Universität Regensburg
Mathematik

Primal-dual active set methods for Allen-Cahn variational inequalities with non-local constraints

Luise Blank, Harald Garcke, Lavinia Sarbu
and Vanessa Styles

Preprint Nr. 06/2009

Primal-dual active set methods for Allen-Cahn variational inequalities with non-local constraints

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

March 20, 2009

Abstract

We propose and analyze a primal-dual active set method for local and non-local Allen-Cahn variational inequalities. An existence result for the non-local variational inequality is shown in a formulation involving Lagrange multipliers for local and non-local constraints. Superlinear local convergence is shown by interpreting the approach as a semi-smooth Newton method. Properties of the method are discussed and several numerical simulations demonstrate its efficiency.

Key words: Allen-Cahn variational inequality, non-local constraints, primal-dual active set methods, semi-smooth Newton methods, finite element approximation.

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

1 Introduction

The Allen-Cahn model describes interface motion with applications including materials science, image processing, biology and geology, see [5, 29, 19, 30, 4, 3, 18, 22, 9, 27]. Here an interface in which a phase field or order parameter rapidly changes its value, is modelled to have a thickness of order ε where $\varepsilon > 0$ is a small parameter. The model is based on a non-convex energy E which has the form

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

where $\Omega \subset \mathbb{R}^d$ is an open and bounded domain, $\gamma > 0$ is a parameter related to the interfacial energy and $u : \Omega \to \mathbb{R}$ is the phase field, also called order parameter. The potential function $\psi : \mathbb{R} \to \mathbb{R}_0^+ \cup \{\infty\}$ is assumed to have two global minima at the

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

points ± 1 and the values ± 1 describe the pure phases. Examples are $\psi(u) = (1-u^2)^2$ or the obstacle potential

$$\psi(u) = \begin{cases} \frac{1}{2}(1-u^2) & \text{if} \quad |u| \leq 1\,, \\ \infty & \text{if} \quad |u| > 1\,. \end{cases}$$

Introducing $\psi_0(u) := \frac{1}{2}(1-u^2)$ and the indicator function

$$I_{[-1,1]}(u) := \begin{cases} 0 & \text{if} \quad |u| \leq 1\,, \\ \infty & \text{if} \quad |u| > 1 \end{cases}$$

we obtain

$$\psi(u) = \psi_0(u) + I_{[-1,1]}(u)\,. \tag{1}$$

In order to have $E(u)$ of moderate size u favours the values ± 1 due to the potential function. On the other hand given the gradient term $\int\limits_\Omega |\nabla u|^2$ oscillations between the values ± 1 are energetically not favourable.
Given an initial phase distribution $u(.,0) = u_0 : \Omega \to \mathbb{R}$ at time $t = 0$ the interface motion can be modelled by the steepest descent of E with respect to the L^2-norm which results into the Allen-Cahn equation. In the case of a smooth potential ψ we obtain after a suitable rescaling of time

$$\varepsilon \partial_t u = \gamma \varepsilon \Delta u - \tfrac{\gamma}{\varepsilon}\psi'(u) \quad \text{for } x \in \Omega \text{ and } t > 0$$

together with Neumann boundary conditions $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$, where ν is the outer unit normal to Ω. If ψ has the form (1) we obtain, see [5],

$$(\varepsilon \partial_t u, \chi - u) + \gamma\varepsilon(\nabla u, \nabla(\chi - u)) + \tfrac{\gamma}{\varepsilon}(\psi_0'(u), \chi - u) \geq 0 \tag{2}$$

which has to hold for almost all t and all $\chi \in H^1(\Omega)$ with $|\chi| \leq 1$. Here and in what follows $(.,.)$ denotes the L^2–inner product.
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 $\fint\limits_\Omega u dx = m$ where $m \in (-1,1)$ is a fixed number and we use the notation $\fint\limits_\Omega f(x)dx := \frac{1}{|\Omega|}\int\limits_\Omega f(x)dx$ with $|\Omega|$ being the Lebesgue measure of Ω. In the case of a smooth potential ψ we obtain (see [29, 19])

$$\varepsilon \partial_t u = \gamma\varepsilon\Delta u - \tfrac{\gamma}{\varepsilon}\psi'(u) + \tfrac{\gamma}{\varepsilon}\fint\limits_\Omega \psi'(u)dx\,.$$

Assuming homogeneous Neumann boundary conditions it follows that

$$\tfrac{d}{dt}\int\limits_\Omega u dx = 0 \quad \text{and} \quad \tfrac{d}{dt}E(u) \leq 0\,.$$

In the case of an obstacle potential we need to solve for given initial data $u_0 \in H^1(\Omega)$ with $|u_0| \leq 1$ a.e. in Ω the following problem [5]:

($\mathbf{P}_m$) *Find* $u \in H^1(\Omega_T)$ *such that* $\fint_\Omega u(x,t)dx = m$, $u(.,0) = u_0$, $|u| \leq 1$ *a.e. in* $\Omega_T := \Omega \times (0,T)$ *and*

$$(\varepsilon \partial_t u, \chi - u) + \gamma\varepsilon(\nabla u, \nabla(\chi - u)) + \tfrac{\gamma}{\varepsilon}(\psi_0'(u), \chi - u) \geq 0 \tag{3}$$

which has to hold for almost all t *and all* $\chi \in H^1(\Omega)$ *with* $|\chi| \leq 1$ *and* $\fint_\Omega \chi = m$.

In this paper we reformulate ($\mathbf{P}_m$) with the help of Lagrange multipliers μ_- and μ_+ for the inequality constraints $u \geq -1$, $u \leq 1$ and a Lagrange multiplier λ for the equality constraint $\fint_\Omega u = m$. Under appropriate assumptions we show the existence of a unique solution u of ($\mathbf{P}_m$) together with unique multipliers, see Section 2.
The main goal of this paper is to introduce and analyze a primal-dual active set method for a finite element discretization of a (semi-)implicit Euler discretization of ($\mathbf{P}_m$) and respectively, (2). Our formulation make use of the three dual variables μ_-, μ_+ and λ, see Section 3. As was demonstrated in [21] the primal-dual active set method can be reformulated as a semi-smooth Newton method. This allows us to show local superlinear convergence of the proposed primal-dual active set method, see Section 3.
When solving ($\mathbf{P}_m$) and (2) in each time step we have to solve a (non-local) obstacle problem. In order to demonstrate specific features of our approach we also discuss elliptic obstacle problems. Given $F \in (H^1(\Omega))^*$ and $u_D \in H^1(\Omega)$ we want to solve

$$\mathcal{F}(u) := \int_\Omega \tfrac{1}{2}|\nabla u|^2 dx - F(u - u_D) \longrightarrow \min! \text{ for } u \in \hat{\mathcal{K}} \tag{4}$$

where $\hat{\mathcal{K}}$ is either $\mathcal{K} := \{u \in H^1(\Omega) \mid u - u_D \in H_0^1(\Omega),\ \varphi \leq u \leq \psi\}$ or $\mathcal{K}_m := \{u \in \mathcal{K} \mid \fint_\Omega u dx = m\}$ with $m \in \mathbb{R}$ where $\varphi, \psi : \Omega \to \mathbb{R} \cup \{-\infty, \infty\}$. Similar problems appear in the modelling of water tanks [7] and in the calculus of variations [20, 13].
We show that there are cases where the active set only moves one mesh point per primal-dual active set iteration leading to a very slow convergence. For bilateral constraints global convergence cannot be expected. In fact there are situations in which iterates can oscillate between the two constraints. Numerical experiments demonstrate this behaviour. Also we suggest and discuss how a nested approach can speed up the method.
Finally in Section 5 we present numerical simulations for the non-local as well as for the local Allen-Cahn variational inequality, i.e. with and without mass constraints. In the local situation we observe that our approach is faster than other proposed methods. Furthermore, larger time steps can be taken. For the non-local Allen-Cahn variational inequality so far only explicit methods have been used, see [5, 19]. In this paper we present a first method to numerically solve an implicit time discretization

of the Allen-Cahn variational inequality with integral constraint. We show efficiency and accuracy of our method for a problem where the explicit solution is known. In particular, it turns out that large time steprs are possible, that implicit discretization in time lead to higher accuracy than the semi-implicit one without loss of efficiency and that the computation time for the non-local Allen-Cahn variational inequality remains almost the same as in the local case. Finally we present two numerical simulations for the Allen-Cahn variational inequality with integral constraint.

2 Existence theory

In this section we show existence and uniqueness to the Allen-Cahn variational inequality with integral constraint, see (3). As a first step we reformulate the problem in the following lemma with the help of Langrange multipliers μ_+ and μ_- corresponding to the inequality constraints $u \leq 1$ and $u \geq -1$ and a Lagrange multiplier λ corresponding to the constraint $\fint_\Omega u dx = m := \fint_\Omega u_0 dx$.
As a general assumption we require:

($\mathcal{A}_m$) *The domain $\Omega \subset \mathbb{R}^d$ is bounded and either convex or has a $C^{1,1}$-boundary. Furthermore the initial data $u_0 \in H^1(\Omega)$ fulfill $|u_0| \leq 1$ a.e. and $\fint_\Omega u_0 = m$ for a given $m \in (-1,1)$.*

Lemma 2.1 *Let $T > 0$ be a positive time and let the assumptions ($\mathcal{A}_m$) hold. A function $u \in L^2(0,T;H^2(\Omega)) \cap H^1(\Omega_T)$ solves ($\mathbf{P}_m$) if there exists $\mu_+, \mu_- \in L^2(\Omega_T)$ and $\lambda \in L^2(0,T)$ such that*

$$\lambda = \varepsilon^2 \partial_t u - \gamma\varepsilon^2 \Delta u + \gamma \psi_0'(u) + \mu_+ - \mu_- \quad \text{a.e. in } \Omega_T\,, \tag{5}$$

$$u(0) = u_0,\ \frac{\partial u}{\partial \nu} = 0 \quad \text{a.e. on } (\partial\Omega)_T := \partial\Omega \times (0,T)\,, \tag{6}$$

$$\fint_\Omega u dx = m \quad \text{for almost all } t \in [0,T]\,, \tag{7}$$

$$|u| \leq 1 \quad \text{a.e. in } \Omega_T\,, \tag{8}$$

$$\mu_+(u-1) = 0, \mu_-(u+1) = 0 \quad \text{a.e. in } \Omega_T\,, \tag{9}$$

$$\mu_+ \geq 0, \mu_- \geq 0 \quad \text{a.e. in } \Omega_T\,. \tag{10}$$

Proof: Let $\eta \in L^2(0,T;H^1(\Omega))$ be such that $|\eta| \leq 1$ a.e. in Ω_T and $\fint_\Omega \eta(x,t)dx = m$ for almost all $t \in [0,T]$. Multiplying (5) by $(\eta - u)$ gives

$$0 = \int_{\Omega_T} (\varepsilon^2 \partial_t u - \varepsilon^2 \gamma \Delta u + \gamma \psi_0'(u))(\eta - u) + \int_{\Omega_T} \mu_+(\eta - u) - \int_{\Omega_T} \mu_-(\eta - u)\,.$$

Using the properties of η and (8)-(10) gives

$$\mu_+(\eta - u) \leq 0,\ \mu_-(\eta - u) \geq 0 \quad \text{a.e. in } \Omega_T\,.$$

After integration by parts in space we obtain

$$0 \leq \int\limits_{\Omega_T} (\varepsilon \partial_t u + \tfrac{\gamma}{\varepsilon} \psi_0'(u))(\eta - u) + \gamma\varepsilon \int\limits_{\Omega_T} \nabla u \cdot \nabla(\eta - u)\,.$$

Now (3) follows after localization in time. □

The primal-dual active set method that we propose will heavily depend on the Lagrange multipliers μ_-, μ_+, λ. We hence show the existence of a solution u together with unique Lagrange multipliers μ_-, μ_+ and λ. In the proof of the existence theorem we handle the linear equality constraint $\fint\limits_\Omega u = m$ by projection and use a penalty approach for the inequality constraint $|u| \leq 1$, see also [1, 6]. In particular we replace the indicator function in ψ by terms penalizing deviations of u from the interval $[-1, 1]$.

