a.e. in Ω_T. In order to show that $(\boldsymbol{\mu},\mathbf{u}) = 0$ we first note that

$$(\boldsymbol{\mu}_\delta, \mathbf{u}_\delta) = -\tfrac{\gamma}{\delta}(\hat{\boldsymbol{\phi}}(\mathbf{u}_\delta), \mathbf{u}_\delta) \leq 0,$$

and using that $\mathbf{u}_\delta \to \mathbf{u}$ and $\boldsymbol{\mu}_\delta \rightharpoonup \boldsymbol{\mu}$ in $L^2(0,T,\mathbf{L}^2(\Omega))$ it follows that $(\boldsymbol{\mu},\mathbf{u}) \leq 0$. However, since $\mathbf{u} \geq \mathbf{0}$ and $\boldsymbol{\mu} \geq \mathbf{0}$ we have that $(\mathbf{u},\boldsymbol{\mu}) = 0$ a.e. in $(0,T)$. Note that $\sum_{i=1}^N (\lambda_\delta)_i = 0$ and hence $\sum_{i=1}^N \lambda_i = 0$, i.e. $\boldsymbol{\lambda} \in S$.

It remains to show uniqueness. Assume that there are two solutions $(\mathbf{u}^1,\boldsymbol{\mu}^1,\boldsymbol{\lambda}^1,\Lambda^1)$ and $(\mathbf{u}^2,\boldsymbol{\mu}^2,\boldsymbol{\lambda}^2,\Lambda^2)$. Then we define $\bar{\mathbf{u}} = \mathbf{u}^1 - \mathbf{u}^2, \bar{\boldsymbol{\mu}} = \boldsymbol{\mu}^1 - \boldsymbol{\mu}^2$. Multiplying the difference of the equation (5) for $\mathbf{u}^1$ and $\mathbf{u}^2$ with $\bar{\mathbf{u}}$ gives, after integration and using $\mathbf{1}\cdot\bar{\mathbf{u}} = 0$ and $\fint \mathbf{u} = \mathbf{0}$, that

$$\varepsilon\tfrac{d}{dt}\int_\Omega |\bar{\mathbf{u}}|^2 + \gamma\varepsilon\int_\Omega |\nabla\bar{\mathbf{u}}|^2 - \tfrac{1}{\varepsilon}\int_\Omega (\boldsymbol{\mu}^1 - \boldsymbol{\mu}^2)\cdot(\mathbf{u}^1 - \mathbf{u}^2) \leq \tfrac{\gamma\sigma_{max}(\mathbf{A})}{\varepsilon}\int_\Omega |\bar{\mathbf{u}}|^2. \tag{38}$$

The complementarity conditions (7)-(8) imply that $(\boldsymbol{\mu}^1 - \boldsymbol{\mu}^2)\cdot(\mathbf{u}^1 - \mathbf{u}^2) \leq 0$ and hence we deduce that

$$\varepsilon\tfrac{d}{dt}\int_\Omega |\bar{\mathbf{u}}|^2 + \gamma\varepsilon\int_\Omega |\nabla\bar{\mathbf{u}}|^2 \leq \tfrac{\gamma\sigma_{max}(\mathbf{A})}{\varepsilon}\int_\Omega |\bar{\mathbf{u}}|^2.$$

Using a Grönwall argument now gives uniqueness of $\mathbf{u}$. Hence $\boldsymbol{\mu}+\boldsymbol{\lambda}+\Lambda\mathbf{1}$ is uniquely given through equation (5).
Now we show the uniqueness of the Lagrange multipliers $\boldsymbol{\lambda}, \Lambda$ and $\boldsymbol{\mu}$. For what follows we fix $t \in (0,T)$ such that $\mathbf{u}(t) \in \mathbf{H}^2(\Omega)$ and define the inactive sets $I_i := \{x \in \Omega \mid u_i(x,t) > 0\}$, the interface between phases i and j as $I_{ij} := I_i \cap I_j$ and the measure $|I_{ij}|$ of I_{ij}.
We claim: $\lambda_i - \lambda_j$ is uniquely defined for all pairs (i,j) with $|I_{ij}| > 0$.
Using (5) and recalling that $\mathbf{e}^k$ is the k-th unit vector we obtain for (i,j) with $|I_{ij}| > 0$ that

$$(\varepsilon\tfrac{\partial \mathbf{u}}{\partial t} - \gamma\varepsilon\Delta\mathbf{u} - \tfrac{\gamma}{\varepsilon}\mathbf{A}\mathbf{u} - \tfrac{1}{\varepsilon}\boldsymbol{\lambda}) \cdot (\mathbf{e}^i - \mathbf{e}^j) = 0 \quad \text{on} \quad I_{ij} \tag{39}$$

where we have used that $\mu_i = \mu_j = 0$ on I_{ij}. We hence conclude

$$\lambda_i - \lambda_j = \tfrac{1}{|I_{ij}|} \int\limits_{I_{ij}} \left[\varepsilon^2 \tfrac{\partial u_i}{\partial t} - \gamma\varepsilon^2\Delta u_i - \gamma(\mathbf{A}\mathbf{u})_i - \varepsilon^2 \tfrac{\partial u_j}{\partial t} + \gamma\varepsilon^2\Delta u_j + \gamma(\mathbf{A}\mathbf{u})_j\right].$$

This implies that the difference $\lambda_i - \lambda_j$ is uniquely defined if there exists an interface between phases i and j, i.e. if $|I_{ij}| > 0$.
We now define a graph over $\{1,...,N\}$ with the edges $\mathcal{E} = \{\{i,j\} : |I_{ij}| > 0\}$. If the graph is connected, which we show in the following, the differences $\lambda_i - \lambda_j$ are for all $i,j \in \{1,...,N\}$ uniquely defined. Together with the condition $\sum\limits_{i=1}^{N} \lambda_i = 0$ we obtain the uniqueness of $\boldsymbol{\lambda}$.
In order to show that the graph is connected, we define the following sets of indices

$$\mathcal{L} = \{i \in \{1,...,N\} : \text{ there is a path from 1 to } i\} \quad \text{and} \quad \mathcal{M} = \{1,...,N\} \setminus \mathcal{L}.$$

We need to show that $\mathcal{M} = \emptyset$ and therefore we assume $\mathcal{M} \neq \emptyset$. We set

$$v = \sum_{i \in \mathcal{L}} u_i \quad \text{and} \quad w = \sum_{j \in \mathcal{M}} u_j$$

and note that $v \geq 0, w \geq 0$ and $v + w = 1$. Now one observes that the set

$$A := \{x \in \Omega : v(x) > 0 \text{ and } w(x) > 0\}$$

has measure zero. This is true because otherwise there exist $i \in \mathcal{L}$ and $j \in \mathcal{M}$ such that $|I_{ij}| > 0$ which contradicts the definition of $\mathcal{L}$ and the assumption $\mathcal{M} \neq \emptyset$. We hence obtain that v only attains the values 0 and 1. Since $\mathcal{M} \neq \emptyset$ we obtain that v is not constant. Since an H^1-function that attains finitely many values has to be constant we obtain a contradiction. Hence $\mathcal{M} = \emptyset$ and the graph is connected.
Now we show the uniqueness of Λ. Since $\sum_{i=1}^{N} u_i = 1$ and $\mathbf{u} \geq \mathbf{0}$, we can find for any $x_0 \in \Omega$ an $i \in \{1,...,N\}$ such that $x_0 \in I_i$ and $|I_i| > 0$. On I_i we know that $\mu_i = 0$ and hence we can define

$$\Lambda = (\varepsilon^2 \tfrac{\partial \mathbf{u}}{\partial t} - \gamma\varepsilon^2\Delta\mathbf{u} - \gamma\mathbf{A}\mathbf{u} - \boldsymbol{\lambda})_i \quad \text{on} \quad I_i. \tag{40}$$

The Lagrange multiplier Λ is well defined and unique since

$$(\varepsilon\tfrac{\partial \mathbf{u}}{\partial t} - \gamma\varepsilon\Delta\mathbf{u} - \tfrac{\gamma}{\varepsilon}\mathbf{A}\mathbf{u} - \tfrac{1}{\varepsilon}\boldsymbol{\lambda})_i(x,t) = (\varepsilon\tfrac{\partial \mathbf{u}}{\partial t} - \gamma\varepsilon\Delta\mathbf{u} - \tfrac{\gamma}{\varepsilon}\mathbf{A}\mathbf{u} - \tfrac{1}{\varepsilon}\boldsymbol{\lambda})_j(x,t)$$

for almost every $x \in I_{ij}$ and since for almost all $x \in \Omega$, there is an $i \in \{1, \dots, N\}$ such that $u_i(x,t) > 0$, i.e. $x \in I_i$. Having shown uniqueness of $\mathbf{u}, \boldsymbol{\lambda}$ and Λ uniqueness of $\boldsymbol{\mu}$ follows from equation (5).

