

NONSMOOTH SCHUR–NEWTON METHODS FOR VECTOR-VALUED CAHN–HILLIARD EQUATIONS

CARSTEN GRÄSER, RALF KORNUBER, AND ULI SACK

ABSTRACT. We introduce nonsmooth Schur–Newton methods for the solution of the nonlinear discrete saddle-point problems arising from discretized vector-valued Cahn–Hilliard equations with logarithmic and obstacle potentials. The discrete problems are obtained by semi-implicit discretization in time and a first order finite element discretization in space. We incorporate the linear constraints that enforce solutions to stay on the Gibbs simplex using Lagrangian multipliers and prove existence of these multipliers under the assumption of a non-trivial initial condition for the order parameters.

1. VECTOR-VALUED CAHN–HILLIARD EQUATIONS

We consider phase separation in isothermal, multi-phase systems on a polygonal (polyhedral) domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$. The concentrations of the different phases $i = 1, \dots, N$ at $(x, t) \in \Omega \times [0, T_0]$, $T_0 > 0$, are represented by the components $u_i(x, t)$ of the order parameter $u = (u_1, \dots, u_N)^T$. The order parameter satisfies the constraints

$$u(x, t) \in G = \{v \in \mathbb{R}^N \mid v_i \geq 0, \sum_{i=1}^N v_i = 1\} \quad \forall (x, t) \in \Omega \times [0, T_0],$$

because concentrations are non-negative and add up to unity. The closed convex set $G \subset \mathbb{R}^N$ is often called *Gibbs simplex*. We assume that the Ginzburg–Landau total free energy of our system takes the form

$$(1.1) \quad \mathcal{E}(u) = \int_{\Omega} \frac{\varepsilon}{2} \sum_{i=1}^N |\nabla u_i|^2 + \frac{1}{\varepsilon} \Psi(u) \, dx$$

with fixed interface parameter $\varepsilon > 0$. While the quadratic interfacial energy is penalizing steep gradients the free energy Ψ gives rise to phase separation. We concentrate on a multi-phase version of the well-known logarithmic free energy [3, 4]. More precisely, $\Psi = \Psi_{\theta}$ is given by

$$(1.2) \quad \Psi_{\theta}(u) = \Phi_{\theta}(u) + \Psi_0(u), \quad u \in G,$$

with the convex function

$$(1.3) \quad \Phi_{\theta}(u) = \begin{cases} \sum_{i=1}^N \theta u_i \ln(u_i) + \chi_{[0, \infty)}(u_i), & \text{for } \theta > 0, \\ \sum_{i=1}^N \chi_{[0, \infty)}(u_i), & \text{for } \theta = 0 \end{cases} \quad u \in G,$$

and the quadratic term

$$\Psi_0(u) = \theta_c \frac{N}{2} u \cdot Cu = \theta_c \frac{N}{2} \sum_{i=1}^N u_i (Cu)_i, \quad (Cu)_i = \sum_{j=1}^N c_{ij} u_j$$

induced by a symmetric interaction matrix $C = (c_{ij})_{i,j=1}^N$ (cf. De Fontaine [9]). Here, θ , θ_c are denoting absolute and critical temperature, respectively. Note that we also incorporate the obstacle potential Ψ_0 as limiting case for $\theta = 0$ using the indicator functional $\chi_{[0,\infty)}$ of $[0, \infty) = \mathbb{R}_0^+$. We also use this for $\theta > 0$ to naturally extend Φ_θ and Ψ_θ to the whole real line. In principle we could also use $\chi_{[0,r]}$ for any $r \geq 1$ instead since $u_i \leq 1$ is ensured by the sum constraint.