We define $\psi_+(z) := \max(z-1, 0)^2$, $\psi_-(z) := \min(z+1, 0)^2$ for all $z \in \mathbb{R}$ and $\psi_\delta(z) := \psi_0(z) + \frac{1}{\delta}(\psi_+(z) + \psi_-(z)) + 1$ for $\delta > 0$. It can be shown that $\psi_\delta'' \geq -1$ and that $\psi_\delta(z) \geq 0$ for all $z \in \mathbb{R}$ holds for $\delta \in (0, 2)$. Now, we consider the L^2–gradient flow of

$$E_\delta(u) := \int\limits_\Omega (\tfrac{\gamma\varepsilon}{2}|\nabla u|^2 + \tfrac{\gamma}{\varepsilon}\psi_\delta(u))dx$$

taking the mean value constraint into account. Hence we want to solve the semilinear parabolic equation

$$\varepsilon^2 \partial_t u_\delta = \gamma\varepsilon^2 \Delta u_\delta - \gamma \psi_\delta'(u_\delta) + \fint\limits_\Omega \gamma \psi_\delta'(u_\delta) dx \text{ in } \Omega_T\,, \tag{11}$$

$$u_\delta(0) = u_0, \qquad \frac{\partial u_\delta}{\partial \nu} = 0 \quad \text{on } (\partial\Omega)_T\,. \tag{12}$$

The main challenge is to control approximate versions of the Lagrange multipliers which is non-standard due to the coupling of non-local equality and local inequality constraints.

Theorem 2.1 *Let the assumptions $(\mathcal{A}_m)$ hold and let $T > 0$. Then there exists a unique solution $(u, \mu_+, \mu_-, \lambda)$ of (5)-(10) with the following properties*

$$u \in L^2(0,T; H^2(\Omega)) \cap L^\infty(0,T; H^1(\Omega)) \cap H^1(\Omega_T)\,,$$
$$\mu_+, \mu_- \in L^2(\Omega_T)\,,$$
$$\lambda \in L^2(0,T)\,.$$

Proof: We show the existence of a solution u_δ to (11), (12) by a Galerkin approach, see e.g. Evans [16]. Since such an approach is standard we only briefly describe the arguments and focus on specific aspects that appear in our case due to the non-local structure. We choose $\{w_j\}_{j \in \mathbb{N}_0}$ to be the eigenfunctions of the problem

$$-\Delta w = \kappa w \quad \text{in} \quad \Omega, \ \frac{\partial w}{\partial \nu} = 0 \quad \text{on} \quad \partial\Omega$$

which are chosen to be normalized such that $(w_i, w_j) = \delta_{ij}$ and we also choose w_0 to be constant. A Galerkin approximation is then given as

$$u^N(x,t) = \sum_{j=0}^{N} c_j(t) w_j(x)\,, \tag{13}$$

$$\varepsilon^2(\partial_t u^N, w_j) + \gamma\varepsilon^2(\nabla u^N, \nabla w_j) + \gamma(\psi_\delta'(u^N), w_j) = 0 \text{ for } j = 1,\dots,N \text{ and } t \geq 0\,, \tag{14}$$

$$c_j(0) = (u_0, w_j) \quad \text{for } j = 0,\dots,N\,, \tag{15}$$

$$c_0(t) = c_0(0) \quad \text{for } t \geq 0\,. \tag{16}$$

The last condition guarantees that the approximating solution u^N fulfills the integral constraint. The mean value term does not appear in (14) because the w_j, $j = 1,\dots,N$, are orthogonal to constants. Standard ODE theory gives local existence upto some time $\tilde{T}$ to the above initial value problem. Multiplying (14) by $c_j'(t)$, summation and integration gives

$$\varepsilon \int\limits_{\Omega_T} (\partial_t u^N)^2 + E_\delta(u^N(\tilde{T})) = E_\delta(u^N(0))\,. \tag{17}$$

This energy estimate can be used to show boundedness of the ODE solution and hence global existence to (13)-(16) on $[0,T]$. Furthermore standard compactness and regularity arguments, see e.g. Evans [16], give in the limit $N \to \infty$ the existence of a solution $u_\delta \in L^2(0,T;H^2(\Omega)) \cap H^{1,2}(\Omega_T) \cap L^\infty(0,T;H^1(\Omega))$ of (11), (12) satisfying $\int\limits_\Omega u_\delta(t) = \int\limits_\Omega u_0$ for all $t \in [0,T]$. In particular we obtain

$$\varepsilon \int\limits_{\Omega_T} (\partial_t u_\delta)^2 + E_\delta(u_\delta(T)) \leq E_\delta(u_0) \tag{18}$$

for all $\delta > 0$. Since $\psi_\delta'(u_\delta) \in L^2(0,T;H^1(\Omega))$ we can multiply (11) by $-\Delta u_\delta$ and integrate. After integration by parts we obtain

$$\varepsilon^2 \tfrac{d}{dt} \tfrac{1}{2} \int\limits_\Omega |\nabla u_\delta|^2 + \varepsilon^2\gamma \int\limits_\Omega |\Delta u_\delta|^2 + \gamma \int\limits_\Omega \tfrac{1}{\delta}(\psi_+''(u_\delta) + \psi_-''(u_\delta))|\nabla u_\delta|^2 = -\int\limits_\Omega \gamma\psi_0''(u_\delta)|\nabla u_\delta|^2 = \int\limits_\Omega \gamma |\nabla u_\delta|^2\,.$$

A Grönwall argument gives that $(u_\delta)_{\delta\in(0,2)}$ is uniformly bounded in $L^\infty(0,T;H^1(\Omega))$ and that Δu_δ is uniformly bounded in $L^2(\Omega_T)$. Now elliptic regularity theory gives that $(u_\delta)_{\delta>0}$ is uniformly bounded in $L^2(0,T;H^2(\Omega)) \cap L^\infty(0,T;H^1(\Omega)) \cap H^1(\Omega_T)$. When passing to the limit in (11) we would like to obtain the Lagrange multipliers $\mu_\pm$ as the limit of $\frac{\gamma}{\delta}\psi_\pm(u_\delta)$ and λ as the limit of $\fint\limits_\Omega \gamma\psi_\delta'(u_\delta)$ as δ tends to zero. We hence need to estimate

$$\lambda_\delta := \fint\limits_\Omega \gamma\psi_\delta'(u_\delta) dx$$

in $L^2(0,T)$ uniformly in δ. Multiplying (11) by $u_\delta \pm 1$ gives after integration and integration by parts using $\int_\Omega \partial_t u_\delta = 0$ and $\fint_\Omega u_\delta = m$:

$$\int_\Omega \varepsilon^2 \partial_t u_\delta u_\delta + \gamma\varepsilon^2 \int_\Omega |\nabla u_\delta|^2 + \int_\Omega \gamma \psi_\delta'(u_\delta)(u_\delta \pm 1) = \lambda_\delta (m \pm 1)|\Omega| \,. \tag{19}$$

Since $\psi_\delta'' \geq -1$ we obtain that $\psi_\delta(z) + \frac{1}{2}z^2$ is convex and hence

$$\psi_\delta'(u_\delta)(u_\delta \pm 1) + \tfrac{1}{2}(u_\delta \pm 1)^2 \geq \psi_\delta(u_\delta) - \psi_\delta(\pm 1) \,.$$

Using (19), $\psi_\delta(\pm 1) = 1$ and $|m| < 1$ we get

$$|\lambda_\delta| \leq \tfrac{C}{(1-|m|)|\Omega|} (\|\partial_t u_\delta\|_{L^2} \|u_\delta\|_{L^2} + \int_\Omega u_\delta^2 + 1) \,,$$

where C is a constant that can depend on ε and γ. Since u_δ is uniformly bounded in $H^1(\Omega_T)$ for $\delta \in (0,2)$ we obtain

$$\lambda_\delta \in L^2(0,T) \quad \text{uniformly for } \delta \in (0,2) \,.$$

In this context we refer to [1] for a similar argument in a different context.
Using standard compactness results we obtain the existence of a subsequence, still denoted by (u_δ), and a $u \in L^2(0,T;H^2(\Omega)) \cap H^1(\Omega_T) \cap L^\infty(0,T;H^1(\Omega))$, such that

$$\begin{aligned}
u_\delta &\rightharpoonup u \quad \text{in } L^2(0,T;H^2(\Omega)) \,, \\
u_\delta &\rightharpoonup u \quad \text{in } H^1(\Omega_T) \,, \\
u_\delta &\to u \quad \text{in } L^2(0,T;H^1(\Omega)) \,, \\
u_\delta &\rightharpoonup u \quad \text{in } L^\infty(0,T;H^1(\Omega)) \,, \\
u_\delta &\to u \quad \text{a.e. in } \Omega_T \,.
\end{aligned}$$

The energy estimate (18) gives furthermore

$$\int_\Omega (\psi_+(u_\delta) + \psi_-(u_\delta)) \leq c(\varepsilon,\gamma,u_0)\delta$$

for almost all $t \in [0,T]$. Since $u_\delta \to u$ a.e. in Ω_T we obtain from Fatou's Lemma

$$\begin{aligned}
\int_\Omega (\max(u,1)^2 + \min(u,-1)^2) &= \int_\Omega \lim_{\delta\to 0} (\max(u_\delta,1)^2 + \min(u_\delta,-1)^2) \\
&\leq \liminf_{\delta\to 0} \int_\Omega (\psi_+(u_\delta) + \psi_-(u_\delta)) \\
&\leq \lim_{\delta\to 0} c(\varepsilon,\gamma,u_0)\delta = 0 \,.
\end{aligned}$$

We hence obtain $|u| \leq 1$ a.e. in Ω_T.
Defining

$$\mu_{\delta,\pm} := \pm \tfrac{\gamma}{\delta} \psi_\pm'(u_\delta)$$

we can rewrite (11) as

$$\lambda_\delta = \varepsilon^2 \partial_t u_\delta - \gamma\varepsilon^2 \Delta u_\delta + \gamma \psi_0'(u_\delta) + \mu_{\delta,+} - \mu_{\delta,-}\,. \tag{20}$$

Since $\mu_{\delta,+} \cdot \mu_{\delta,-} = 0$ we obtain from (20) and the a priori estimates on u_δ and λ_δ that

$$\|\mu_{\delta,+}\|_{L^2(\Omega_T)} + \|\mu_{\delta,-}\|_{L^2(\Omega_T)} \leq c(\varepsilon, \gamma, u_0)\,.$$

Hence there exists $\mu_+, \mu_- \in L^2(\Omega_T)$ such that for a subsequence

$$\mu_{\delta,\pm} \rightharpoonup \mu_\pm \quad \text{in } L^2(\Omega_T) \quad \text{as } \delta \searrow 0\,.$$