□

Remark 2.1 *i) It can be shown that a solution to* $(\mathbf{P_m})$ *is unique. Hence we can conclude that for all solutions to* $(\mathbf{P_m})$ *there exist Lagrange multipliers* $\boldsymbol{\mu}, \boldsymbol{\lambda}, \Lambda$ *such that (5) - (8) hold. Furthermore, problem (5)-(8) is equivalent to* $(\mathbf{P_m})$.

ii) In [35] the existence of a solution for vector-valued Allen-Cahn variational inequalities without volume constraints is shown by using a representation of the Lagrange multipliers which cannot be used directly for a numerical approach. There u_N *is substituted by* $1 - \sum_{i=1}^{N-1} u_i$ *and a system of parabolic variational inequalities in* $\mathbb{R}^{N-1}$ *is considered.*

3 Primal-dual active set approach

For the numerical approximation of solutions $\mathbf{u}$ of $(\mathbf{P_m})$ we introduce a primal-dual active set method or equivalently a semi-smooth Newton method [7, 27]. Both are well known in the context of optimization with partial differential equations as constraints. We present a time discretization of the Allen-Cahn system and reformulate the complementarity conditions using primal-dual active sets. Finally, even though the method is not applicable to the time discretized problem, we present for ease of understanding the idea of the resulting iterative solution procedure for the time discretized problem, which will be applied to the fully discretized problem in the next section.
We denote the time step by τ, which can be a variable time step, $t_0 = 0$, $t_n := t_{n-1} + \tau$ and $\mathbf{u}^{n-1} := \mathbf{u}(., t_{n-1})$. For simplicity we denote by $\mathbf{u}$ the time discrete solution at time t_n. Then possible time discretizations of $(\mathbf{P_m})$ are given as (semi-)implicit or explicit Euler-discretizations. Explicit Euler discretizations for Allen-Cahn obstacle problems have been used for example in [12, 20, 23, 24]. Numerical analysis for (semi-) implicit discretizations of the Allen-Cahn model has been performed in the papers [15, 22, 31, 32, 33, 34] and in works cited in these papers. Fully implicit discretizations are the most accurate, see e.g. [9], but due to the non-monotonicity for large time steps they can be either very expansive or they are not uniquely solvable. It will turn out that this is not the case for the primal dual active set approach as for well developed interfaces also larger time steps can be used, see e.g. Remark 4.3. In this paper we focus on the implicit discretization of the vector-valued Allen-Cahn obstacle problem leading to the following formulation:

$(\mathbf{P}^\tau_{\mathbf{m}})$ *Given* $\mathbf{u}^{n-1} \in \mathcal{G}^{\mathbf{m}}$ *find* $\mathbf{u} = \mathbf{u}^n \in \mathcal{G}^{\mathbf{m}}$ *such that*

$$\tfrac{\varepsilon}{\tau}(\mathbf{u} - \mathbf{u}^{n-1}, \boldsymbol{\eta} - \mathbf{u}) + \gamma\varepsilon(\nabla\mathbf{u}, \nabla\boldsymbol{\eta} - \nabla\mathbf{u}) \geq \tfrac{\gamma}{\varepsilon}(\mathbf{A}\mathbf{u}, \boldsymbol{\eta} - \mathbf{u}) \quad \forall \boldsymbol{\eta} \in \mathcal{G}^{\mathbf{m}}. \tag{41}$$

This discretization can also be seen as the Euler-Lagrange equation of an implicit time discretization of the L_2 gradient flow of the energy E, which is given as

$$\min_{\mathbf{u}\in\mathcal{G}^{\mathbf{m}}} \mathcal{E}(\mathbf{u}) := \int_\Omega \left\{ \tfrac{\gamma\varepsilon}{2}|\nabla\mathbf{u}|^2 + \tfrac{\gamma}{\varepsilon}\psi(\mathbf{u}) \right\} dx + \tfrac{\varepsilon}{2\tau}\|\mathbf{u} - \mathbf{u}^{n-1}\|^2_{\mathbf{L}^2}. \tag{42}$$

As in Lemma 2.1 one can reformulate $(\mathbf{P}^\tau_{\mathbf{m}})$ by using scaled Lagrange-multipliers $\boldsymbol{\mu} \in \mathbf{L}^2(\Omega)$ for the inequality constraint $\mathbf{u} \geq \mathbf{0}$, $\boldsymbol{\lambda} \in S$ for the constraint $\mathbf{P}_S(\fint_\Omega \mathbf{u} - \mathbf{m}) = \mathbf{0}$ and $\Lambda \in L^2(\Omega)$ for the sum constraint $\sum_{i=1}^N u_i = 1$ to obtain:

$$\begin{aligned}
\tfrac{\varepsilon^2}{\tau}(\mathbf{u} - \mathbf{u}^{n-1}) - \gamma\varepsilon^2\Delta\mathbf{u} - \gamma\mathbf{A}\mathbf{u} - \boldsymbol{\mu} - \boldsymbol{\lambda} - \Lambda\mathbf{1} &= \mathbf{0} \qquad \text{a.e. in } \Omega\,, && (43)\\
\tfrac{\partial\mathbf{u}}{\partial\nu} &= \mathbf{0} \qquad \text{a.e. on } \partial\Omega\,, && (44)\\
\mathbf{P}_S(\fint_\Omega \mathbf{u} - \mathbf{m}) &= \mathbf{0} && (45)\\
\sum_{i=1}^N u_i &= 1 && (46)
\end{aligned}$$

together with the complementarity conditions

$$\mathbf{u} \geq \mathbf{0} \quad \text{a.e. in } \Omega, \quad \boldsymbol{\mu} \geq \mathbf{0} \quad \text{a.e. in } \Omega, \quad (\boldsymbol{\mu}, \mathbf{u}) = 0. \tag{47}$$

Now the idea is to reformulate the complementarity conditions using active sets based on the primal variable $\mathbf{u}$ and the dual variable $\boldsymbol{\mu}$. Then, for any $c > 0$, (47) is equivalent to the following: For all $i \in \{1, ..., N\}$:

$$u_i = 0 \ \text{ a.e. in } \mathcal{A}_i; \qquad \mu_i = 0 \ \text{ a.e. in } \mathcal{I}_i := \Omega \setminus \mathcal{A}_i; \tag{48}$$

where the primal-dual active sets are given by

$$\mathcal{A}_i = \{x \in \Omega \mid cu_i(x) - \mu_i(x) < 0\}\,. \tag{49}$$

If the sets $\mathcal{A}_i$ are known, we can determine the solution as follows. First we set $u_i = 0$ on $\mathcal{A}_i$ and $\mu_i = 0$ on $\mathcal{I}_i$, see (48). It now turns out that u_i only needs to be determined in points in $\mathcal{I}_i$ in which another component is inactive, i.e. on

$$\mathcal{D}_i = \mathcal{I}_i \cap (\bigcup_{j\neq i} \mathcal{I}_j)\,.$$

On $\mathcal{I}_i \setminus \mathcal{D}_i$ we observe that i is the only inactive component and hence the constraint $\sum_{j=1}^N u_j = 1$ leads to

$$u_i = 1 \text{ on } \mathcal{I}_i \setminus \mathcal{D}_i\,. \tag{50}$$

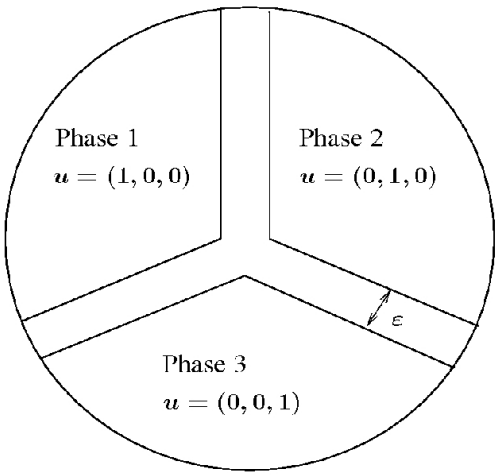


Figure 1: The computational effort in the primal-dual active set method is restricted to the diffuse interface.

Defining now the total diffuse interface region, see Figure 1, as

$$\mathcal{D} := \bigcup_{i=1}^{N} \mathcal{D}_i \supseteq \{x \in \Omega \mid \quad \text{there exists } i \in \{1, \dots, N\} \text{ s.t. } 0 < u_i(x) < 1\}$$

we need to solve the following system: Find, for $i = 1, \dots, N$, u_i on $\mathcal{D}_i$, Λ on $\mathcal{D}$, $\boldsymbol{\lambda} \in \mathbb{R}^N$ such that for $i = 1, \dots, N$

$$-\gamma\varepsilon^2 \Delta u_i + \tfrac{\varepsilon^2}{\tau}(u_i - u_i^{n-1}) - \lambda_i - \Lambda - \gamma(\mathbf{A}\mathbf{u})_i = 0, \quad \text{a.e. on } \mathcal{D}_i\,, \tag{51}$$

$$\frac{\partial u_i}{\partial \nu} = 0 \quad \text{a.e. on } \partial\mathcal{D}_i \cap \partial\Omega, u_i = 0 \text{ on } \partial\mathcal{D}_i \cap \partial\mathcal{A}_i\,, \tag{52}$$

$$\sum_{i=1}^{N} u_i = 1 \text{ on } \mathcal{D} \text{ and } \mathbf{P}_S(\fint_{\Omega} \mathbf{u} - \mathbf{m}) = \mathbf{0}, \sum_{i=1}^{N} \lambda_i = 0. \tag{53}$$