For $\theta < \theta_c$, we assume that Ψ_θ has exactly N distinct local minima on G corresponding to almost pure phases $i = 1, \dots, N$. For example, this is true for the choice

$$(1.4) \quad C = (1 - \delta_{ij})_{i,j=1}^N \quad (\text{Kronecker-}\delta)$$

which means that the interaction of all different phases is equal and no self-interaction occurs. For this setting, we obtain the classical obstacle potential (cf. Blowey & Barrett [3])

$$\Psi_0(u) = \theta_c \frac{N}{2} \sum_{i=1}^N u_i (1 - u_i), \quad u \in G.$$

For $N = 2$ the well-known logarithmic free energy

$$\Psi_\theta(\tilde{u}) = \frac{1}{2} \theta \left[(1 + \tilde{u}) \ln\left(\frac{1+\tilde{u}}{2}\right) + (1 - \tilde{u}) \ln\left(\frac{1-\tilde{u}}{2}\right) \right] + \frac{1}{2} \theta_c (1 - \tilde{u}^2)$$

of the scalar order parameter $\tilde{u} := u_2 - u_1$ is recovered in this way. In the shallow quench, i.e. for $\theta \approx \theta_c$, polynomial free energies generalizing the quartic potential $(1 - \tilde{u}^2)^2$ provide good approximations of the logarithmic free energy Ψ_θ (cf. Steinbach et al. [21]). As polynomials are defined everywhere, the non-differentiable constraints $u_i \geq 0$ are usually skipped in this case. On the other hand, in the deep quench limit $\theta \rightarrow 0$ we obviously have $\Psi_\theta(u) \rightarrow \Psi_0(u)$ uniformly on G .

The vector-valued Cahn–Hilliard equation

$$(1.5) \quad \begin{aligned} u_t &= L\Delta w \\ w &= -\varepsilon^2 \Delta u + P\Psi'_\theta(u) \end{aligned}$$

is thermodynamically consistent in the sense that the total free energy \mathcal{E} defined in (1.1) is monotonically decreasing in course of the evolution. The matrix $L \in \mathbb{R}^{N \times N}$ is symmetric, positive semi-definite and has the one-dimensional kernel $\mathbb{R}\mathbf{1} := \mathbb{R}(1, \dots, 1)^T \subset \mathbb{R}^N$. The orthogonal projection $Pv = v - \frac{1}{N}(\mathbf{1} \cdot v)\mathbf{1}$ maps \mathbb{R}^N onto the linear subspace

$$H = \{v \in \mathbb{R}^N \mid \sum_{i=1}^N v_i = 0\}.$$

It accounts for the fact that admissible variations of $u(x, t) \in G$ must be in H .

For given initial condition $u^0(x)$ with $\int_\Omega u^0(x) < \mathbf{1}|\Omega|$ and $\sum_{i=1}^N u_i^0 \equiv 1$ and Neumann boundary conditions for u and w , the Cahn–Hilliard equation (1.5) is conservative in the sense that

$$\int_\Omega u(x, t) dx = \int_\Omega u_0(x) dx, \quad \sum_{i=1}^N u_i(x, t) = 1, \quad \sum_{i=1}^N w_i(x, t) = 0$$

holds for all $t > 0$.

In the deep quench limit $\theta = 0$ the second equation in (1.5) becomes the variational inclusion

$$(1.6) \quad w \in -\varepsilon^2 \Delta u + P\Psi'_0(u) + \partial\chi_G(u).$$

Here $\partial\chi_G$ denotes the subdifferential of the indicator functional χ_G of the Gibbs simplex G .

Concerning existence, uniqueness and sharp interface limits of (1.5), we refer to Elliott & Luckhaus [8], Blowey & Barrett [3] and Bronsard et al. [5].

2. DISCRETIZATION

2.1. Finite element spaces and notation. The following notation for finite elements is essentially copied from [15]. The problems in each time step are discretized in space by piecewise linear conforming finite elements

$$\mathcal{S}(\mathcal{T}) = \{v \in C(\overline{\Omega}) \mid v|_e \text{ is affine } \forall e \in \mathcal{T}\}$$

on a simplicial partition \mathcal{T} of Ω . We assume that $\mathcal{T} = \mathcal{T}_j$ for an underlying grid hierarchy $\mathcal{T}_0, \dots, \mathcal{T}_j$ that is obtained by successive refinement of a conforming intentionally coarse partition \mathcal{T}_0 . Note that, also \mathcal{T}_0 is conforming, we allow \mathcal{T} to have so-called ‘‘hanging nodes’’ on edge mid points. The conforming nodal basis of $\mathcal{S}(\mathcal{T})$ is denoted by λ_p , $p \in \mathcal{N}(\mathcal{T})$, where $\mathcal{N}(\mathcal{T})$ is the set of non-hanging nodes in \mathcal{T} . For a precise definition of hanging nodes and the conforming nodal basis we refer to [11]. The conforming space of vector-valued linear finite element functions with N components is denoted by