Since $\mu_\pm$ are the weak limits of nonnegative functions we obtain $\mu_\pm \geq 0$ a.e.. Passing to the limit in a weak formulation of (20) now gives (5) and (6). (7) follows since $u_\delta \to u$ in $L^2(\Omega_T)$. In addition we obtain using the monotonicity of ψ_+' and $\psi_+'(1) = 0$

$$\begin{aligned} \mu_{\delta,+}(u_\delta - 1) &= \tfrac{\gamma}{\delta}\psi_+'(u_\delta)(u_\delta - 1) \\ &= \tfrac{\gamma}{\delta}(\psi_+'(u_\delta) - \psi_+'(1))(u_\delta - 1) \geq 0\,. \end{aligned}$$

Since $u_\delta \to u$ and $\mu_{\delta,+} \rightharpoonup \mu$ in $L^2(\Omega_T)$ we obtain

$$\int\limits_{\Omega_T} \mu_+(u-1) = \lim_{\delta\to 0} \int\limits_{\Omega_T} \mu_{\delta,+}(u_\delta - 1) \geq 0\,.$$

Since $(u-1) \leq 0$ and $\mu_+ \geq 0$ we hence deduce

$$\mu_+(u-1) = 0 \text{ a.e. in } \Omega_T\,.$$

It remains to show uniqueness. Assume that there are two solutions $(u^1, \mu_+^1, \mu_-^1, \lambda^1)$ and $(u^2, \mu_+^2, \mu_-^2, \lambda^2)$. Then we define $\overline{u} = u^1 - u^2$, $\overline{\mu}_\pm = \mu_\pm^1 - \mu_\pm^2$, $\overline{\lambda} = \lambda^1 - \lambda^2$. Multiplying the difference of the equation (5) for u^1 and u^2 with $\overline{u}$ gives after integration and using $\int\limits_\Omega \overline{u} = 0$

$$\begin{aligned} \varepsilon^2 \tfrac{d}{dt} \int\limits_\Omega (\overline{u})^2 + \gamma\varepsilon^2 \int\limits_\Omega |\nabla \overline{u}|^2 + \int\limits_\Omega (\mu_+^1 - \mu_+^2)(u^1 - u^2) \\ - \int\limits_\Omega (\mu_-^1 - \mu_-^2)(u^1 - u^2) = \gamma \int\limits_\Omega (\overline{u})^2\,. \end{aligned}$$

The complementary conditions (8)-(10) imply that the terms $(\mu_+^1 - \mu_+^2)(u^1 - u^2)$ and $-(\mu_-^1 - \mu_-^2)(u^1 - u^2)$ are non-negative. We hence deduce

$$\varepsilon^2 \tfrac{d}{dt} \int\limits_\Omega |\overline{u}|^2 + \gamma\varepsilon^2 \int\limits_\Omega |\nabla \overline{u}|^2 \leq \gamma \int\limits_\Omega |\overline{u}|^2\,.$$

A Grönwall argument now gives uniqueness of u. Hence $\mu_+ - \mu_- - \lambda$ is uniquely given through the equation (5). For all $t \in [0,T]$ we find a $\rho > 0$ such that

$$A_\rho := |\{x \in \Omega |\, |u(x,t)| < 1 - \rho\}| > 0\,.$$

Hence we obtain that for almost all $t \in [0,T]$ the Lagrange multiplier $\lambda(t)$ is uniquely given through

$$\lambda(t) = \tfrac{1}{A_\rho} \int_\Omega (\varepsilon^2 \partial_t u - \gamma\varepsilon^2 \Delta u + \gamma \psi_0'(u)) dx \,.$$

Finally, we obtain that μ_+ and μ_- are uniquely given as

$$\begin{aligned} \mu_+ &= (\lambda - \varepsilon^2 \partial_t u + \gamma\varepsilon^2 \Delta u - \gamma \psi_0'(u))_+ \,, \\ \mu_- &= (-\lambda + \varepsilon^2 \partial_t u - \gamma\varepsilon^2 \Delta u + \gamma \psi_0'(u))_+ \end{aligned}$$

where $(z)_+ := \max(z, 0)$. Here we use the fact that

$$\lambda = \varepsilon^2 \partial_t u - \gamma\varepsilon^2 \Delta u + \gamma \psi_0'(u) \quad \text{a.e. on } \{|u| < 1\} \,.$$

□

Remark 2.1 *i) The variational inequality* $(\mathbf{P}_m)$ *has a unique solution. This follows from a testing procedure similar to the one in Theorem 2.1.*
ii) If $(\mathbf{A}_m)$ *holds we obtain that there exists a solution of* $(\mathbf{P}_m)$ *if and only if (5)-(10) is solvable. In particular, if there is a solution* u *of* $(\mathbf{P}_m)$ *Lagrange multipliers* μ_+, μ_-, λ *exist such that (5)-(10) hold. This follows from the unique solvability of* $(\mathbf{P}_m)$*, Theorem 2.1 and Lemma 2.1.*

Similar to the integral constrained case, see Lemma 2.1 and Theorem 2.1, we can derive the following theorem for (2) without local constraints:

Theorem 2.2 *Let* $\Omega \subset \mathbb{R}^d$ *be a bounded domain which is either convex or has a* $C^{1,1}$*–boundary and let* $u_0 \in H^1(\Omega)$ *such that* $|u_0| \leq 1$ *a.e.. Then for any solution to the Allen-Cahn variational inequality (2) without mass conservation Lagrange multipliers* $\mu_+, \mu_- \in L^2(\Omega_T)$ *exist such that the initial and boundary conditions (6), the complementary conditions (8)-(10) and*

$$0 = \varepsilon^2 \partial_t u - \gamma\varepsilon^2 \Delta u + \gamma \psi_0'(u) + \mu_- - \mu_- \quad \text{a.e. in } \Omega_T$$

hold.

3 Primal-dual active set approach

For the numerical approximation of solutions u of $(\mathbf{P}_m)$ we introduce a primal-dual active set method or equivalently a semi-smooth Newton method [21]. 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 with 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 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 $u^n := u(., t_n)$. Then the time discretization of $(\mathbf{P}_m)$ is given as an Euler-discretization. In this paper we focus on the implicit discretization leading to the following formulation:
$(\mathbf{P}_m^\tau)$ *Given* u^{n-1} *find* $u = u^n \in H^1(\Omega)$ *such that* $\fint_\Omega u dx = m$, $|u| \leq 1$ *a.e. in* Ω *and*

$$(\tfrac{\varepsilon}{\tau}(u - u^{n-1}), \chi - u) + \gamma\varepsilon(\nabla u, \nabla(\chi - u)) + \tfrac{\gamma}{\varepsilon}(\psi_0'(u), \chi - u) \geq 0 \tag{21}$$

for all $\chi \in H^1(\Omega)$ *with* $|\chi| \leq 1$ *and* $\fint_\Omega \chi dx = m$.
For simplicity we denote by u the time discrete solution at time t_n. This disretization 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

$$\begin{aligned} &\min \left\{ \tfrac{\gamma\varepsilon}{2}\|\nabla u\|_{L^2}^2 + \tfrac{\gamma}{\varepsilon}\int_\Omega \psi_0(u) + \tfrac{\varepsilon}{2\tau}\|u - u^{n-1}\|_{L^2}^2 \right\} \\ &\text{s.t.} \quad |u| \leq 1 \ \text{ and } \ \fint_\Omega u dx = m. \end{aligned} \tag{22}$$

As in Lemma 2.1 one can reformulate $(\mathbf{P}_m^\tau)$ by using $1/\varepsilon$ scaled Lagrange-multipliers $\mu_\pm$ on Ω for the inequality constraints $|u| \leq 1$, $\mu := \mu_+ - \mu_-$ and $\lambda \in \mathbb{R}$ for $\fint_\Omega u = m$ and obtain:

$$\mu = \lambda - \tfrac{\varepsilon^2}{\tau}(u - u^{n-1}) + \varepsilon^2\gamma\Delta u - \gamma\psi_0'(u) \qquad \text{a.e. in } \Omega\,, \tag{23}$$

$$\frac{\partial u}{\partial \nu} = 0 \qquad \text{a.e. on } \partial\Omega\,, \tag{24}$$

$$\fint_\Omega u dx = m\,, \tag{25}$$

together with the complementarity conditions

$$|u| \leq 1 \quad \text{a.e. in } \Omega\,, \tag{26}$$

$$\mu_+(u-1) = 0, \mu_-(u+1) = 0 \quad \text{a.e. in } \Omega\,, \tag{27}$$

$$\mu_+ \geq 0, \mu_- \geq 0 \quad \text{a.e. in } \Omega\,. \tag{28}$$

Now the idea is to reformulate the complementarity conditions using active sets based on the primal variable u and the dual variables $\mu_\pm$. Then, for any $c > 0$, (26)-(28) is equivalent to:

$$u = 1 \ \text{ a.e. in } \mathcal{A}^+; \quad u = -1 \ \text{ a.e. in } \mathcal{A}^-; \tag{29}$$

$$\mu = 0 \ \text{ a.e. in } \mathcal{I} := \Omega \setminus (\mathcal{A}^+ \cup \mathcal{A}^-) \tag{30}$$

$$\text{with} \quad \begin{aligned} \mathcal{A}^+ &= \{x \in \Omega \mid c\,(u(x) - 1) + \mu(x) > 0\} \\ \mathcal{A}^- &= \{x \in \Omega \mid c\,(u(x) + 1) + \mu(x) < 0\} \end{aligned}\,. \tag{31}$$

A further equivalent formulation of (26)-(28) is given by the following semi-smooth equation

$$\mathcal{H}(u,\mu) := \mu - \max(0, \mu + c\,(u-1)) + \min(0, \mu + c\,(u+1)) = 0. \tag{32}$$

If the sets $\mathcal{A}^{\pm}$ were known, we would only have to solve a system of equations, namely (23)-(25) together with (29), (30). In particular, given (29), (30) the system (23)-(25) on Ω reduces to an equation for $\lambda \in \mathbb{R}$ and for u on the interface $\mathcal{I}$:

$$0 = \lambda - \tfrac{\varepsilon^2}{\tau}(u - u^{n-1}) + \varepsilon^2\gamma\Delta u - \gamma\psi_0'(u) \qquad \text{a.e. in } \mathcal{I}\,, \tag{33}$$

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{a.e. on } \partial\mathcal{I} \cap \partial\Omega\,, \qquad u = \pm 1 \quad \text{a.e. on } \partial\mathcal{I} \cap \partial\mathcal{A}^{\pm}\,, \qquad \fint_{\Omega} u = m\,. \tag{34}$$

Given now u and λ one can determine μ on $\mathcal{A}^{\pm}$ using (23).
This leads to the idea of the **Primal-Dual Active Set (PDAS) algorithm**:
Given initial active sets $\mathcal{A}_0^{\pm}$ iterate the following steps for $k \geq 0$

1. *Set $u^k = \pm 1$ on $\mathcal{A}_k^{\pm}$ and $\mu^k = 0$ on $\mathcal{I}_k$,*
2. *Solve (33)-(34) for $\lambda^k \in \mathbb{R}$ and u^k on $\mathcal{I}_k$,*
3. *Determine μ^k on $\mathcal{A}_k^{\pm}$ using (23),*
4. *Determine the new active sets $\mathcal{A}_{k+1}^{\pm}$,*
5. *Stop the iteration if $\mathcal{A}_{k+1}^{\pm} = \mathcal{A}_k^{\pm}$, otherwise set $k = k+1$ and goto 1.*

This algorithm is formally equivalent to a Newton algorithm applied to (5)-(7) and (32) (see e.g. [21]). However, we would like to mention that the method is more driven by the current active set than by the current values of u, μ and λ in the sense that different u, μ can lead to the same active sets and hence to the same iteration step. The view point of an active set method also enhances the choice of this method instead of other numerical approaches for solving the Allen-Cahn system. Namely, due to the evolution in time good initial active sets are given. This is much in favour for an active set approach instead of an interior point ansatz where good initial data is difficult to exploit. For analytical reasons the formulation as a Newton iteration is often more convenient and will be presented and used in the next section.
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 $\mu_{\pm} \in L^2(\Omega)$ 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 [23]. Therefore, the pointwise definition of the active sets $\mathcal{A}^{\pm}$ according to (30) is not possible. However we show in the next section that the application of the PDAS-method to the fully discretized problem is possible and the convergence locally superlinear.