Then we have to determine Λ on $\Omega \setminus \mathcal{D}$. Since in each point $x \in \Omega$ at least one component is inactive there exists for a given $x \in \Omega \setminus \mathcal{D}$ an i such that $x \in \mathcal{I}_i \setminus \mathcal{D}$ and given $\mu_i = 0$ on $\mathcal{I}_i$ we set

$$\Lambda = -\gamma\varepsilon^2 \Delta u_i + \tfrac{\varepsilon^2}{\tau}(u_i - u_i^{n-1}) - \lambda_i - \gamma(\mathbf{A}\mathbf{u})_i\,. \tag{54}$$

Then Λ is completely determined. Finally we set

$$\mu_i = -\gamma\varepsilon^2 \Delta u_i + \tfrac{\varepsilon^2}{\tau}(u_i - u_i^{n-1}) - \lambda_i - \gamma(\mathbf{A}\mathbf{u})_i - \Lambda \text{ on } \mathcal{A}_i\,. \tag{55}$$

This leads to the idea of the **Primal-Dual Active Set (PDAS) algorithm**:
Given initial active sets $\mathcal{A}_i^0$ for all $i \in \{1, ..., N\}$ iterate the following steps for $k \geq 0$ (where we define $I_i^k, \mathcal{D}_i^k, \mathcal{D}^k$ analogous to the discussion above)

i) *Set* $u_i^k = 0$ *on* $\mathcal{A}_i^k$ *and* $\mu_i^k = 0$ *on* $\mathcal{I}_i^k$ *for all* $i \in \{1, ..., N\}$, $u_i^k = 1$ *on* $\mathcal{I}_i^k \setminus \mathcal{D}_i^k$.

ii) *Solve the coupled system (51)-(53) for* $\boldsymbol{\lambda}^k$, Λ^k *on* $\mathcal{D}$ *and* u_i^k *on* $\mathcal{D}_i^k$ *for all* $i \in \{1, ..., N\}$.

iii) *Determine* Λ^k *on* $\Omega \setminus \mathcal{D}$ *using (54).*

iv) *Determine* μ_i^k *on* $\mathcal{A}_i^k$ *using (55) for all* $i \in \{1, ..., N\}$.

v) *Determine the new active sets* $\mathcal{A}_i^{k+1} = \{x \in \Omega \,|\, cu_i^k(x) - \mu_i^k(x) < 0\}$ *for all* $i \in \{1, ..., N\}$.

vi) *Stop the iteration if* $\mathcal{A}_i^{k+1} = \mathcal{A}_i^k$ *for all* $i \in \{1, ..., N\}$, *otherwise set* $k = k + 1$ *and goto 1.*

Except for the sign conditions for $\mathbf{u}$ and $\boldsymbol{\mu}$ all conditions (44)-(48) hold in each iteration. As mentioned in the beginning of this section we cannot apply the method to the time discretized Allen-Cahn variational inequality. The reason is that although one can show the existence of the Langrange-multipliers and the regularity $\boldsymbol{\mu} \in \mathbf{L}^2(\Omega)$ this regularity does in general not hold in each iteration of the PDAS-algorithm. Then the multipliers may still exist but are only measures. This effect is also known for obstacle problems, see [29], and is discussed in more detail for Cahn-Hilliard problems in [8]. Therefore, the pointwise definition of the active sets $\mathcal{A}_i^k$ is not possible. However we show in the next section that the application of the PDAS-method to the fully discretized problem is possible.
The feature that the method is not applicable in function space may lead to mesh dependence of the PDAS iteration numbers for a fixed time step. Analysis for mesh independence is still lacking and needs further research. However, our numerical investigations clearly indicate mesh independence if the time and space discretizations are reduced simultaneously, see Section 5.

4 Finite element approximation

For space discretization we employ a finite element approximation which we present in this section. Furthermore, we present the PDAS-algorithm for the fully discretized system.

4.1 Notation

For simplicity we assume that Ω is a polyhedral domain. Let $\mathcal{T}_h$ be a regular triangulation of Ω into disjoint open simplices, i.e. in particular $\Omega = \cup_{T \in \mathcal{T}_h} \overline{T}$. Furthermore, we define $h := \max_{T \in \mathcal{T}_h} \{\text{diam } T\}$ to be the maximal element size of $\mathcal{T}_h$ and we set

$\mathcal{J}$ to be the set of nodes of $\mathcal{T}_h$ and $\{p_j\}_{j\in\mathcal{J}}$ to be the coordinates of these nodes. Associated with $\mathcal{T}_h$ is the piecewise linear finite element space

$$S_h := \left\{\varphi \in C^0(\overline{\Omega}) \Big| \, \varphi\big|_T \in P_1(T) \ \forall\, T \in \mathcal{T}_h\right\} \subset H^1(\Omega),$$

where we denote by $P_1(T)$ the set of all affine linear functions on T. Furthermore we denote the standard nodal basis functions of S_h by χ_j for all $j \in \mathcal{J}$ and we set $\mathbf{S}_h = (S_h)^N$. Then $\boldsymbol{u}_j \in \mathbb{R}^N$ for $j \in \mathcal{J}$ denotes the coefficients of the basis representation of $\mathbf{u}_h$ in $\mathbf{S}_h$ which is given by $\mathbf{u}_h = \sum_{j\in\mathcal{J}} \boldsymbol{u}_j\chi_j$.
In order to derive a discretization of the Allen-Cahn model we set

$$\mathcal{G}_h := \{\boldsymbol{\chi} \in \mathbf{S}_h |\ \boldsymbol{\chi} \geq \mathbf{0} \text{ and } \sum_{i=1}^{N}(\chi_i)_j = 1 \ \ \forall j \in \mathcal{J}\}$$

and

$$\mathcal{G}_h^{\mathbf{m}} := \{\boldsymbol{\eta} \in \mathcal{G}_h |\ \fint_\Omega \boldsymbol{\eta} = \mathbf{m}\}.$$

Here $(\chi_i)_j$ denotes the i-th component χ_i of $\boldsymbol{\chi}$ at the j-th node. We introduce also the lumped mass semi-inner product $(f,g)_h = \int_\Omega I_h(fg)$ instead of (f,g), where $I_h : C^0(\overline{\Omega}) \to S_h$ is the standard interpolation operator such that $(I_h\, f)(p_j) = f(p_j)$ for all nodes $j \in \mathcal{J}$.
Defining $m_j := (1, \chi_j)$ we have $\fint_\Omega u_i = \sum_{j\in\mathcal{J}} m_j(u_i)_j / \sum_{j\in\mathcal{J}} m_j$ where $u_i \in S^h$ and $i \in \{1, ..., N\}$. Moreover, we define the stiffness matrix as $\boldsymbol{S} := (s_{ij})$ with $s_{ij} = (\nabla\chi_j, \nabla\chi_i)$, the mass matrix $\boldsymbol{M} := ((\chi_j, \chi_i)_h) = diag(m_j)$ and the mass vector $\boldsymbol{m} := (m_j)_{j\in\mathcal{J}}$. Also we denote the entries of $\mathbf{A}$ by a_{ij}, $i, j = \{1, \ldots, N\}$. Recall that the total spatial amount to be conserved is denoted by $\mathbf{m} = (m^i)_{i=1,\ldots,N}$ which should not be confused with the mass vector $\boldsymbol{m}$.

4.2 Finite element approximation and the PDAS-algorithm

We now introduce the following finite element approximations of $(\mathbf{P}^\tau_{\mathbf{m}})$ given by (41). In the following we consider a fixed time step $\tau = t_n - t_{n-1}$ and omit in some places the superscript n:
$(\mathbf{P}^\tau_{\mathbf{m},h})$ *Given* $\mathbf{u}_h^{n-1} \in \mathcal{G}_h^{\mathbf{m}}$ *find* $\mathbf{u}_h = \mathbf{u}_h^n \in \mathcal{G}_h^{\mathbf{m}}$ *such that*

$$(\tfrac{\varepsilon}{\tau}(\mathbf{u}_h - \mathbf{u}_h^{n-1}) - \tfrac{\gamma}{\varepsilon}\mathbf{A}\mathbf{u}_h, \boldsymbol{\chi} - \mathbf{u}_h)_h + \gamma\varepsilon(\nabla\mathbf{u}_h, \nabla(\boldsymbol{\chi} - \mathbf{u}_h)) \geq 0 \ \ \forall \boldsymbol{\chi} \in \mathcal{G}_h^{\mathbf{m}}. \tag{56}$$

Due to the use of piecewise linear finite elements and nodal basis functions the reformulation of $(\mathbf{P}^\tau_{\mathbf{m},h})$ with Lagrange multipliers $\boldsymbol{\mu}_h \in \mathbf{S}_h$ can be stated as follows:

$(\mathbf{Q}^{\tau}_{\mathbf{m},h})$ *Find* $\mathbf{u}_h \in \mathbf{S}_h$, $\boldsymbol{\mu}_h \in \mathbf{S}_h$, $\boldsymbol{\lambda}_h \in \mathbb{R}^N$ *and* $\Lambda_h \in S_h$ *such that*

$$\begin{aligned}\tfrac{\varepsilon^2}{\tau}(\mathbf{u}_h,\boldsymbol{\varphi})_h - \gamma(\mathbf{A}\mathbf{u}_h,\boldsymbol{\varphi})_h + \gamma\varepsilon^2(\nabla\mathbf{u}_h,\nabla\boldsymbol{\varphi}) - (\boldsymbol{\mu}_h,\boldsymbol{\varphi})_h - (\boldsymbol{\lambda}_h,\boldsymbol{\varphi}) - (\Lambda_h\mathbf{1},\boldsymbol{\varphi})_h \\ = \tfrac{\varepsilon^2}{\tau}(\mathbf{u}_h^{n-1},\boldsymbol{\varphi})_h \quad \forall \boldsymbol{\varphi} \in \mathbf{S}_h,\end{aligned} \tag{57}$$

$$\sum_{i=1}^{N}(u_i)_j = 1 \quad \forall\, j \in \mathcal{J}, \tag{58}$$

$$\mathbf{P}_S\Big(\sum_{j\in\mathcal{J}} m_j\boldsymbol{u}_j - \sum_{j\in\mathcal{J}} m_j\mathbf{m}\Big) = \mathbf{0}, \tag{59}$$

$$\sum_{i=1}^{N}\lambda_i = 0, \tag{60}$$

$$\boldsymbol{\mu}_j \geq \mathbf{0},\ \boldsymbol{u}_j \geq \mathbf{0} \quad \forall\, j \in \mathcal{J}, \qquad (\boldsymbol{u}_h,\boldsymbol{\mu}_h)_h = 0. \tag{61}$$

In the following we eliminate λ_N by using (60) and hence obtain $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_{N-1}, -\lambda_1 - ... - \lambda_{N-1})^T$. Using $\mathbf{u} \in \mathcal{G}^m$ and $\sum\limits_{i=1}^{N} m^i = 1$ we obtain that $\mathbf{P}_S$ in (59) can be replaced by the identity. To avoid redundancy in (59) we use (58) and drop the condition on $i = N$ and now obtain symmetry in the system by restating (59) as

$$\sum_{j\in\mathcal{J}} m_j((u_i)_j - (u_N)_j) = \sum_{j\in\mathcal{J}} m_j(m^i - m^N) \quad \text{for} \quad i \in \{1, ..., N-1\}. \tag{62}$$

Applying the PDAS-method presented in Section 3 to $(\mathbf{Q}^{\tau}_{\mathbf{m},h})$ we obtain the following algorithm. Here we use the notation $\mathbf{u}_i^k$ and $\mathbf{u}_i^{n-1}$ where the k denotes the k-th iteration in the PDAS algorithm and $n-1$ is the $(n-1)$-st time step. This is of course a misuse of notation for $k = n-1$.

Primal-Dual Active Set Algorithm (PDAS-I):

0. Set $k = 0$ and initialize $\mathcal{A}_i^0 \subset \mathcal{J}$ for all $i \in \{1, ..., N\}$.

1. Define $\mathcal{I}_i^k = \mathcal{J} \setminus \mathcal{A}_i^k$ for all $i \in \{1, ..., N\}$.
   Set $(u_i^k)_j = 0$ for $j \in \mathcal{A}_i^k$, $(u_i^k)_j = 1$ for $j \in \mathcal{I}_i^k \setminus \mathcal{D}_i^k$ and $(\mu_i^k)_j = 0$ for $j \in \mathcal{I}_i^k$ for all $i \in \{1, ..., N\}$.

2. Set $\mathcal{D}_i^k := \mathcal{I}_i^k \cap (\bigcup\limits_{\substack{j=1\\ j\neq i}}^{N} \mathcal{I}_j^k)$, $\mathcal{D}^k := \bigcup\limits_{i=1}^{N} \mathcal{D}_i^k$. Solve the discretized PDE (57) on the interface $\mathcal{D}^k$ with the constraints (58), (60), (62) to obtain $(u_i^k)_j$ for all (i,j) such that $j \in \mathcal{D}_i^k, i \in \{1, ..., N\}$ and Λ_j^k for all $j \in \mathcal{D}^k$ and λ_i^k for all

$i \in \{1, ..., N\}$. More precisley we solve

$$\frac{\varepsilon^2}{\tau}(u_i^k)_j - \gamma \sum_{m=1}^{N} a_{im}(u_m^k)_j + \frac{\gamma\varepsilon^2}{m_j} \sum_{l\in\mathcal{J}} s_{lj}(u_i^k)_l - \lambda_i^k - \Lambda_j^k = \frac{\varepsilon^2}{\tau}(u_i^{n-1})_j \qquad (63)$$

$$\text{for } j \in \mathcal{D}_i^k \quad \text{and } i \in \{1, \dots, N\}\,,$$

$$\sum_{j\in\mathcal{J}} m_j((u_i^k)_j - (u_N^k)_j) = \sum_{j\in\mathcal{J}} m_j(m^i - m^N) \quad \text{for } i = 1, ..., N-1\,, \qquad (64)$$

$$\sum_{i=1}^{N} (u_i^k)_j = 1 \quad \text{for } j \in \mathcal{D}^k \qquad (65)$$

where $\lambda_N^k = -\lambda_1^k - ... - \lambda_{N-1}^k$ is used in (63).

3. Define Λ_j^k for all $j \in \mathcal{I}_i^k \setminus \mathcal{D}^k$ as

$$\Lambda_j^k = \frac{\varepsilon^2}{\tau}(u_i^k)_j - \gamma \sum_{m=1}^{N} a_{im}(u_m^k)_j + \frac{\gamma\varepsilon^2}{m_j} \sum_{l\in\mathcal{J}} s_{lj}(u_i^k)_l - \lambda_i^k - \frac{\varepsilon^2}{\tau}(u_i^{n-1})_j\,.$$

4. Determine $(\mu_i^k)_j$ for $j \in \mathcal{A}_i^k$ using (57) for all $i = 1, ..., N$ as

$$(\mu_i^k)_j = \frac{\varepsilon^2}{\tau}(u_i^k)_j - \gamma \sum_{m=1}^{N} a_{im}(u_m^k)_j + \frac{\gamma\varepsilon^2}{m_j} \sum_{l\in\mathcal{J}} s_{lj}(u_i^k)_l - \lambda_i^k - \Lambda_j^k - \frac{\varepsilon^2}{\tau}(u_i^{n-1})_j\,.$$

5. Set $\mathcal{A}_i^{k+1} := \{j \in \mathcal{J} : c(u_i^k)_j - (\mu_i^k)_j < 0\}$ for $i = 1, ..., N$.

6. If $\mathcal{A}_i^{k+1} = \mathcal{A}_i^k$ for all $i \in \{1, ..., N\}$ stop, otherwise set $k = k + 1$ and goto 1.

Remark 4.1

i) In each iteration all conditions (57)-(60) hold, but the sign conditions in (61) may not be true.

ii) In the above algorithm only the equations (63)-(65) require the solution of a sparse linear system – in all other steps of the algorithm simple assignments take place. We remark that (63)- (65) is a linear system with degrees of freedom on the diffuse interface only, see Figure 1.

iii) In each node p_j for $j \in \mathcal{J}$ some components of $\mathbf{u}_h$ are active and the others are inactive. The number of components which are active can vary from point to point. Only for each individual component u_i we can split the set of nodes into nodes which are active and inactive for this component. This results in a quite complex linear system (63)-(65).

4.3 Convergence as a semi-smooth Newton method

In order to show that the PDAS method converges, we reformulate the method as a semi-smooth Newton method. Instead of introducing active and inactive sets we can reformulate (61) using the semi-smooth function $\mathcal{H}(y,z) := z - \max(0, z - cy)$ with $c > 0$ as follows

$$\mathcal{H}((u_i)_j, (\mu_i)_j) = 0 \tag{66}$$

for $i \in \{1, ..., N\}$ and $j \in \mathcal{J}$.
The mapping $y \mapsto \max(0, y)$ from $\mathbb{R}$ to $\mathbb{R}$ is slantly differentiable and a possible slanting function is given by $G : \mathbb{R} \to \mathbb{R}$ with $G(y) = 1$ for $y > 0$ and $G(y) = 0$ for $y \leq 0$, see [27]. Hence for the above function $\mathcal{H}$ one derives the slanting function $\mathcal{G} : \mathbb{R}^2 \to \mathbb{R}^2$ with $\mathcal{G}(y,z) = (c, 0)$ if $z - cy > 0$ and $\mathcal{G}(y,z) = (0, 1)$ if $z - cy \leq 0$. We hence obtain that

$$\mathcal{G}(\overline{y}, \overline{z}) \begin{pmatrix} y - \overline{y} \\ z - \overline{z} \end{pmatrix} = -\mathcal{H}(\overline{y}, \overline{z}) \tag{67}$$

is equivalent to

$$y = 0 \quad \text{if} \quad \overline{z} - c\overline{y} > 0 \quad \text{and} \quad z = 0 \quad \text{if} \quad \overline{z} - c\overline{y} \leq 0\,. \tag{68}$$

The equation (67) will later be part of one Newton step for the overall system. We now consider the system (57)-(60) together with the semi-smooth equation (66) as one large algebraic system of the form $\mathbf{F}(\boldsymbol{\eta}) = \mathbf{F}(\mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Lambda}) = \mathbf{0}$ where $\boldsymbol{\eta} := (\mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Lambda})$, $\mathbf{u} := (\mathbf{u}_1, \ldots, \mathbf{u}_N)$, $\boldsymbol{\mu} := (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_N)$ and $\mathbf{u}_i, \boldsymbol{\mu}_i, \boldsymbol{\Lambda}$ are the coefficient vectors of u_i, μ_i and Λ. Hence the system is defined in an Euclidean space of dimension $2N|\mathcal{J}| + N + |\mathcal{J}|$. We now use a semi-smooth Newton method (SSN) for the equation and consider

$$\mathbf{G}(\boldsymbol{\eta}^{k-1})(\boldsymbol{\eta}^k - \boldsymbol{\eta}^{k-1}) = -\mathbf{F}(\boldsymbol{\eta}^{k-1}) \tag{69}$$

where $\mathbf{G}$ is the slanting function of $\mathbf{F}$ using $\mathcal{G}$. One observes that the equation in (69) related to (66) leads to $(u_i^k)_j = 0$ if $(\mu_i^{k-1})_j - c(u_i^{k-1})_j > 0$ and $(\mu_i^k)_j = 0$ if $(\mu_i^{k-1})_j - c(u_i^{k-1})_j \leq 0$, compare (67), (68).