$$\mathcal{S}^N(\mathcal{T}) = \{v \in C(\overline{\Omega})^N \mid v|_e \text{ is affine } \forall e \in \mathcal{T}\}.$$

A basis of this $N|\mathcal{N}(\mathcal{T})|$ -dimensional space is given by

$$\mathcal{B}^N(\mathcal{T}) := \{e^i \lambda_p \mid p \in \mathcal{N}(\mathcal{T}), 1 \leq i \leq N\},$$

where $e^i \in \mathbb{R}^N$ with $(e^i)_j = \delta_{ij}$ denotes the i -th Euclidean basis vector. Besides this we will need the space

$$\mathcal{H}_c(\mathcal{T}) = \{v \in \mathcal{S}^N(\mathcal{T}) \mid v(p) \cdot \mathbf{1} = c \forall p \in \mathcal{N}(\mathcal{T})\}$$

for $c = 0$ and $c = 1$ and the convex set

$$\mathcal{K}(\mathcal{T}) = \{v \in \mathcal{S}^N(\mathcal{T}) \mid v(p) \geq 0 \forall p \in \mathcal{N}(\mathcal{T})\}.$$

Note that this implies

$$\mathcal{G}(\mathcal{T}) := \mathcal{H}_1(\mathcal{T}) \cap \mathcal{K}(\mathcal{T}) = \{v \in \mathcal{S}^N(\mathcal{T}) \mid v(p) \in G \forall p \in \mathcal{N}(\mathcal{T})\}.$$

The lumped L^2 -product in $\mathcal{S}^N(\mathcal{T})$ is defined by

$$(u, v)^{\mathcal{T}} = \int_{\Omega} I^{\mathcal{T}}(u \cdot v)(x) dx$$

where the dot product is taken point wise and $I^{\mathcal{T}} : C(\overline{\Omega}) \rightarrow \mathcal{S}(\mathcal{T})$ is the linear interpolation operator given by

$$I^{\mathcal{T}}(v) = \sum_{p \in \mathcal{N}(\mathcal{T})} v(p) \lambda_p.$$

The approximate nonsmooth nonlinear functional $\phi_{\mathcal{T}}$,

$$\phi_{\theta}^{\mathcal{T}}(v) = (\Phi_{\theta}(v), \mathbf{1})^{\mathcal{T}}$$

is obtained by replacing exact integration by a quadrature rule based the nodal interpolation operator in $\mathcal{S}(\mathcal{T})$.

2.2. Semi-implicit time discretization. To simplify the notation, we will not use indices for different time steps but denote solutions from the current time step by u, w and those from the previous one by $u^{\text{old}}, w^{\text{old}}$. Furthermore \mathcal{T} denotes the grid in the current time step that might in general differ from grid in the previous step.

For the logarithmic potential with $\theta > 0$ and a fixed spatial grid Blowey et al. [4] analyzed a fully discrete version using a fully implicit time discretization (FI).

(FI) Find $u \in \mathcal{G}(\mathcal{T}), w \in \mathcal{S}^N(\mathcal{T})$ such that

$$\begin{aligned} \varepsilon^2 (\nabla u, \nabla v) + (P\Psi'_\theta(u), v)^\mathcal{T} &= (w, v)^\mathcal{T} & \forall v \in \mathcal{S}^N(\mathcal{T}), \\ (u, v)^\mathcal{T} + \tau (L\nabla w, \nabla v) &= (u^{\text{old}}, v)^\mathcal{T} & \forall v \in \mathcal{S}^N(\mathcal{T}). \end{aligned}$$

Existence and uniqueness is shown by Blowey et al. [4, Theorem 2.4] under the constraint

$$\tau < 4\varepsilon^2 / (N\theta_c^2 \lambda_C^2 \|L\|)$$

where λ_C is the largest positive eigenvalue of C .