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 and discuss the local convergence by employing the formulation as a semi-smooth Newton method. At the end of this section we show with the help of obstacle problems some features of the method.

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. $\Omega = \cup_{T\in\mathcal{T}_h}\overline{T}$. Furthermore we define $h := \max_{T\in\mathcal{T}_h}\{\text{diam } T\}$ 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

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

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 $j \in \mathcal{J}$. Then u_j for $j = 1, \dots \mathcal{J}$ denote the coefficients of the basis representation of u_h in S_h which is given by $u_h = \sum_{j\in\mathcal{J}} u_j\chi_j$ and the vector of coefficients is denoted by $\boldsymbol{u}$.
In order to derive a discretization of our models we set

$$\mathcal{K}_h := \{\eta \in S_h \,|\, |\eta(x)| \le 1 \text{ for all } x \in \Omega\}, \qquad \mathcal{K}_h^m := \{\eta \in \mathcal{K}_h | \fint_\Omega \eta dx = m\}.$$

We introduce also the lumped mass scalar 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 $\fint_\Omega u_h = \sum_{j\in\mathcal{J}} m_j u_j / \sum_{j\in\mathcal{J}} m_j$. Moreover we define the stiffness matrix as $\boldsymbol{A} := (a_{ij})$ with $a_{ij} = (\nabla\chi_j, \nabla\chi_i)$, the mass matrix $\boldsymbol{M} := ((\chi_j, \chi_i)_h) = diag(m_j)$ and the mass vector $\boldsymbol{m} := (m_j)$.

4.2 Finite element approximation and the PDAS-algorithm

We now introduce the following finite element approximations of $(\mathbf{P}_m^\tau)$ given by (21) using $\psi_0'(u) = -u$. In the following we consider a fixed time step $\tau = t_n - t_{n-1}$ and omit in some places the superscript n:
$(\mathbf{P}_{m,h}^\tau)$ *Given* $u_h^{n-1} \in \mathcal{K}_h^m$ *find* $u_h = u_h^n \in \mathcal{K}_h^m$ *such that*

$$(\tfrac{\varepsilon}{\tau}(u_h - u_h^{n-1}) - \tfrac{\gamma}{\varepsilon}u_h, \chi - u_h)_h + \gamma\varepsilon(\nabla u_h, \nabla(\chi - u_h)) \ge 0 \ \ \forall \chi \in \mathcal{K}_h^m. \tag{35}$$

Due to the use of piecewise linear finite elements and nodal basis functions the reformulation of $(\mathbf{P}^{\tau}_{m,h})$ with Lagrange multipliers $\mu_h \in S_h$ and $\lambda \in \mathbb{R}$ can be stated as follows:

$(\mathbf{Q}^{\tau}_{m,h})$ *Find* $u_h \in S_h$, $\mu_h \in S_h$ *and* $\lambda \in \mathbb{R}$ *such that*

$$(\tfrac{\varepsilon^2}{\tau} - \gamma)(u_h, \varphi)_h + \gamma\varepsilon^2(\nabla u_h, \nabla\varphi) + (\mu_h, \varphi)_h - \lambda(1, \varphi) = \tfrac{\varepsilon^2}{\tau}(u_h^{n-1}, \varphi)_h \quad \forall \varphi \in S_h \tag{36}$$

$$\sum_{j\in\mathcal{J}} m_j u_j = m \sum_{j\in\mathcal{J}} m_j, \tag{37}$$

$$(\mu_j)_- \geq 0, \ (\mu_j)_+ \geq 0, \ |u_j| \leq 1, \tag{38}$$

$$(u_j + 1)(\mu_j)_- = (u_j - 1)(\mu_j)_+ = 0 \quad \forall\ j \in \mathcal{J}. \tag{39}$$

As in (32) the complementarity conditions can be rewritten as

$$\mathcal{H}(u_j, \mu_j) = 0 \qquad \forall j \in \mathcal{J}\,. \tag{40}$$

Applying now the PDAS-method presented in Section 3 to $(\mathbf{Q}^{\tau}_{m,h})$ we obtain the following algorithm:

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

0. Set $k = 0$ *and initialize* $\mathcal{A}_0^{\pm}$.

1. Define $\mathcal{I}_k = \mathcal{J} \setminus (\mathcal{A}_k^+ \cup \mathcal{A}_k^-)$.
Set $u_j^k = \pm 1$ *for* $j \in \mathcal{A}_k^{\pm}$ *and* $\mu_j^k = 0$ *for* $j \in \mathcal{I}_k$.

2. Solve the discretized PDE (36) with the non-local constraint (37) to obtain u_j^k *for* $j \in \mathcal{I}_k$ *and* $\lambda^k \in \mathbb{R}$*:*

$$(\tfrac{\varepsilon}{\tau} - \tfrac{\gamma}{\varepsilon})m_j u_j^k + \gamma\varepsilon \sum_{i\in\mathcal{I}_k} a_{ij} u_i^k - \tfrac{1}{\varepsilon} m_j \lambda^k \tag{41}$$
$$= \tfrac{\varepsilon}{\tau} m_j u_j^{n-1} + \gamma\varepsilon\Big(\sum_{i\in\mathcal{A}_k^-} a_{ij} - \sum_{i\in\mathcal{A}_k^+} a_{ij}\Big) \qquad \forall j \in \mathcal{I}_k$$

$$\sum_{i\in\mathcal{I}_k} m_i u_i^k = m \sum_{i\in\mathcal{J}} m_i - \sum_{i\in\mathcal{A}_k^+} m_i + \sum_{i\in\mathcal{A}_k^-} m_i. \tag{42}$$

3. Determine μ_j^k *for* $j \in \mathcal{A}_k^{\pm}$ *using (36):*

$$\mu_j^k = -(\tfrac{\varepsilon^2}{\tau} - \gamma)u_j^k - \gamma\varepsilon^2 \tfrac{1}{m_j} \sum_{i\in\mathcal{J}} a_{ij} u_i^k + \lambda^k + \tfrac{\varepsilon^2}{\tau} u_j^{n-1}.$$

4. Set $\mathcal{A}_{k+1}^+ := \{j \in \mathcal{J} : u_j^k + \frac{\mu_j^k}{c} > 1\}$, $\mathcal{A}_{k+1}^- := \{j \in \mathcal{J} : u_j^k + \frac{\mu_j^k}{c} < -1\}$.

5. If $\mathcal{A}_{k+1}^{\pm} = \mathcal{A}_k^{\pm}$ *stop, otherwise set* $k = k + 1$ *and goto 1.*

Remark 4.1 *We solve (41), (42) by multiplying (42) with* $\frac{\varepsilon}{\tau}$ *and using the conjugate gradient method.*

4.3 Convergence as a semi-smooth Newton method

In [21] it is shown that the mapping $y \to max(0, y)$ from $\mathbb{R}$ to $\mathbb{R}$ is slantly differentiable and one possible slanting function is $G(y) = 1$ for $y > 0$ and $G(y) = 0$ for $y \leq 0$. Setting $\mathcal{A}_h^+ := \{j \in \mathcal{J} : u_j + \frac{\mu_j}{c} > 1\}$, $\mathcal{A}_h^- := \{j \in \mathcal{J} : u_j + \frac{\mu_j}{c} < -1\}$, $\mathcal{I}_h := \mathcal{J} \setminus (\mathcal{A}_h^+ \cup \mathcal{A}_h^-)$ and adapting the above for the *min-max*-function $\mathcal{H}(u_j, \mu_j)$ one derives the slanting function $\mathcal{G}(u_j, \mu_j) = (-c, 0)$ for $j \in \mathcal{A}_h^\pm$ and $\mathcal{G}(u_j, \mu_j) = (0, 1)$ for $j \in \mathcal{I}_h$. We now consider the system (36), (37) and (40) as a problem of finding a root of $\boldsymbol{F}(\boldsymbol{u}, \boldsymbol{\mu}, \lambda) = \mathbf{0}$, by using a semi-smooth Newton method (SSN) $(\boldsymbol{u}^{k+1}, \boldsymbol{\mu}^{k+1}, \lambda^{k+1}) = (\boldsymbol{u}^k, \boldsymbol{\mu}^k, \lambda^k) - \boldsymbol{G}(\boldsymbol{u}^k, \boldsymbol{\mu}^k, \lambda^k)^{-1} \boldsymbol{F}(\boldsymbol{u}^k, \boldsymbol{\mu}^k, \lambda^k)$. We set

$$\mathcal{G}_u(\boldsymbol{u}, \boldsymbol{\mu}) := (g_{ij}^u) \qquad \text{with} \qquad g_{ij}^u := \begin{cases} -c & \text{for } i = j \in \mathcal{A}_h^\pm \\ 0 & \text{elsewhere} \end{cases} \tag{43}$$

$$\mathcal{G}_\mu(\boldsymbol{u}, \boldsymbol{\mu}) := (g_{ij}^\mu) \qquad \text{with} \qquad g_{ij}^\mu := \begin{cases} 1 & \text{for } i = j \in \mathcal{I}_h \\ 0 & \text{elsewhere} \end{cases} \tag{44}$$

and derive for $\boldsymbol{F}$ the slanting function

$$\mathbf{G}(\boldsymbol{u}, \boldsymbol{\mu}, \lambda) = \begin{pmatrix} (\frac{\varepsilon^2}{\tau} - \gamma)\mathbf{M} + \gamma\varepsilon^2\mathbf{A} & \mathbf{M} & -\mathbf{m} \\ \mathcal{G}_u(\boldsymbol{u}, \boldsymbol{\mu}) & \mathcal{G}_\mu(\boldsymbol{u}, \boldsymbol{\mu}) & 0 \\ -\mathbf{m}^t & 0 & 0 \end{pmatrix}. \tag{45}$$

Moreover, it is easy to show the equivalence of the Newton algorithm to the PDAS-I method using as a starting set $\mathcal{A}_0^\pm$ the set given by an initial guess (u^{-1}, μ^{-1}).