Remark 4.2 *It is straightforward to show that the semi-smooth Newton method (SSN) is equivalent to the discrete primal-dual active set algorithm (PDAS-I), see e.g. [27] for a similar situation.*

In order to show local convergence of the semi-smooth Newton method we need to show invertibility of $\boldsymbol{G}$ in some neighborhood of a solution to $\boldsymbol{F}(\boldsymbol{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Lambda}) = \mathbf{0}$.
To proceed we need a discrete Poincaré inequality in a situation where we have for $i = 1, \ldots, N$ given inactive sets $\mathcal{I}_i$ and corresponding active sets $\mathcal{A}_i := \mathcal{J} \setminus \mathcal{I}_i$. There exists a Poincaré constant $c_h^p(\boldsymbol{K}) > 0$ such that

$$(\mathbf{v}, \mathbf{v})_h \leq c_h^p(\nabla \mathbf{v}, \nabla \mathbf{v}) \quad \forall\, \mathbf{v} \in \boldsymbol{K} \tag{70}$$

with $\boldsymbol{K} := \{\mathbf{v} \in \mathbf{S}_h \mid \int\limits_\Omega \mathbf{v} = \mathbf{0}, \sum\limits_{i=1}^N (v_i)_j = 0 \ \forall j \in \mathcal{J},\ v_i(p_j) = 0 \text{ if } j \in \mathcal{A}_i \text{ and } i = 1, ..., N\}$, see e.g. [2] or [21]. We remark here that the typical situation in applications is that the interfacial region has order ε. In conclusion, it was discussed in [9] that in fact only functions with thin support need to be considered for the Poincaré inequality which leads to good Poincaré constants. As in [9] we can also conclude here that this leads to situations in which we can solve the discrete Allen-Cahn system (63)-(65) also for large time steps, see [9] for more details.
To show invertibility we need in addition a discrete analogue of the graph theoretic argument used in the proof of Theorem 2.2. Assume that inactive sets $\mathcal{I}_i$ for $i = 1, \ldots, N$ are given. We then choose a graph over $\{1, ..., N\}$ with edges

$$\mathcal{E} = \{\{i, l\} : \ \mathcal{I}_i \cap \mathcal{I}_l \neq \emptyset\}.$$

Theorem 4.1 *Assume that given inactive sets $\mathcal{I}_i$, $i = 1, \ldots, N$ are connected in the sense that the corresponding graph is connected and assume in addition that $\mathcal{J} = \bigcup\limits_{i=1}^N \mathcal{I}_i$. Assume furthermore*

$$\tau(\sigma_{max}(\mathbf{A}) - \tfrac{\varepsilon^2}{c_h^p(\boldsymbol{K})}) < \tfrac{\varepsilon^2}{\gamma} \tag{71}$$

where $c_h^p(\boldsymbol{K})$ is the Poincaré constant given by (70). Then the linear mapping $\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Lambda})$ is invertible which is equivalent to the unique solvability of (63)- (65).

Proof: We show that the kernel of $\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Lambda})$ contains only $\mathbf{0}$. The equation

$$\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Lambda})(\boldsymbol{v}, \boldsymbol{\kappa}, \boldsymbol{\alpha}, \boldsymbol{\beta})^t = \mathbf{0} \tag{72}$$

is equivalent to

$$\tfrac{\varepsilon^2}{\tau}(v_i)_j - \gamma \sum_{m=1}^N a_{im}(v_m)_j + \tfrac{\gamma\varepsilon^2}{m_j} \sum_{l \in \mathcal{J}} s_{lj}(v_i)_l - (\kappa_i)_j - \alpha_i - \beta_j = 0 \text{ for } j \in \mathcal{J}, \tag{73}$$

$$\sum_{j \in \mathcal{J}} m_j(v_i)_j = 0, \qquad \sum_{i=1}^N (v_i)_j = 0 \text{ for } j \in \mathcal{J}, \qquad \sum_{i=1}^N \alpha_i = 0, \tag{74}$$

$$(v_i)_j = 0 \text{ for } j \in \mathcal{A}_i \text{ and } (\kappa_i)_j = 0 \text{ for } j \in \mathcal{I}_i \tag{75}$$

which has to hold for all $i \in \{1, ..., N\}$.
The equations in (73) related to inactive $(i, j) \in \{1, \ldots, N\} \times \mathcal{J}$ together with (74) are the first order necessary conditions of the quadratic optimization problem in $\mathbf{v}$, where $\mathbf{v}$ is the coefficient vector $\boldsymbol{v}$, on the set $\boldsymbol{K}$

$$\min_{\mathbf{v} \in \boldsymbol{K}} \left[\tfrac{\varepsilon^2}{2\tau}(\mathbf{v}, \mathbf{v})_h - \tfrac{\gamma}{2}(\mathbf{A}\mathbf{v}, \mathbf{v})_h + \tfrac{\gamma\varepsilon^2}{2}(\nabla\mathbf{v}, \nabla\mathbf{v}) \right] . \tag{76}$$

Here $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ play the role of Lagrange multipliers for the first two equations in the definition of $\boldsymbol{K}$. To show that $\boldsymbol{v}$ equals $\mathbf{0}$ we prove that $\mathbf{v} \equiv \mathbf{0}$ is the unique solution of (76).
The minimization problem (76) is strictly convex if $\tau \leq \frac{\varepsilon^2}{\gamma\sigma_{max}(\mathbf{A})}$. For larger τ we need to control $(\mathbf{v},\mathbf{v})_h$ on $\boldsymbol{K}$. Using the Poincaré inequality (70) we obtain

$$\tfrac{\gamma\varepsilon^2}{2}(\nabla\mathbf{v},\nabla\mathbf{v}) + \tfrac{1}{2}(\tfrac{\varepsilon^2}{\tau} - \gamma\sigma_{max}(\mathbf{A}))(\mathbf{v},\mathbf{v})_h \geq (\tfrac{\gamma\varepsilon^2}{2} + \tfrac{1}{2}c_h^p(\tfrac{\varepsilon^2}{\tau} - \gamma\sigma_{max}(\mathbf{A}))(\nabla\mathbf{v},\nabla\mathbf{v})\,.$$

Hence if (71) holds then the minimization problem (76) is strictly convex and (76) is uniquely solvable. Thus $\boldsymbol{v} \equiv \mathbf{0}$. Therefore (63) reduces to $(\kappa_i)_j + \alpha_i + \beta_j = 0$.
Now we show that $\boldsymbol{\alpha} = \mathbf{0}$. For $j \in \mathcal{I}_i \cap \mathcal{I}_l$ we have $(\kappa_l)_j = (\kappa_i)_j = 0$ and hence

$$\beta_j = -\alpha_i \quad \text{and} \quad \beta_j = -\alpha_l\,.$$

Consequently $\alpha_i = \alpha_l$ for all i,l with $\mathcal{I}_i \cap \mathcal{I}_l \neq \emptyset$. Since the inactive sets are connected we obtain that $\alpha_i = \alpha_l$ for all $i,l \in \{1,...,N\}$ and since $\sum_{i=1}^{N} \alpha_i = 0$ we obtain $\boldsymbol{\alpha} \equiv \mathbf{0}$.
Furthermore this yields $\beta_j = 0$ if there exists an i with $j \in \mathcal{I}_i$. Having $\mathcal{J} = \bigcup_{i=1}^{N} \mathcal{I}_i$ it follows $\boldsymbol{\beta} = \mathbf{0}$ which also gives that $\boldsymbol{\kappa} = \mathbf{0}$. □
We are now in a position to prove a local convergence result for the (PDAS-I)-algorithm.