In order to avoid such severe stability restrictions on the time step, the expanding linear part of Ψ'_θ is often discretized explicitly (cf. [4] and others). This approach leads to an unconditionally stable semi-implicit scheme of the form:

(SI) Find $u \in \mathcal{G}(\mathcal{T}), w \in \mathcal{S}^N(\mathcal{T})$ such that

$$\begin{aligned} \varepsilon^2 (\nabla u, \nabla v) + (P\Phi'_\theta(u), v)^\mathcal{T} \\ + (\Psi'_0(u^{\text{old}}) + (P - I)\Psi'_0(u), v)^\mathcal{T} &= (w, v)^\mathcal{T} & \forall v \in \mathcal{S}^N(\mathcal{T}), \\ (u, v)^\mathcal{T} + \tau (L\nabla w, \nabla v) &= (u^{\text{old}}, v)^\mathcal{T} & \forall v \in \mathcal{S}^N(\mathcal{T}). \end{aligned}$$

For convergence results and error estimates of the fully implicit version (FI), we refer to Blowey et al. [4] and Barrett and Blowey [2] who also treat the obstacle potential [3] and concentration-dependent mobilities [1].

While it is easy to see that (FI) satisfies $w \in \mathcal{H}_0(\mathcal{T})$ this is no longer true for (SI) where we get

$$\mathbf{1} \cdot w = \mathbf{1} \cdot (\Psi'_0(u^{\text{old}}) - \Psi'_0(u)).$$

Both formulations do not allow for a direct generalization to the deep quench limit $\theta = 0$ due to the use of the projected derivative $P\Phi'_\theta(\cdot)$. In the following we will derive semi-implicit discretizations that also incorporate $\theta = 0$. Corresponding fully implicit discretizations can be obtained analogously.

To this end we test the first equation in (SI) with $v \in \mathcal{H}_0$ only to obtain

$$\varepsilon^2 (\nabla u, \nabla v) + (\Phi'_\theta(u), v)^\mathcal{T} + (\Psi'_0(u^{\text{old}}), v)^\mathcal{T} = (Pw, v)^\mathcal{T} \quad \forall v \in \mathcal{H}_0(\mathcal{T}).$$

It is easy to see that this equation complemented by the second one in (SI) already guarantees uniqueness of u . Since only Pw but not $(I - P)w$ is determined uniquely

by these two equations we introduce $w_0 = Pw \in \mathcal{H}_0(\mathcal{T})$ as new variable. Furthermore it is sufficient to test the second equation in (SI) by $v \in \mathcal{H}_0(\mathcal{T})$ only. The resulting system can also be written using a variational inequality leading to

$$\begin{aligned} \varepsilon^2 (\nabla u, \nabla(v - u)) + \phi_\theta^\mathcal{T}(v) - \phi_\theta^\mathcal{T}(u) \\ - (w_0, v - u)^\mathcal{T} &\geq - (\Psi'_0(u^{\text{old}}), v - u)^\mathcal{T} \quad \forall v \in \mathcal{H}_1(\mathcal{T}), \\ (u, v)^\mathcal{T} + \tau (L\nabla w_0, \nabla v) &= (u^{\text{old}}, v)^\mathcal{T} \quad \forall v \in \mathcal{H}_0(\mathcal{T}). \end{aligned}$$

This formulation has the advantage that it does naturally include the deep quench limit $\theta = 0$. For $\theta > 0$ it is equivalent to (SI) in the sense that for any solution (u, w) of (SI) a solution of this system is given by (u, Pw) .

Since solutions vary strongly in space we are interested in different adaptive spatial discretizations in each time step. Because the usage of lumped L^2 products would either destroy symmetry or mass conservation if subsequent spatial grids differ we will not use lumping besides the approximation of the convex nonlinearity. For a detailed discussion of the selection of discrete L^2 products with adaptive grids we refer to [11].