Theorem 4.1 *The PDAS-I algorithm converges locally superlinear to the coefficient vector $\boldsymbol{u}^*$ of the solution u_h of the discretized Allen-Cahn variational inequality with mass constraints (35), if at least one mesh point p_j of u_h exists such that $|u_h(p_j)| < 1$ and τ is sufficiently small. (The precise assumption on τ is given in Theorem 4.2).*

Proof: Given at least one inactive mesh point p_j of u_h there exists an open neighborhood U where $j \in \mathcal{I}_h$ for all $(\boldsymbol{u}, \boldsymbol{\mu}, \lambda) \in U$. Since for a fixed discretization there exists only a finite number of possible active sets, only finitely many different $\mathcal{G}_u(\boldsymbol{u}, \boldsymbol{\mu})$ and $\mathcal{G}_\mu(\boldsymbol{u}, \boldsymbol{\mu})$ can occur, as is the case for $\boldsymbol{G}$. In the following theorem we show that $\boldsymbol{G}$ is invertible for all possible active sets with $\mathcal{I}_h \neq \emptyset$. Hence the inverses of $\boldsymbol{G}$ are uniformly bounded on U. The convergence result for the semismooth Newton method in [11, 21] then provides the local superlinear convergence. □

Remark 4.2 *One has to keep in mind that the convergence radius is unknown. In the worst case the radius is so small that the active set of the initial guess is identical to the active set of the solution. Then, however, the solution is obtained in one step. In practice we always observed convergence for a larger convergence radius.*

In order to show the invertibility of $\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{\mu}, \lambda)$ we need a discrete Poincaré inequality: There exists a Poincaré constant $c_h^p > 0$ such that

$$(v, v)_h \leq c_h^p (\nabla v, \nabla v) \quad \forall v \in \boldsymbol{K} \tag{46}$$

with $\boldsymbol{K} := \{v \in S^h \mid \int\limits_\Omega v = 0, v(p_j) = 0 \text{ if } j \in \mathcal{A}_h^\pm\}$. We then obtain the following theorem.

Theorem 4.2 *Assume $\mathcal{I}_h \neq \emptyset$ and let τ be such that*

$$\tau(1 - \tfrac{\varepsilon^2}{c_h^p}) < \tfrac{\varepsilon^2}{\gamma}\,. \tag{47}$$

Then the matrix $\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{\mu}, \lambda)$ is invertible which is equivalent to the unique solvability of (41), (42).

Proof: We show that the kernel of $\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{\mu}, \lambda)$ contains only $\mathbf{0}$ provided that $\mathcal{I}_h \neq \emptyset$. The equation

$$\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{\mu}, \lambda)(\boldsymbol{v}, \boldsymbol{\kappa}, \alpha)^t = \mathbf{0} \tag{48}$$

implies $\boldsymbol{v} \equiv \mathbf{0}$ on the active sets $\mathcal{A}_h^\pm$, $\boldsymbol{\kappa} = 0$ on $\mathcal{I}_h$ and $\boldsymbol{m} \cdot \boldsymbol{v} = \mathbf{0}$. If one can show that $\boldsymbol{v} = 0$ has to hold, we can conclude by using a row $j \in \mathcal{I}_h$ that $\alpha = 0$. Then the first block of rows yield $\boldsymbol{\kappa} = \mathbf{0}$, and hence the assertion holds.
To show $\boldsymbol{v} = 0$ we prove that $\boldsymbol{v} = 0$ is the unique solution of the quadratic optimization problem in $\boldsymbol{v}$ on the inactive set

$$\min_{v \in \boldsymbol{K}} \left[\tfrac{1}{2}(\tfrac{\varepsilon^2}{\tau} - \gamma)(v, v)_h + \tfrac{\gamma\varepsilon^2}{2}(\nabla v, \nabla v)\right] \tag{49}$$

where the first order necessary conditions are given by (48).
We show that (49) is a strictly convex minimization problem. If $\tau \leq \frac{\varepsilon^2}{\gamma}$ this follows immediately. In the other case we need to control $(v, v)_h$ on $\boldsymbol{K}$. Using the Poincaré inequality (46) we obtain

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

We hence obtain that (49) is uniquely solvable if (47) holds. □

In the Allen-Cahn model interfaces have to leading order in ε a thickness $\pi\varepsilon$. Hence we expect that the typical Poincaré constants c_h^p in (46) which depends due to $\boldsymbol{K}$ only on $\mathcal{I}_h$ scales like ε^2. Then the time step restriction is much less severe than $\tau < \frac{\varepsilon^2}{\gamma}$ which is usually taken, see [12], in fully implicit time discretizations of the Allen-Cahn equation. This shall be illustrated in the following remark.

Remark 4.3 *i) The Poincaré constant c_h^p and therefore the maximal possible time step τ, see (47), can be estimated by a Poincaré constant $c_p(\mathcal{I}_h)$ for a continuous*

problem. We introduce $\Omega_{\mathcal{I}_h} := int\{x \in \Omega \mid x \in \ supp\, \chi_j,\ j \in \mathcal{I}_h\}$*where* $intA$ *is the interior of a set* A*. Let* $c_p(\mathcal{I}_h)$ *be a Poincaré constant such that*

$$(v,v) \leq c_p(\mathcal{I}_h)\,(\nabla v, \nabla v) \quad \forall v \in H_0^1(\Omega_{\mathcal{I}_h}) \ \ \textit{with} \fint_\Omega v dx = 0\,.$$

Since $(v,v)_h \leq (d+2)(v,v)$ *for all* $v \in S^h$ *where* d *is the space dimension, see [24] Lemma 11, we obtain* $(v,v)_h \leq c_p(\mathcal{I}_h)\,(d+2)(\nabla v, \nabla v)$ *for* $v \in \boldsymbol{K}$ *and hence*

$$c_h^p \leq c_p(\mathcal{I}_h)\,(d+2)\,.$$

ii) In one dimension it is shown in [17] Lemma 6.2 that for an interface $\mathcal{I}$ *of width* $\pi\varepsilon$ *a Poincaré constant* $c_p(\mathcal{I}) = \frac{1}{4}\varepsilon^2$ *is obtained.*
Then, given a good numerical approximation $\mathcal{I}_h$ *of* $\mathcal{I}$ *no restriction at all has to be enforced for the time step* τ *in order to show unique solvability.*

Remark 4.4 *We can also solve a semi-implicit discretization, i.e. in case that we replace* $\psi_0(u_h)$ *by* $\psi_0(u_h^{n-1}) + \psi_0'(u_h^{n-1})(u_h - u_h^{n-1})$*, respectively* $\psi_0'(u_h)$ *by* $\psi_0'(u_h^{n-1})$ *and in (36)* $\gamma(u_h, \varphi)_h$ *by* $\gamma(u_h^{n-1}, \varphi)_h$*, with a primal dual active set algorithm. In this case the resulting linear systems are always solvable, since in (49) the negative term disappears. However it will turn out that the fully implicit time discretization is much more accurate.*

4.4 Further features of the PDAS-approach

Although local superlinear convergence is shown, we can face, as mentioned in Remark 4.2, slow convergence or oscillatory behavior of the method, if the starting sets or the choice of c is not appropriate. For both we give here examples.

i) *An example with slow convergence.*

There are situtations in which the active set approach for the discrete obstacle problem belonging to (4) converges very slowly. To demonstrate this we discuss the following explicit example. We choose $\Omega = (-1,1)$, $u_D \equiv 1$, $\varphi \equiv 0$, $\psi \equiv \infty$, $F(u) = -8\int_{-1}^1 u dx$ and $\hat{\mathcal{K}} = \mathcal{K}$, i.e. $\min \int_{-1}^1 \frac{1}{2}|\nabla u|^2 - 8(u-1)dx$ for $(u-1) \in H_0^1(-1,1)$ and $u \geq 0$. The explicit solution is given as $u(x) = 4max(x - \frac{1}{2}, 0)^2$ for $x > 0$ and for $x < 0$ by its reflection. The discretization shall be given by the help of S_h and a uniform grid with step size $h = 1/N$, $N \in \mathbb{N}$ which leads to $u_0 = u_N = 1$ and $-\frac{1}{h^2}(u_{j+1} - 2u_j + u_{j-1}) - 8 = \mu_j$ for $j = 1, \ldots, N-1$ together with $u_j = 0$ for $j \in \mathcal{A} = \{j \,|\, cu_j + \mu_j < 0\}$ and $\mu_j = 0$ for $j \notin \mathcal{A}$. Taking as initial active set $\mathcal{A}_0$ all interior nodes we obtain that the first iterate u^0 is equal to zero at all interior nodes. Consequently for all nodes j with two active neighbours we obtain $\mu_j^0 = -8$. In conclusion we obtain that the active set can shrink in this and later iterations only by one mesh point at each boundary. This makes the algorithm very slow for small h.

Similar situations often appear in practical computations and hence either regularization as for example discussed in [23] has to be applied, or a nested approach, as we propose and employ in Section 5.1, has to be used.

ii) *Oscillations in the bilateral case.*
In the obstacle problem (4) we choose $\Omega = (-1,1)$, $u_D \equiv 0$, $\varphi \equiv -1$, $\psi \equiv 1$, $F \equiv 0$ and $\hat{\mathcal{K}} = \mathcal{K}$, i.e. $\min \int_{-1}^{1} \frac{1}{2}|\nabla u|^2 dx$ for $u \in H_0^1(-1,1)$ with $|u| \leq 1$. The exact solution is of course $u \equiv 0$. We again discretize the problem in S^h on a uniform grid with step size $h = \frac{1}{N}$, $N \in \mathbb{N}$ and obtain $\mu_j = \frac{1}{h^2}(u_{j+1} - 2u_j + u_{j-1})$ for $j = -N+1, \ldots, N-1$, $u_{_N} = u_N = 0$ together with $u_j = 1$ for $j \in \mathcal{A}^+ = \{j \,|\, c(u_j - 1) + \mu_j > 0\}$, $u_j = -1$ for $j \in \mathcal{A}^- = \{j \,|\, c(u_j + 1) + \mu_j < 0\}$ and $\mu_j = 0$ for $j \notin \mathcal{A}^{\pm}$. Choosing $\mathcal{A}_0^+ = \{0\}$, $\mathcal{A}_0^- = \{1\}$ we obtain $\mu_j^0 = 0$ if $j \neq 0, 1$,

$$u^0(x) = \begin{cases} 1 + x & \text{if } x \in [-1, 0]\,, \\ 1 - \frac{2}{h}x & \text{if } x \in [0, h]\,, \\ \frac{1}{1-h}(x-1) & \text{if } x \in [h, 1] \end{cases}$$

and as a consequence $\mu_0^0 = -\frac{2+h}{h^2} < \frac{2}{h^2}$, $\mu_1^0 = \frac{1}{h^2}(2 + \frac{h}{1-h}) > \frac{2}{h^2}$. In conclusion we obtain $\mathcal{A}_1^+ = \{1\}$, $\mathcal{A}_1^- = \{0\}$ if $c < \frac{1}{h^2}$. A similar argument shows that in this case $\mathcal{A}_2^+ = \{0\} = \mathcal{A}_0^+$ and $\mathcal{A}_2^- = \{1\} = \mathcal{A}_0^-$ and hence the iterates oscillate between two states.
Hence c has to be chosen larger than $\frac{1}{h^2}$ in order to avoid such unwanted oscillation. This shows clearly that the local superlinear convergence radius of the PDAS-method depends on the choice of c. In fact, we observed similar oscillations in our numerical computations also in higher space dimensions. In addition the above example shows that for bilateral constraints the parameter c will not drop out after one iteration, compare [21] for the unilateral case.