Theorem 4.2 *Assume $(\mathbf{u}_h, \boldsymbol{\mu}_h, \boldsymbol{\lambda}_h, \Lambda_h)$ is a solution of the discretized Allen-Cahn problem (57) - (61). Assume that the inactive sets $\hat{\mathcal{I}}_i = \{j \in \mathcal{J} : (\mathbf{u}_i)_j > 0\}$ are connected and assume that $\tau < \frac{\varepsilon^2}{\gamma\sigma_{max}(\mathbf{A})}$. Then (PDAS-I) converges locally in a neighborhood of $(\mathbf{u}_h, \boldsymbol{\mu}_h, \boldsymbol{\lambda}_h, \Lambda_h)$.*

Proof: We choose a neighborhood of $(\mathbf{u}_h, \boldsymbol{\mu}_h, \boldsymbol{\lambda}_h, \Lambda_h)$ such that the corresponding inactive sets are all connected and such that $\mathcal{J} = \bigcup_{i=1}^{N} \mathcal{I}_i$. Now Theorem 4.1 guarantees invertibility of $\boldsymbol{G}$ in this neighborhood. Since only finitely many constellations with active sets are possible we can deduce that $\boldsymbol{G}^{-1}$ is uniformly bounded in this neighborhood. Hence convergence results in [16, 27] now give the local convergence result.

□

Remark 4.3 *i) The graph describing all possible interfaces with respect to the space-time continuous solution* $\mathbf{u}$ *is connected, see Theorem 2.1. Arguing similar as in the proof of Theorem 2.1 one observes that the graph corresponding to the discrete solution is also connected provided h is small enough. In practice we only need to ensure that h is so small that there are enough mesh points on the interface.*

ii) The condition $\mathcal{J} = \bigcup_{i=1}^{N} \mathcal{I}_i^h$ *holds in a neighborhood of the solution* $\mathbf{u}_h$ *of (56). This follows from* $\sum_{i=1}^{N} (u_i)_j = 1$ *and* $(u_i)_j \geq 0$ *which guarantees that for all* $j \in \mathcal{J}$ *there exists an* $i \in \{1, \dots, N\}$ *such that* $(u_i)_j > 0$*. Hence one can find a neighborhood of* $\mathbf{u}_h$ *where this holds true too.*

iii) Of course the condition on the time step in Theorem 4.2 can be relaxed taking Theorem 4.1 into account. As the Poincaré constants of all possible active sets in the neighborhood would enter into a precise assumption, we did not state such a result in a precise way.

5 Computational results

Most of the existing literature on numerical methods for systems of Allen-Cahn variational inequalities concentrates on the explicit discretization in time where with the use of mass lumping a non-linear system of equations has to be solved [23, 24]. Using explicit time discretization leads to the usual stability restriction for parabolic PDEs, $\tau \leq Ch^2$. For Allen-Cahn variational inequalities we need that $h \ll \varepsilon$ which makes this time step restriction very severe. In [33] a multigrid algorithm based on a subspace correction approach is used where a subspace is decomposed into smaller spaces leading to a polygonal Gauss-Seidel relaxation as the fine grid smoother. A semi-implicit time discretization is employed which is unconditionally stable. However, for the scalar Allen-Cahn equation it has been observed that the semi-implicit discretization can lead to inaccurate approximations especially for large time steps [9]. We use an implicit discretization in time which has a time step restriction (71) that is less severe than the one for the explicit discretization and leads to a higher accuracy.

In this section we discuss some computational results including mesh independency of the PDAS iteration numbers if time and spatial discretization parameters are decreased simultaneously. In addition we study the influence of an increasing number of phases, convergence of the discretization and present some numerical simulations in 3D. Further applications of the approach in topology optimization, imaging and materials science will be the subject of a forthcoming paper. In Subsection 5.1 we apply the PDAS-method to the vector-valued Allen-Cahn equation without volume constraints. First, we compute numerical solutions which approximate an analytical solution of a corresponding sharp interface problem and then we discuss some properties of our method. Second, we present some numerical simulations for three and more order parameters. Finally, numerical simulations of grain growth with many order parameters, i.e. N large are presented.

In Subsection 5.2 the PDAS-method is applied to the vector-valued Allen-Cahn model with volume constraints. An explicit solution for a corresponding sharp in-

terface model is derived and the convergence and accuracy of our method is analyzed.

We note that since the interfacial thickness is proportional to ε in order to resolve the interfacial layer we need to choose $h \ll \varepsilon$ (see [18, 20] for details). Away from the interface h can be chosen larger and hence adaptivity in space can heavily speed up computations. In fact we use the finite element toolbox Alberta 2.0 (see Schmidt and Siebert [37]) for adaptivity and we implemented the same mesh refinement strategy as in Barrett, Nürnberg and Styles [5], i.e. a fine mesh is constructed where $0 < (u_h^{n-1})_i < 1$ with a coarser mesh present in the bulk regions $(u_h^{n-1})_i = 0$ and $(u_h^{n-1})_i = 1$ for $i \in \{1, ..., N\}$.
In all our computations we set the matrix $\mathbf{A}$ in the multi-obstacle potential to be the identity matrix and we take $\gamma = 1$. We set the time step $\tau = 1 \cdot 10^{-4}$ unless otherwise stated. For the computations in two space dimensions we take $\Omega = (-1, 1)^2$, $\varepsilon = \frac{1}{16\pi}$, the minimal diameter of an element $h_{min} = 3.91 \cdot 10^{-3}$ and the maximal diameter $h_{max} = 6.25 \cdot 10^{-2}$. For computations in three space dimensions we take $\Omega = (-1, 1)^3$, $\varepsilon = \frac{1}{12\pi}$, $h_{min} = 7.81 \cdot 10^{-3}$, and $h_{max} = 1.25 \cdot 10^{-1}$. This is necessary due to memory restrictions. We solve (63)-(65) in two space dimensions using the direct solver UMFPACK [17] and in three space dimensions we use MINRES. For a more efficient solver with preconditioning, we refer to [10].

5.1 Vector-valued Allen-Cahn variational inequality without volume constraints

The Allen-Cahn model approximates motion by curvature, see [12]. We hence consider circles for which the radius $R(t)$ at time t is given by the ODE $\frac{d}{dt}R(t) = -\frac{1}{R(t)}, R(0) = 0.4$, see [12].
We set $N = 3$, i.e. three phases are present, and take the simple problem of two shrinking circles with initial radii $R(0) = 0.4$ and centres $(-0.5, 0)$ and $(0.5, 0)$. Two order parameters $(u_h)_1$ and $(u_h)_2$ are each set to be 1 on one circle and 0 anywhere else. The third order parameter $(u_h)_3$ is set to be 1 outside the circles and 0 inside. We take smooth transition layers of width $\varepsilon\pi$. Both circles shrink with the same velocity $-\frac{1}{R(t)}$.
As mentioned previously we cannot show analytically that the number of PDAS iterations is mesh independent. However, in our application a good initial data on the current time step is given from the solution of the previous time step. Therefore, the mesh independence is only of interest if both the mesh size h and also the time step τ are driven to 0. We use a uniform mesh of size h and the same initial data as previously. Table 1 shows that when h and τ both are decreased simultaneously according to $\tau \approx h^2$, the number of PDAS iterations remains stable. Table 1 also gives the processing times up to $T = 0.03$ and the average number of unknowns in the linear system (63), (64).

[1]Due to memory restrictions an adaptive mesh was used with $h_{min} = h$

τ	h	DOFs	PDAS-iter.	Unknowns	CPU [s]
$1 \cdot 10^{-3}$	1/128	66049	4.77	12057	28
$2.5 \cdot 10^{-4}$	1/256	263169	4.86	51720	522
$6.25 \cdot 10^{-5}$	1/512	1050625	4.70	214263	10263
$1.5625 \cdot 10^{-5}$	1/1024	319230 [1]	4.57	872760	209998

Table 1: Average number of PDAS iterations up to $T = 0.03$ for varying mesh and time step sizes.

In the next computation we examine the number of PDAS-iterations for increasing phases N. We take circles of radius 0.3 and position them such that they do not intersect. For three order parameters we take two circles (one phase for each circle and one phase outside the circles); for $N = 4$ we take three circles, and so on up to $N = 7$ where six circles are needed. Moreover we consider $N = 2$ (one circle) for the vector-valued Allen-Cahn equation with two order parameters and the scalar Allen-Cahn equation. In this case the scalar equation is obtained by taking $u_2 = 1 - u_1$. Figure 2 shows the average number of PDAS-iterations for t between 0 and 0.04 with fixed timestep size $\tau = 1 \cdot 10^{-4}$. For $N = 1$ and $N = 2$ the number of PDAS-iterations is considerably lower than for larger N. This could be because both order parameters are inactive on the interface. For larger N we have that two order parameters are inactive on each of the interfaces whilst the other order parameters may be active. For $N \geq 3$ the average number of PDAS-iterations remains almost stable. We conclude that the number of PDAS-iterations is driven by the change of the active and inactive sets only, while the number of phases does not seem to make much difference.

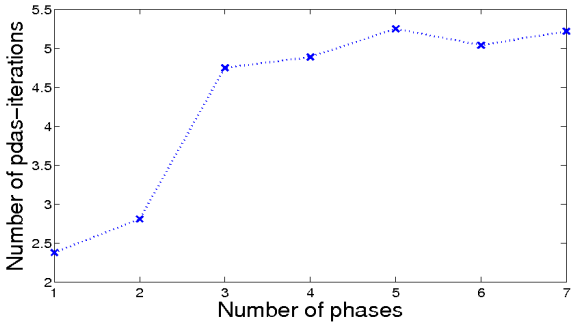


Figure 2: Average number of PDAS-iterations for increasing number of phases N

To demonstrate the efficiency of the method we also performed a computation with thirty order parameters. In this case the Allen-Cahn system models grain growth and at triple junctions a $\frac{2\pi}{3}$ angle condition has to hold, see [13, 24] for details. For the computation in Figure 3 we use a Voronoi partitioning algorithm to randomly fill the 2D computational domain. At the beginning of the computation cell edges are not smooth and at triple junctions angle conditions are not fulfilled, but already

at time $t = 0.004$ the partitioning becomes regular and fulfills the angle conditions approximately. Each of the thirty phases describes a different orientational variant in a crystalline material.