5 Computational results

In this section we discuss some computational results. In Subsection 5.1 we apply the PDAS-method to one-dimensional obstacle problems with mass constraints. Here exact solutions can be derived and convergence properties can be analyzed. We give one example, where one sees slow convergence similar to the example in 4.4(i) and suggest and apply a nested iteration approach to overcome this problem. Moreover, we apply the method also in a case where the Langrange multiplier is only a measure and discuss its results.
In Subsection 5.2 we consider the Allen-Cahn variational inequality without integral constraint. We compare our PDAS-method to the widely used projected SOR method and show its efficiency. Furthermore, we see that the PDAS-method allows larger time steps resulting in a speed up without loss of accuracy. In Subsection 5.3 we present some results for the non-local Allen-Cahn equation. The computation times obtained are similiar to the local Allen-Cahn variational inequalities and

again large time steps are possible. Furthermore, we consider one example where the explicit solution is known and show that implicit discretization provides higher accuracy than the semi-implicit one. Finally, we present numerical simulations in two and three space dimensions for the Allen-Cahn variational inequality with volume constraint.

Unless it is otherwise stated, we take $\varepsilon = \frac{1}{16\pi}$ and $\gamma = 1$. Furthermore we present numerical results in two and three space dimensions, where we take $\Omega = (-1,1)^2$ and, respectively, $\Omega = (-1,1)^3$.

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 [14, 15] 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 1.2 (see Schmidt and Siebert [28]) for adaptivity and we implemented the same mesh refinement strategy as in Barrett, Nürnberg and Styles [2], i.e. a fine mesh is constructed where $|u_h^{n-1}| < 1$ with a coarser mesh present in the bulk regions $u_h^{n-1} = \pm 1$. For the computations in Subsections 5.2 and 5.3 we take the minimal diameter of an element $h_{min} = 3.91 \times 10^{-3}$ and the maximal diameter $h_{max} = 6.25 \cdot 10^{-2}$, unless otherwise stated. The time step is chosen as $\tau = 6.25 \cdot 10^{-5}$.

For the standard Allen-Cahn variational inequality, i.e. without mass constraint, there is no Lagrange multiplier λ and (42) does not need to be considered. In each PDAS iteration one has to solve the linear system (41) without the variable λ present. For both the local and the non-local Allen-Cahn variational inequality we use the conjugate gradient method to solve the linear system (41)-(42), see also Remark 4.1.

5.1 Non-local obstacle problem

We consider the one-dimensional version of the non-local obstacle problem (4) introduced in Section 1. In particular, we set $\Omega = (-1,1)$, $u_D \equiv 1$, $\varphi \equiv 0$, $\psi \equiv \infty$ and $m = 2$, i.e. we study

$$\min \textstyle\int_{-1}^{1} \frac{1}{2}|u'|^2 dx - F(u-1) \text{ for } (u-1) \in H_0^1 \text{ with } u \geq 0 \text{ and } \fint_{-1}^{1} u dx = 2.$$

The numerical results presented are obtained from a piecewise linear finite element approximation of u and the Lagrange multiplier μ on a uniform grid with mesh size $h = \frac{2}{N+1}$, where again mass lumping is used. Then, a primal-dual active set method like the one introduced in Section 4 is applied. In the computations the active set is initialized to be all nodes j with $-0.5 \leq p_j \leq 1.6$.

We consider two cases for the forcing term F, for which the obstacle problems can be solved explicitely. In both cases the solutions are symmetric. Hence we only list the solutions on $[0,1]$.

Case 1: $F(u) = \overline{f} \int_{-1}^{1} u(x)\chi_{[-a,a]}(x)dx$ where $\bar{f}$ is a constant and $a \in (0,1)$.
One can show similar to the presented Allen-Cahn problem, that there exist Lagrange multipliers $\lambda \in \mathbb{R}$ and $\mu \in L^2(-1,1)$ such that $-u'' - \bar{f}\chi_{[-a,a]} = \lambda + \mu$ and $u \geq 0,\ \mu \geq 0,\ u\mu = 0$ on $[-1,1]$. Herewith, we obtain

$$u(x) = \begin{cases} 0 & 0 \leq x < b, \\ -\frac{\lambda+\bar{f}}{2}(b-x)^2 & b < x < a, \\ -\frac{\lambda}{2}(1-x)^2 + 1 + c_l(1-x) & a < x \leq 1, \end{cases}$$

where λ, b and c_l can be obtained using the continuity of u and u' at $x = a$ and the integral constraint $\int_{-1}^{1} u dx = 2$.
In Figure 1 we display plots of the approximate solution u_h obtained for several $\bar{f}$ and with $a = 0.3$. In Table 1 the errors in u, λ and μ as well as the number of

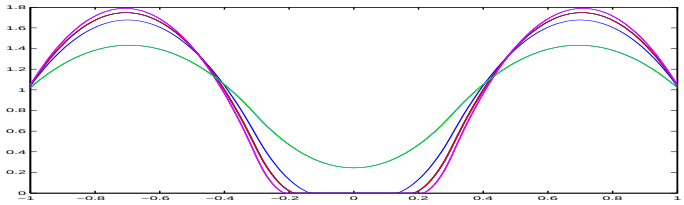


Figure 1: Case 1 with $a = 0.3$ and $\overline{f} = -20$ (green), $\overline{f} = -40$ (blue), $\overline{f} = -60$ (red) and $\overline{f} = -80$ (magenta)

h	$\|\|u-u_h\|\|_{H^1}$	$\|\lambda - \lambda_h\|$	$\|\|\mu - \mu_h\|\|_2$	# of Newton iter.	CPU[s] time
$1.00 \cdot 10^{-2}$	$2.407 \cdot 10^{-1}$	$3.496 \cdot 10^{-1}$	$2.126 \cdot 10^{-1}$	12	1.10
$5.00 \cdot 10^{-3}$	$1.206 \cdot 10^{-1}$	$1.749 \cdot 10^{-1}$	$1.071 \cdot 10^{-1}$	22	14.05
$2.50 \cdot 10^{-3}$	$6.048 \cdot 10^{-2}$	$8.778 \cdot 10^{-2}$	$5.393 \cdot 10^{-2}$	43	223.17
$1.25 \cdot 10^{-3}$	$3.024 \cdot 10^{-2}$	$4.382 \cdot 10^{-2}$	$2.688 \cdot 10^{-2}$	85	4182.30
$1.25 \cdot 10^{-3}$	$3.024 \cdot 10^{-3}$	$4.382 \cdot 10^{-2}$	$2.688 \cdot 10^{-2}$	16 nested	185.20

Table 1: Case 1 with $\bar{f} = -80$ and $a = 0.3$

PDAS-iterations and CPU-times are listed for various mesh sizes h for $\bar{f} = -80$. We see the convergence order h of the H^1-norm in u, in λ, of the L^2-norm in μ and the approximation of b is essentially only limited by the location of the mesh points. The number of Newton iterations approximately doubles as h is halved. This is because we are in a similar situation to the one in Section 4.4. From the second iteration on the active set only moves two mesh point per primal dual active set iteration, hence leading to a very slow convergence.
We overcome this issue by applying nested iteration, i.e. we first solve the problem on a coarse mesh and then use this solution as initial data for the problem on the

next finer mesh and repeat this until the finest mesh is reached. As coarsest mesh we used $h = 10^{-2}$ and then halved the mesh sizes up to $h = 1.25 \cdot 10^{-3}$. With this nested approach we drastically reduced the CPU time for $h = 1.25 \cdot 10^{-3}$ from 4182.3 seconds to 185.2 seconds. The number of Newton iterations that were used, starting with the coarse grid, were 12, 1, 1 and 2. One clearly notices that the main work is done on the coarse grid. We suggest this approach for all cases where the number of PDAS iterations is large already on a coarse grid.

Case 2: $F(u) = gu(0) = g\delta_0(u)$, where δ_0 is the Dirac distribution at 0 and $g \in \mathbb{R}$.

It can be shown, see for example [25], that there exists Lagrange multipliers $\lambda \in \mathbb{R}$, $\mu \in (H_0^1(-1,1))^*$ such that $-u'' - [u'(0)]_-^+ \delta_0 - g\delta_0 = \lambda + \mu$, and $u \geq 0,\ \mu \geq 0,\ \mu(u) = 0$ where $[.]_-^+$ denotes the jump of a quantity. We want to emphasize that μ may be only a measure in this case, and hence the complementarity conditions cannot be formulated pointwise.
Nevertheless, for $g \in [-8, 0]$ we obtain

$$u(x) = \tfrac{3}{8}g(1-x)^2 + 1 - \tfrac{g}{4}(1-x) \quad \text{if } x > 0\,,$$

$\lambda = -\frac{3}{4}g$ and the Lagrange function $\mu \equiv 0$. The bound is never strictly active.
For $g < -8$ we have

$$u(x) = -3(1-x)^2 + 2(1-x) + 1 \quad \text{if } x > 0\,,$$

$\lambda = 6$ and the measure $\mu = (-8-g)\delta_0$.
In Figure 2 we display plots of the approximate solution u_h obtained for several values of g including a case of $g < -8$. In Table 2 we list the results for $g = -2$.

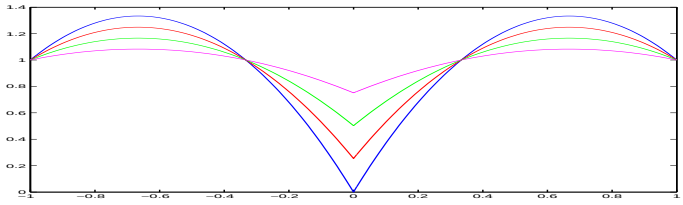


Figure 2: Case 2 with $g = -2$ (magenta), $g = -4$ (green), $g = -6$ (red), $g = -12$ (blue)

As in Case 1 we obtain a convergence rate of h for u and λ. The Lagrange function $\mu = 0$ is always determined exactly. Due to the generic choice of the initial active set, which we chose not to be empty, we always need 2 Newton iterations.

h	$\|\|u - u_h\|\|_{H^1}$	$\|\lambda - \lambda_h\|$	# of Newton iter.
$1.000 \cdot 10^{-2}$	$1.178 \cdot 10^{-2}$	$7.463 \cdot 10^{-3}$	2
$5.000 \cdot 10^{-3}$	$5.900 \cdot 10^{-3}$	$3.741 \cdot 10^{-3}$	2
$2.500 \cdot 10^{-3}$	$2.953 \cdot 10^{-3}$	$1.873 \cdot 10^{-3}$	2
$1.250 \cdot 10^{-3}$	$1.477 \cdot 10^{-3}$	$9.371 \cdot 10^{-4}$	2

Table 2: Case 2 with $g = -2$

h	$\|\|u-u_h\|\|_{H^1}$	$\|\lambda-\lambda_h\|$	$\|8+g+h\mu_h(0)\|$	# of Newton iter.
$1.0\cdot10^{-2}$	$4.229\cdot10^{-2}$	$6.000\cdot10^{-4}$	$6.060\cdot10^{-2}$	1
$5.0\cdot10^{-3}$	$2.118\cdot10^{-2}$	$1.500\cdot10^{-4}$	$3.015\cdot10^{-2}$	1
$2.5\cdot10^{-3}$	$1.060\cdot10^{-2}$	$3.750\cdot10^{-5}$	$1.504\cdot10^{-2}$	1
$1.25\cdot10^{-3}$	$5.301\cdot10^{-3}$	$9.368\cdot10^{-6}$	$7.509\cdot10^{-3}$	1

Table 3: Case 2 with $g=-12$

The results for $g=-12$ are given in Table 3. Here, we have the specific situation that μ is a measure and we have a primal-dual active set, which is given by the single point $x=0$. The approximations always determine this set exactly for any tested h in one iteration. Moreover, the convergence rate in u is h and in λ it is even h^2. As approximation for the point mass at $x=0$ with mass $-8-g$ we get $\mu_h=0$ everywhere except on $(-h,h)$ where the mass is given by $h\mu_h(0)$. The corresponding error behaves like h.