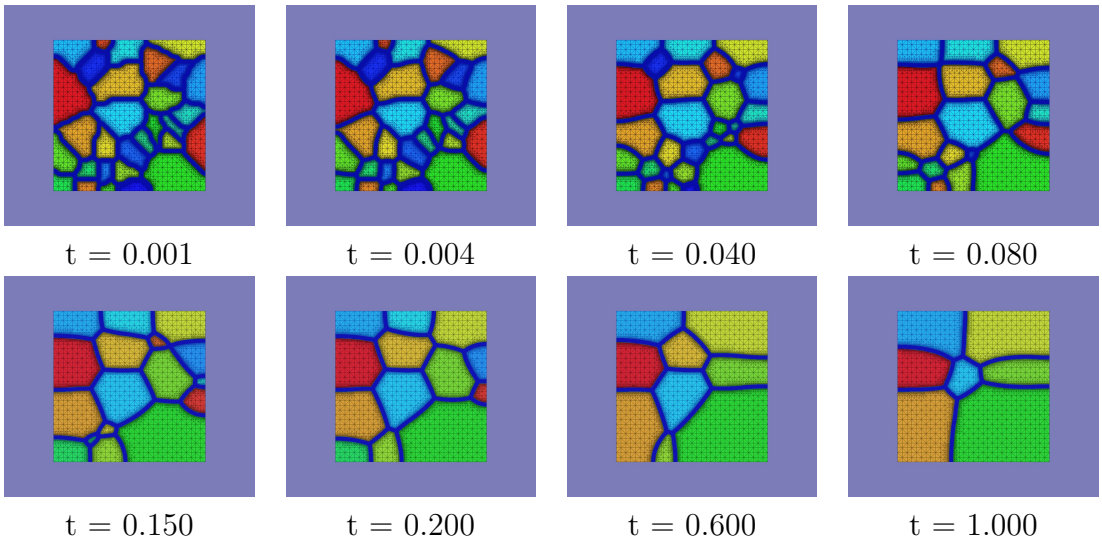


Figure 3: Vector-valued Allen-Cahn equation with Voronoi partitioning as initial data (thirty order parameters).

5.2 Vector-valued Allen-Cahn variational inequality with volume constraints

Now we consider the vector-valued Allen-Cahn variational inequality with volume constraints. In the following we compare the approximation obtained by the introduced method with an exact sharp interface solution. For the scalar Allen-Cahn variational inequality (and also for the vector-valued Allen-Cahn variational inequality with $N = 2$) we obtain an explicit solution for the following problem: Given two circles with radii r_1 and r_2 which do not intersect, then the sharp interface problem for volume conserved motion by curvature results in the following system of ODEs: $r_1' = -\frac{1}{r_1} + \lambda$, $r_2' = -\frac{1}{r_2} + \lambda$ together with the condition of volume conservation $0 = \frac{1}{2}(r_1^2 + r_2^2)'$, where the initial radii $r_1(0)$ and $r_2(0)$ are known. This can be solved analytically, see [36]. We can use this problem in the case of $N = 3$ by considering two decoupled systems, that is four circles that do not intersect where phase 1 occupies two circles, phase 2 occupies the other two circles and phase 3 is present outside these four circles, see Figure 4. For the first order parameter (blue) we take the initial radii of the two circles to be $r_1(0) = 0.2$ and $r_2(0) = 0.3$ and for the second order parameter (red) we take $r_3(0) = 0.4$ and $r_4(0) = 0.25$. Figure 5 shows the approximate solution and the exact solution for the two big circles for two different time steps $\tau = 1 \cdot 10^{-3}$ and $\tau = 1 \cdot 10^{-4}$. The behaviour of the two small circles is

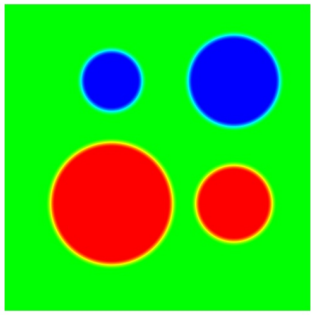


Figure 4: Initial configuration for vector-valued Allen-Cahn ($N = 3$) with volume constraints for comparison with explicitely known solutions.

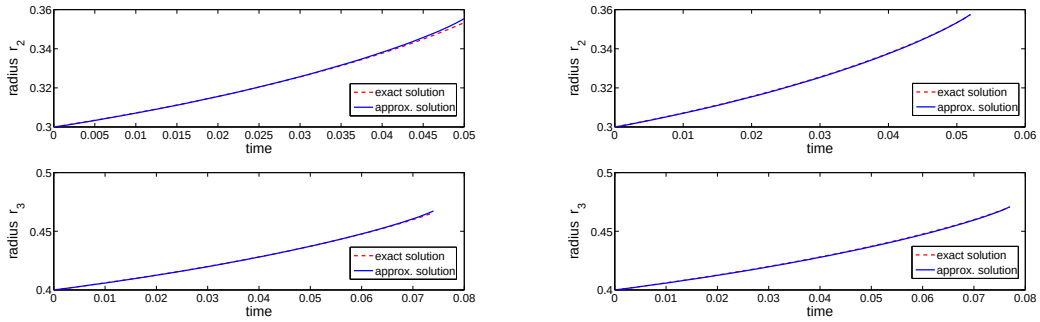


Figure 5: Exact solution of the sharp interface problem and approximate radii of the Allen-Cahn solution for time steps $\tau = 2 \cdot 10^{-3}$ (left) and $\tau = 1 \cdot 10^{-4}$ (right).

essentially the same and therefore omitted. For both time steps the approximations are very good. Larger time steps can also be taken as long as the curvature does not become too big. An adaptive time step strategy might speed up computations for the volume-conserved vector-valued Allen-Cahn variational inequality.

In order to demonstrate that the PDAS approach can be used for 3D computations as well we computed solutions for the volume constrained case in three space dimensions. The volume constrained Allen-Cahn model can be used to compute soap bubbles as long time limits which are steady states of the Allen-Cahn model, see [23] for details.

In nature, soap bubble configurations enclose and separate several regions of space. They have fixed volumes and tend to minimize the total surface area. This observation leads to the following basic problem: How can one enclose and separate n regions of $\mathbb{R}^3$ having volumes $v_1, v_2, ..., v_n$ with the smallest possible surface area. For $n = 1$, i.e. a single region it is well known that a sphere is the optimal configuration. It has been proved that for two regions the optimal configuration is a double bubble [28].

Due to the integral constraints the regions to be separated have fixed volume whilst

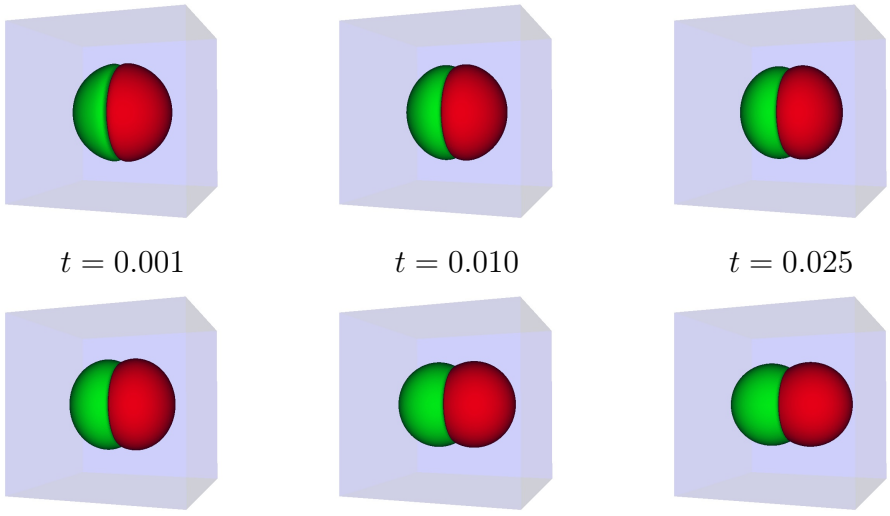
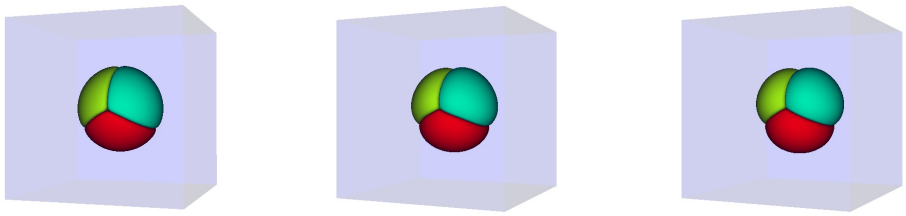
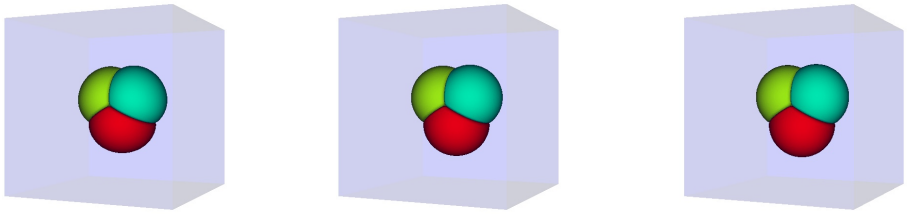