5.2 Allen-Cahn variational inequality without mass constraints

We begin by comparing the PDAS algorithm with the standard projected SOR (pSOR) algorithm that is often used to solve $(\mathbf{P}_h)$, see [12]. In particular we take the simple problems of a shrinking circle in $\mathbb{R}^2$ and a shrinking sphere in $\mathbb{R}^3$ with radius 0.45 and centre 0 and we compare the CPU times as well as the relative error at $T=0.01$ of the two algorithms. In the tables CPU total gives the computation time needed for the program to reach the time T, whereas CPU solver gives the computation time needed for the solver only. For the projected SOR method this is the computation time needed to solve the system of equations using a projected SOR algorithm; for the PDAS method it is the time needed to solve for u_j^k the linear system of equations (41) without λ (there is no mass constraint present) using the conjugate gradient. We calculate the relative error by taking the zero level set of u_h as approximation of the sharp interface and compare it to the solution of the sharp interface formulation for which the radius $R(t)$ at time t is given by the ODE $\frac{d}{dt}R(t)=-\frac{1}{R(t)}, R(0)=0.45$ [5]. To be precise we measure the error between the intersection points of the positive x_1-axis with the circle and with the zero level set of u_h. There may be minor variances of u_h in the other directions, but they have been negligible in our experiments. At $T=0.01$ we have $R(T)=0.4272$ and the circle disappears at $t=0.10125$. For the sphere the exact solution is given by the ODE $\frac{d}{dt}R(t)=-\frac{1}{2R(t)}, R(0)=0.45$, since the mean curvature is defined to be the sum of the principle curvatures. Hence, the sphere shrinks twice as fast as the circle. We expect essentially the relative errors for the projected SOR and the PDAS to be of the same size, since we are solving the same equation on the interface. However, for the SOR method we use $|u_j^{k,l}-u_j^{k,l-1}|<tol$ as stopping criterion while for the

cg-method we use the residual. We chose and fixed the tolerances in such a way that the relative errors are almost the same for the smallest time step. For the larger time step the PDAS method resulted in slightly higher accuracy. Futhermore, given averages of the degree of freedoms (DOFs) and the PDAS-iterations are averages over the time.

time step	CPU [s] total		CPU [s] solver		DOFs circa	rel. error		PDAS-iter. ø
	pSOR	PDAS	pSOR	PDAS		pSOR	PDAS	
$6.25{\cdot}10^{-5}$	23.57	20.59	7.47	2.77	15500	$9.64{\cdot}10^{-3}$	$7.61{\cdot}10^{-3}$	2.1
$5.00{\cdot}10^{-4}$	13.90	7.43	10.12	3.27	16500	$1.63{\cdot}10^{-2}$	$5.23{\cdot}10^{-3}$	3.2
$1.00{\cdot}10^{-3}$	-	6.59	-	3.44	17700	-	$6.11{\cdot}10^{-3}$	4.2

Table 4: Projected SOR method vs. PDAS-method - CPU and error for various time steps in 2D at $T = 0.01$

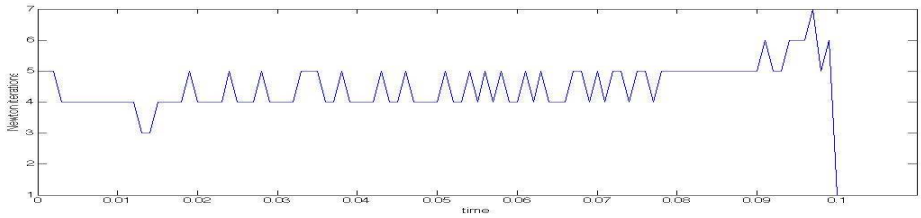


Figure 3: Number of Newton iterations for a shrinking circle

In the two dimensional case the results are listed in Table 4 and Figure 3. For both time steps $\tau = 6.25 \cdot 10^{-5}$ and $\tau = 5 \cdot 10^{-4}$ our PDAS-method is faster. In case of $\tau = 5 \cdot 10^{-4}$ it is even about 46% faster than the pSOR method. Furthermore, it is more stable since for the pSOR method the time step has to be less than $6 \cdot 10^{-4}$ to converge. In Theorem 4.2 and the following remark we have shown that the time step can be much larger for the PDAS-method, which is also observed in the above calculations. In particular, in the numerical experiments the PDAS-method worked for time steps as long as the mean curvature does not become too high. If the solution is only of interest at time T, large time steps are favoured and with the PDAS-method the computation time can be reduced significantly.

In addition, larger time steps can be taken for the PDAS method without loosing accuracy. This is in contrast to the projected SOR method which gives a higher error for $\tau = 5 \cdot 10^{-4}$.

We note that the CPU time for the SOR solver in the projected SOR method increases for larger time steps although there are less equations to be solved. This is due to the initial guess, which is the solution of the previous time step, being further away from the solution of the linear system in the current time step. The conjugate gradient solver does not depend that strongly on a good initial guess due to a higher convergence rate and converges faster than the SOR method for larger time steps.

time step	CPU [s] total		CPU [s] solver		DOFs ø mill.	rel. error		PDAS-iter. ø
	pSOR	PDAS	pSOR	PDAS		pSOR	PDAS	
$6.25{\cdot}10^{-5}$	12447	11350	3609	1697	1.9	$8.48{\cdot}10^{-3}$	$8.95{\cdot}10^{-3}$	4.2
$5.00{\cdot}10^{-4}$	6072	2500	4713	1029	2.1	$4.22{\cdot}10^{-3}$	$2.63{\cdot}10^{-3}$	4.7
$1.00{\cdot}10^{-3}$	-	2216	-	1290	2.3	-	$3.18{\cdot}10^{-3}$	6.1

Table 5: Projected SOR method vs. PDAS method - CPU and error for various time steps for 3D at $T = 0.01$

Moreover, the CPU time for the solver in the PDAS-method is significantly lower than for the pSOR method, since the system of equations are considerably smaller. However, the total CPU time does not decrease quite that much because extra time is needed to determine the Lagrange multiplier μ_j^k and to set the active and inactive sets.
Note that in contrast to the PDAS-method for the pSOR method we are restricted to use the SOR solver. In this paper we are using the conjugated gradient algorithm in the PDAS-method. However, a different solver may additionally speed up the PDAS-method. This is a subject for further research.
The average number of Newton iterations increases for larger time steps, since the interface, and hence the active set moves more in a single time step. The dependence of the number of Newton iterations on the speed of the interface movement can also be observed in time (see Figure 3). When the radius becomes smaller the number of Newton iterations increases. Note that the circle disappears at $t = 0.10125$.

For the three dimensional test problem the same behaviour is observed as in the two dimensional case, see Table 5. We note in particular that for $\tau = 5.00 \cdot 10^{-4}$ our method is 2.4 times faster than the pSOR method. This is mostly due to the decreased computation time needed to solve the linear systems. Since the mean curvature is defined as the sum of all principle curvatures the curvature of the sphere is twice as high as the curvature of a circle with the same radius. This has the effect that the interface changes more rapidly and hence, the number of Newton iterations is significantly higher for the three dimensional case. Moreover, we see that in higher dimensions it is essential to be able to use large time steps for the study of the interface at a specific time T, which is possible with the PDAS-method. In comparison to the pSOR method we obtain u_h with a speed up of 64%. Further time reduction could be possible with another choice of linear algebra solver, as mentioned before.
Finally, we would like to mention, that most existing literature concentrates up to now on the explicit discretization where with the use of mass lumping a non-linear system of equations has to be solved [5, 19, 26]. The explicit discretization has the usual stability restriction for parabolic problems, $\tau \leq Ch^2$. Since we need $h \ll \varepsilon$ this time step restriction is very strong. In contrast, for the implicit discretization in time in combination with the suggested PDAS-method, which we study in this

time step	CPU [s] total	CPU [s] solver	DOFs circa	rel. error	Newton iter. ø
$6.25 \cdot 10^{-5}$	19.75	4.42	16000	$4.08 \cdot 10^{-4}$	1.07
$5.00 \cdot 10^{-4}$	7.06	3.46	17000	$4.04 \cdot 10^{-4}$	1.50
$1.00 \cdot 10^{-3}$	6.74	3.95	17500	$4.00 \cdot 10^{-4}$	2.00

Table 6: PDAS-method for the volume conserved Allen-Cahn equation at $T = 0.01$

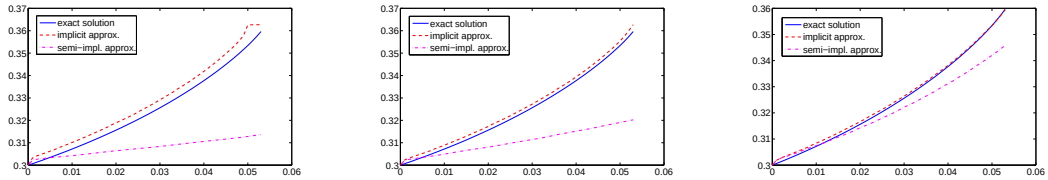
paper, we only have the restriction (47). In our computations we were able to reduce the time step down to $\tau = 1.0 \cdot 10^{-3}$.

5.3 Allen-Cahn variational inequality with integral constraint

5.3.1 Numerical comparison with analytically known solutions

Example 1: To clarify the difference between problems without and with integral constraints we consider the same model problem as in the previous section, i.e. a circle of radius 0.45 as initial interface, but now with volume conservation. Now the circle should keep its shape and stay stationary. Therefore we expect only minor changes in the active set due to discretization errors. This behaviour can be seen in Table 6 looking at the numbers of Newton iterations averaged over time, which are between 1 and 2. Only for the first iteration we need 3 or 4 Newton iterations after which the number of Newton iterations goes down to 2 and soon stabilizes at only one Newton iteration per time step iteration. The CPU time does not increase compared to the computations for the shrinking circle. As before we use the conjugate gradient method and essentially we only add one more row and column to the linear system of equations. Again large time steps can be used and speed up the calculation without loss of accuracy.
Example 2: Next, we take two spheres with radii r_1 and r_2 which do not intersect, here $r_1(0) = 0.3$ and $r_2(0) = 0.2$ with centres $(-0.5, 0)$ and $(0.5, 0)$. This results for the sharp interface problem in $r_1' = -\frac{1}{r_1} + \lambda$, $r_2' = -\frac{1}{r_2} + \lambda$ together with the condition of volume conservation $0 = \frac{1}{2}(r_1^2 + r_2^2)'$, which we also can solve analytically [27]. The larger circle grows while the smaller one disappears roughly at time $T = 0.053$s. Again we compared the radii of the sharp interface solutions with the approximations we obtained using the PDAS-method. We employed a semi-implicit as well as an implicit discretisation in time. In Figure 4 the radii of the larger circle over time are displayed for all three solutions for three different time step sizes. Here the radii of the approximations are determined as in Section 5.2. The behaviour for the smaller one is essentially the same and therefore omitted. The semi-implicit approximation leads to very poor accuracy for larger time steps, in particular at growing time t. Although accuracy improves for smaller time steps it remains worse than the approximation obtained for implicit discretisation in time. Hence, even though there is no time step restriction for the semi-implicit time discretisation

the time step still needs to be very small to achieve accurate approximations. The implicit discretisation leads to better results.



$\tau = 1.0 \cdot 10^{-3}$ $\tau = 5.0 \cdot 10^{-4}$ $\tau = 6.25 \cdot 10^{-5}$

Figure 4: Comparison: sharp interface solution vs. semi-implicit and implicit PDAS-approximation.

For the implicit discretization the absolute error in time is given in Figure 5. The errors are of order 10^{-3} but increase significantly when the circle with initial radius $r_2(0) = 0.2$ becomes very small. Close to this singularity a smaller time step achieves higher accuracy, whereas if the radii of the circles are big enough larger time steps can be chosen. This clearly indicates that an adaptive choice of the time step would be favourable, which shall be studied in the future.

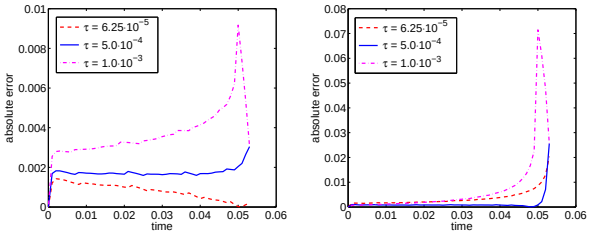


Figure 5: Absolute error between sharp interface solution and approximations (using implicit discretisation in time) of r_1 (left) and r_2 (right) for various time steps

5.3.2 Simulations for the Allen-Cahn equation with integral constraint

For the first simulation (Figure 6) of interface evolution with volume constraint in two space dimensions we set the initial values for the order parameter u randomly between -0.1 and 0.1, i.e. there are no pure phases present. Already at time $t = 0.002$ grains start to form and grow and at $t = 0.003$ we have two phases (red and blue) separated by a diffuse interface. Now the interface moves according to motion by mean curvature but preserving the volume of both phases. That means that closed curves turn into circles and shapes with less volume shrink and disappear while at the same time shapes with the highest volume will grow. At the end (i.e. when the problem becomes stationary) there are three different shapes we can obtain: a circle, a quarter of a circle in one of the corners (see Figure 6) and a straight vertical or horizontal line dividing the two phases.
For the computation in Figure 7 we use a three dimensional domain with one of the phases being a dumbbell. For this computation we had to take a coarser mesh

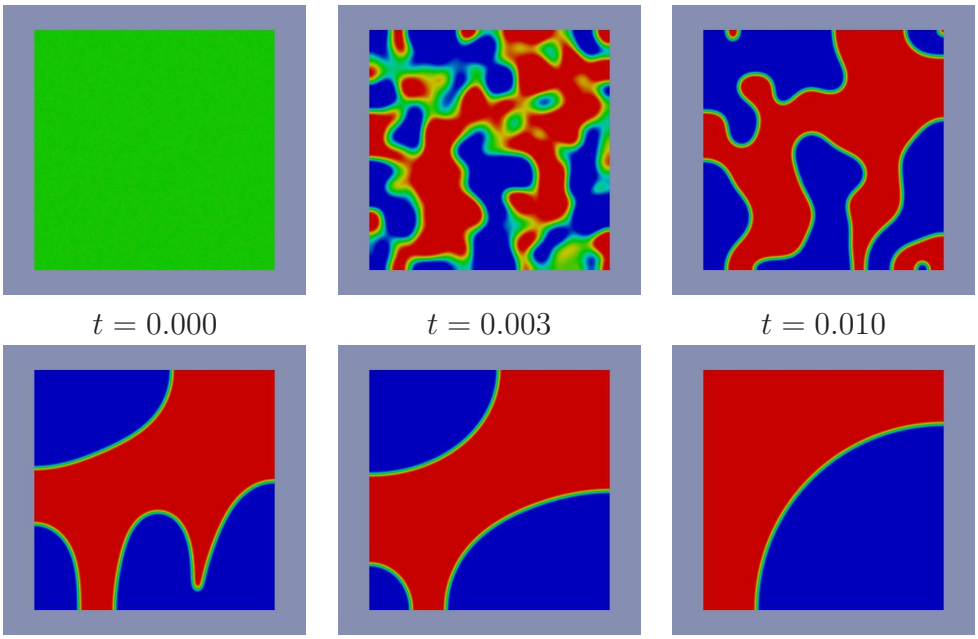


Figure 6: Volume controlled Allen-Cahn equation (2d) with random initial data (varying between -0.1 and 0.1)

due to memory restrictions. We used $h_{min} = 7.81 \cdot 10^{-3}$ and $h_{max} = 6.25 \cdot 10^{-2}$. Without the volume conservation the dumbbell would dissect and the two spheres would shrink and disappear. The volume conservation forces the dumbbell to turn into an ellipsoid before turning into a sphere and finally becoming stationary.

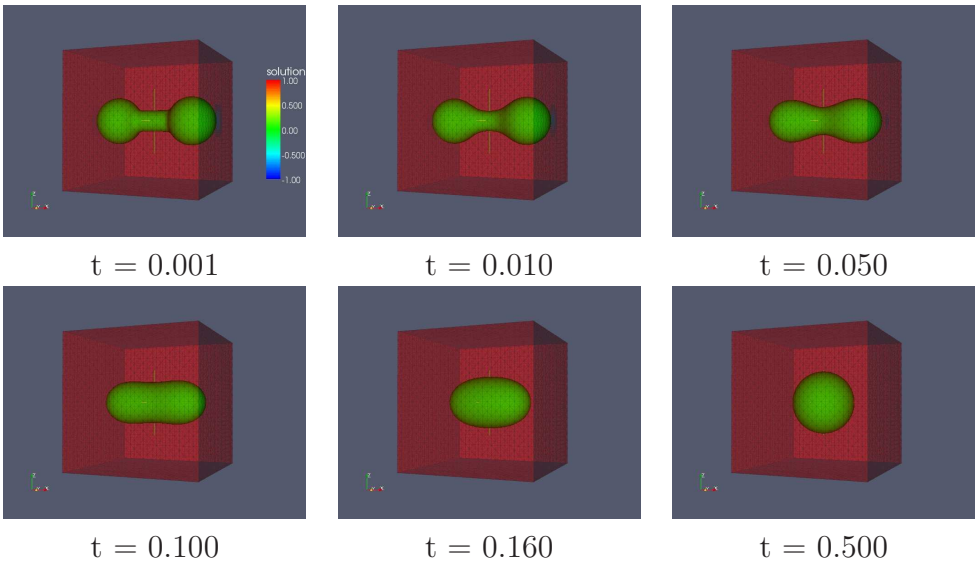


Figure 7: Volume controlled Allen-Cahn equation with a dumbbell as initial data

5.4 Conclusions

In this paper we introduced, analyzed and applied a primal-dual active set method for local and non-local Allen-Cahn variational inequalities. Our approach enables us to use implicit discretisation in time and hence large time steps can be taken as well as higher accuracy is obtained. Only close to singularities smaller time steps are needed to achieve accurate approximations. Therefore, an adaptive time step strategy may be used which will be subject for further research. Thus far we have used the conjugate gradient method to solve the linear problem in each iteration, but a different linear algebra solver would lead to a further speed up of our method.

Acknowledgement.
This work was supported by the SPP 1253 "Optimization with Partial Differential Equations" of the German science foundation (DFG) through the grant BL 433/2-1 and by the Vielberth foundation. Also the fourth author was supported by the EPSRC grant EP/D078334/1.

References

[1] 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.

[2] 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.

[3] 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.

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

[5] 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

[6] Blowey, J. F. and Elliott, C. M., *The Cahn-Hilliard gradient theory for phase separation with nonsmooth free energy. I*, Mathematical Analysis. European J. Appl. Math. 2, no. 3 (1991), 233–280.

[7] Bonnans, J.F., Bessi Fourati, R., and Smaoui, H., *The obstacle problem for water tanks*, J. Math. Pures Appl. 82 (2003), 1527–1553.

[8] Bourdin, B. and Chambolle, A., *Design-dependent loads in topology optimization*, ESAIM Control Optim. Calc. Var. 9 (2003), 19–48 (electronic).

[9] Bronsard, L. and Stoth, B. *Volume-preserving mean curvature flow as a limit of a nonlocal Ginzburg-Landau equation*, SIAM J. Math. Anal. 28, no. 4 (1997), 769–807.

[10] Burger, M. and Stainko, R., *Phase-field relaxation of topology optimization with local stress constraints*, SIAM J. Control Optim. 45, no. 4 (2006), 1447–1466 (electronic).

[11] 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.

[12] Deckelnick, K., Dziuk, G. and Elliott C.M., *Computation of geometric partial differential equations and mean curvature flow*, Acta Numerica (2005), 139–232.

[13] Eisen, G., *On the obstacle problem with a volume constraint*, Manuscripta Math. 43, no. 1 (1983), 73–83.

[14] 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).

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

[16] Evans, L.C., *Partial differential equations*, Graduate Studies in Mathematics 19. American Mathematical Society, Providence, RI (1998).

[17] Garcke, H., Ito, K., and Kohsaka, Y., *Linearized stability analysis of stationary solutions for surface diffusion with boundary conditions*, SIAM J. Math. Anal. 36, no 6 (2005), 1031–1056.

[18] 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.

[19] Garcke, H., Nestler, B., Stinner, B., and Wendler, F., *Allen-Cahn systems with volume constraints*, M^3AS: Mathematical Models and Methods in Applied Sciences 10 (18) (2008) (to appear)

[20] Hildebrandt, S. and Meier, M., *On variational problems with obstacles and integral constraints for vector-valued functions*, Manuscripta Math. 28, no. 1-3 (1979), 185–206.

[21] 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 (electronic) (2003).

[22] Italo Capuzzo, D., 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.

[23] 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.

[24] Kikuchi, F., Nakazato, K., and Ushijima, T., *Finite element approximation of a nonlinear eigenvalue problem related to MHD equilibria*, Japan J. Appl. Math. 1, no. 2 (1984), 369–403.

[25] Kinderlehrer, D. and Stampacchia, G., *An introduction to variational inequalities and their applications*, Pure and Applied Mathematics 88. Academic Press, Inc., New York-London (1980), xx+313 pp.

[26] Nochetto, R. H., Paolini, M. and Verdi, C. *A dynamic mesh algorithm for curvature dependent evolving interfaces*, J. Comput. Phys. 123 (1996), no. 2, 296–310.

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

[28] 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.

[29] Taylor, J. and Cahn, J.W., *Linking anisotropic sharp and diffuse surface motion laws via gradient flows*, J. Statist. Phys. 77 (1-2) (1994), 183–197.

[30] F. Wendler, J. K. Becker, B. Nestler, and P. Bons, *Phase-field simulations of partial melts in geological materials*, Computer & Geosciences, (2009)

[31] Zhou, S. and Wang, M.Y., *Multimaterial structural topology optimization with a generalized Cahn-Hilliard model of multiphase transition*, Struct. Multidisc. Optim. 33 (2007), 89–111.