Figure 6: Double bubble; vector-valued Allen-Cahn with volume constraints, three order parameters.

the evolution tends to minimize the surface energy, and hence the surface area, see [13, 23]. In the first computation we use three order parameters and start with a sphere where the left half is occupied by phase 1 and the right half is occupied by phase 2. We note that first very rapidly the $2\pi/3$ angle condition is attained. Then the two halves gradually move outwards whilst staying attached in the middle, see Figure 6. The movement ends when the steady state, a double bubble, is reached. Figure 7 shows a similar computation for $N = 4$. We begin the computations with a sphere that is divided into three equal spherical wedges. Each of these wedges is represented by a different phase, i.e. we have three phases in the sphere and one phase outside. As before, first the angle condition is attained and then the three parts move until a triple bubble is reached.



$t = 0.001$ $t = 0.010$ $t = 0.020$



$t = 0.040$ $t = 0.100$ $t = 0.500$

Figure 7: Triple bubble; vector-valued Allen-Cahn with volume constraints, four order parameters.

References

[1] Allen, S.M. and Cahn, J.W., *A microscopic theory for antiphase motion and its application to antiphase domain coarsening*, Acta Metall. Mater. 27 (1979), 1085 – 1095.

[2] Alt, H.W., *Lineare Funktionalanalysis*, Springer Verlag 2006.

[3] Barrett, J.W. and Blowey, J.F., *An error bound for the finite element approximation of a model for phase separation of a multi-component alloy*, IMA Journal of Numerical Analysis 16 (1996), 257–287.

[4] Barrett, J.W. and Blowey, J.F., *Finite element approximation of a model for phase separation of a multi-component alloy with non-smooth free energy*, Numerische Mathematik 77 (1997), 11–34.

[5] Barrett, J.W, Nürnberg, R. and Styles, V., *Finite element approximation of a phase field model for void electromigration*, SIAM J. Numer. Anal. 46 (2004), 738–772.

[6] Benes, M., Chalupecky, V. and Mikula, K., *Geometrical image segmentation by the Allen-Cahn equation*, Appl. Numer. Math. 51 (2004), no. 2-3, 187–205.

[7] Bergounioux, M., Ito, K. and Kunisch, K., *Primal-dual active set strategy for constrained optimal control problems*, SIAM J. Control Optim. 37, no. 4 (1999), 1176–1194.

[8] Blank, L., Butz, M. and Garcke. H., *Solving the Cahn-Hilliard variational inequality with a semi-smooth Newton method*, to appear in ESAIM: Control, Optimization and Calculus of Variations (2010).

[9] Blank, L., Garcke, H., Sarbu, L. and Styles, V., *Primal-dual active set methods for Allen-Cahn variational inequalities with non-local constraints*, DFG priority program "Optimization with PDEs", Preprint SPP1253-09-01.

[10] Blank, L., Sarbu, L. and Stoll, M., *Preconditioning for Allen-Cahn variational inequalities with non-local constraints*, Preprint 11/2010, University of Regensburg.

[11] Blesgen, T. and Weikard, U., *Multi-component Allen-Cahn equation for elastically stressed solids*, Electron. J. Differential Equations 2005, no. 89, 1–17.

[12] Blowey, J. F. and Elliott, C. M., *Curvature dependent phase boundary motion and parabolic double obstacle problems*, Degenerate diffusions (Minneapolis, MN, 1991), 19–60, IMA Vol. Math. Appl., 47, Springer, New York, 1993

[13] Bronsard, L. and Reitich, F., *On three-phase boundary motion and the singular limit of a vector-valued Ginzburg-Landau equation*, Arch. Rat. Mech. Anal. 124 (1993), 355–379.

[14] Capuzzo Dolcetta, I., Finzi Vita, S. and March, R., *Area-preserving curve-shortening flows: from phase separation to image processing*, Interfaces Free Bound. 4, no 4 (2002), 325–343.

[15] Chen, X, Elliott, C.M., Gardiner, A., and Zhao, J.J., *Convergence of numerical solutions to the Allen-Cahn equation*, Applic. Anal. 69 (1998), 47–56.

[16] Chen, X., Nashed, Z. and Qi, L., *Smoothing methods and semismooth methods for non-differentiable operator equations*, SIAM J. Numer. Anal. 38 (2000), 1200–1216.

[17] Davis, T.A., *Umfpack version 4.4 user guide*, tech. report, Dept. of Computer and Information Science and Engineering, Univ. of Florida (2005).

[18] Elliott, C. M. *Approximation of curvature dependent interface motion*, *State of the art in Numerical Analysis, IMA Conference Proceedings*, 63, pp. 407–440. Clarendon Press, Oxford (1997).

[19] Elliott, C.M. and Luckhaus. S., *A generalised diffusion equation for phase separation of a multi-component mixture with interfacial free energy*, SFB256 University Bonn, Preprint 195 (1999).

[20] Elliott, C. M. and Styles, V., *Computations of bi-directional grain boundary dynamics in thin films*, J. Comput. Phys. 187 (2003) 524–543.

[21] Evans, L.C., *Partial differential equations*, Graduate Studies in Mathematics 19. American Mathematical Society, Providence, RI (2002).

[22] Feng, X and Prohl, A., *Numerical analysis of the Allen-Cahn equation and approximation for mean curvature flows*, Numer. Math. 94 (2003), 33–65.

[23] Garcke, H., Nestler, B., Stinner, B. and Wendler, F., *Allen-Cahn systems with volume constraints*, M^3AS: Math. Models Methods in Appl. Sci. 18, no. 8 (2008), 1347–1381.

[24] Garcke, H., Nestler, B., and Stoth, B., *A multi phase field concept: numerical simulations of moving phase boundaries and multiple junctions*, SIAM J. Appl. Math. 60, no. 1 (1999), 295–315.

[25] Garcke, H., Nürnberg, R. and Styles, V., *Stress- and diffusion-induced interface motion: modelling and numerical simulations*, European J. Appl. Math. 18, no. 6 (2007), 631–657.

[26] Grisvard, P., *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics 24. Pitmann, Boston, MA, 1985.

[27] Hintermüller, M., Ito, K. and Kunisch, K., *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim. 13, no. 3 (2002), 865–888.

[28] Hutchings, M., Morgan, F., Ritore, M. and Ros, A. *Proof of the double bubble conjecture*, Annals of Mathematics 155, no. 2 (2002), 459–489.

[29] Ito, K. and Kunisch, K., *Semi-smooth Newton methods for variational inequalities of the first kind.*, M2AN Math. Model. Numer. Anal. 37, no. 1 (2003), 41–62.

[30] Kay, D.A. and Tomasi, A., *Color image segmentation by the vector-valued Allen-Cahn phase-field model: a multigrid solution*, IEEE Trans. Image Processing 18, no. 10 (2009), 2330–2339.

[31] Kessler, D., Nochetto, R., and Schmidt, A, *A posteriori error control for the Allen-Cahn problem: circumventing Gronwals inequality*, M2AN 38 (2004), 129–142.

[32] Kornhuber,R. and Krause, R., *On multigrid methods for vector-valued Allen- Cahn equations*, In I. Herrera et al., editor, Domain Decomposition Methods in Science and Engineering, pages 307–314, Mexico City, Mexico, 2003. UNAM.

[33] Kornhuber, R. and Krause, R., *Robust multigrid methods for vector-valued Allen-Cahn equations with logarithmic free energy*, Comput. Vis. Sci. 9 (2006), 103–116.

[34] Nochetto, R. and Verdi, C., *Convergence past singularities for a fully discrete approximation of curvature-driven interfaces*, SIAM J. Numer. Anal. 34 (1997), no. 2, 490–512.

[35] Rodrigues, J. F. and Santos, L., *On a constrained reaction-diffusion system related to multiphase problems* Discrete and continuous dynamical systems 25, no.1 (2009), 299–319.

[36] Rubinstein, J. and Sternberg, P., *Nonlocal reaction diffusion equations and nucleation*, IMA J. Appl. Math. 48 (1992), 249–264.

[37] Schmidt, A. and Siebert, K.G., *Design of adaptive finite element software. The finite element toolbox ALBERTA*, Lecture Notes in Computational Science and Engineering 42. Springer-Verlag, Berlin (2005) xii+315.

[38] Simon, J., *Compact sets in the space $L^p(0,T;B)$*, Ann. Mat. Pura Appl. (4) 146 (1987), 65–96.

[39] Wendler, F., Becker, J.K., Nestler, B., Bons, P. and Walte, N.P., *Phase-field simulations of partial melts in geological materials*, Comput. Geosci. 35, no. 9 (2009), 1907–1916.