(SI2) Find $u \in \mathcal{H}_1(\mathcal{T})$, $w_0 \in \mathcal{H}_0(\mathcal{T})$ such that

$$\begin{aligned} \varepsilon^2 (\nabla u, \nabla(v - u)) + \phi_\theta^\mathcal{T}(v) - \phi_\theta^\mathcal{T}(u) \\ - (w_0, v - u) &\geq - (\Psi'_0(u^{\text{old}}), v - u) \quad \forall v \in \mathcal{H}_1(\mathcal{T}), \\ - (u, v) - \tau (L\nabla w_0, \nabla v) &= - (u^{\text{old}}, v) \quad \forall v \in \mathcal{H}_0(\mathcal{T}). \end{aligned}$$

Here we have replaced the nodal convex constraints $u \in \mathcal{G}(\mathcal{T})$ by the nodal linear constraints $u \in \mathcal{H}_1(\mathcal{T})$. Note that this does not change solutions because $u \in \mathcal{K}(\mathcal{T})$ is already enforced by the characteristic functional in $\phi_\theta^\mathcal{T}$.

If we want to apply a nonlinear saddle point solver to (SI2) we would still have to deal with the nodal linear constraints directly. In the following we introduce Lagrangian multipliers for those constraints in order to avoid this.

First we note that the local linear constraints $u(p) \cdot \mathbf{1}$ can be expressed equivalently in variational form by

$$(u, \mathbf{1}v) = (u \cdot \mathbf{1}, v) = (1, v) \quad \forall v \in \mathcal{S}(\mathcal{T}).$$

Introducing the Lagrangian multiplier $\eta \in \mathcal{S}(\mathcal{T})$ for this constraint we end up with:

(SI3) Find $u \in \mathcal{S}^N(\mathcal{T})$, $w_0 \in \mathcal{H}_0(\mathcal{T})$, $\eta \in \mathcal{S}(\mathcal{T})$ such that

$$\begin{aligned} \varepsilon^2 (\nabla u, \nabla(v - u)) + \phi_\theta^\mathcal{T}(v) - \phi_\theta^\mathcal{T}(u) \\ - (w_0 + \mathbf{1}\eta, v - u) &\geq - (\Psi'_0(u^{\text{old}}), v - u) \quad \forall v \in \mathcal{S}^N(\mathcal{T}), \\ - (u, v) - \tau (L\nabla w_0, \nabla v) &= - (u^{\text{old}}, v) \quad \forall v \in \mathcal{H}_0(\mathcal{T}) \\ - (u, \mathbf{1}v) &= - (u^{\text{old}}, \mathbf{1}v) \quad \forall v \in \mathcal{S}(\mathcal{T}). \end{aligned}$$

Setting $\tilde{w} = w_0 + \mathbf{1}\eta$ and adding the last two equations and using $L\mathbf{1} = 0$ this turns out to be the same as:

(SI4) Find $u \in \mathcal{S}^N(\mathcal{T})$, $\tilde{w} \in \mathcal{S}^N(\mathcal{T})$ such that

$$\begin{aligned} \varepsilon^2 (\nabla u, \nabla(v - u)) + \phi_\theta^{\mathcal{T}}(v) - \phi_\theta^{\mathcal{T}}(u) \\ - (\tilde{w}, v - u) &\geq - (\Psi'_0(u^{\text{old}}), v - u) \quad \forall v \in \mathcal{S}^N(\mathcal{T}), \\ - (u, v) - \tau (L\nabla\tilde{w}, \nabla v) &= - (u^{\text{old}}, v) \quad \forall v \in \mathcal{S}^N(\mathcal{T}). \end{aligned}$$

It is straight forward to prove that (SI2) and (SI4) are equivalent in the sense that if (u^1, w^1) and (u^2, w^2) are solutions of (SI2) and (SI4), respectively, then we have $u^1 = u^2$ and $Pw^1 = Pw^2$. It is easy to see that (SI4) is a saddle point problem where the primal operator $\varepsilon^1 (\nabla \cdot, \nabla \cdot)$ is only coercive on the space that satisfies the second variational equation. However, we can avoid this using the technique used in [10, 11, 14] for the two component case. As in the continuous case we observe the mass conservation

$$\int_{\Omega} u_i(x) dx = \int_{\Omega} u_i^{\text{old}}(x) dx.$$

Hence we can add the terms

$$\varepsilon^2 \int_{\Omega} u \cdot \int_{\Omega} (v - u) = \varepsilon^2 \int_{\Omega} u^{\text{old}} \cdot \int_{\Omega} (v - u)$$

on both sides of the variational inequality without changing the solution. Adding this bilinear integral operator leading to:

(SI5) Find $u \in \mathcal{S}^N(\mathcal{T})$, $\tilde{w} \in \mathcal{S}^N(\mathcal{T})$ such that

$$\begin{aligned} \varepsilon^2 (\nabla u, \nabla(v - u)) + \varepsilon^2 \int_{\Omega} u \cdot \int_{\Omega} (v - u) + \phi_\theta^{\mathcal{T}}(v) - \phi_\theta^{\mathcal{T}}(u) \\ - (\tilde{w}, v - u) &\geq \varepsilon^2 \int_{\Omega} u^{\text{old}} \cdot \int_{\Omega} (v - u) - (\Psi'_0(u^{\text{old}}), v - u) \quad \forall v \in \mathcal{S}^N(\mathcal{T}), \\ - (u, v) - \tau (L\nabla\tilde{w}, \nabla v) &= - (u^{\text{old}}, v) \quad \forall v \in \mathcal{S}^N(\mathcal{T}). \end{aligned}$$

Note that this reformulation is equivalent to adding the term

$$\frac{\varepsilon^2}{2} \left| \int_{\Omega} (u - u^{\text{old}}) \right|^2$$

to the Lagrangian functional associated with (SI4).

Theorem 2.1. *Assume that $0 < \int_{\Omega} u_i^{\text{old}} < |\Omega|$ holds true for all $1 \leq i \leq N$. Then there exists a solution (u, \tilde{w}) of problem (SI5).*

Proof. First we note that (u, \tilde{w}) is a solution of (SI5) iff it is a saddle point of the Lagrangian functional

$$\begin{aligned} \mathcal{L}(v, z) &:= \mathcal{J}(v) + (u^{\text{old}} - v, z) - \frac{\tau}{2} (L\nabla z, \nabla z), \\ \mathcal{J}(v) &:= \frac{\varepsilon^2}{2} (\nabla v, \nabla v) + \frac{\varepsilon^2}{2} \left| \int_{\Omega} (v - u^{\text{old}}) \right|^2 + \phi_\theta^{\mathcal{T}}(v) + (\Psi'_0(u^{\text{old}}), v) \end{aligned}$$

defined on the set

$$\text{dom}(\mathcal{J}) \times \mathcal{S}^N(\mathcal{T}) = \{v \in \mathcal{S}^N(\mathcal{T}) \mid v \geq 0\} \times \mathcal{S}^N(\mathcal{T}).$$

For such saddle point problems Proposition 2.4 in [7] provides existence if

- (1) \mathcal{L} is finite-valued,
- (2) $\mathcal{L}(\cdot, z)$ is convex, lower semi-continuous,

- (3) $\mathcal{L}(v, \cdot)$ is concave, upper semi-continuous,
- (4) there is a $z_0 \in \mathcal{S}^N(\mathcal{T})$ such that $L(\cdot, z_0)$ is coercive,
- (5) the function $-\inf_{v \in \text{dom}(\mathcal{J})} \mathcal{L}(v, \cdot)$ is coercive.

While (1), (2), (3) are obvious, (4) follows from the Poincaré inequality (see [11]).

To show (5) let $z \in \mathcal{S}^N(\mathcal{T})$ be arbitrary and define $v^{(1)}, v^{(2)} \in \mathcal{S}^N(\mathcal{T})$ by

$$v_i^{(1)} = \frac{1 + \text{sgn}(z_i^{(c)})}{2} = \text{const}, \quad v^{(2)}(p) = \left(1 + \text{sgn}[\mathbf{1} \cdot z(p)]\right) \rho \mathbf{1}$$

where $\rho > 0$ is arbitrary by now. Then $0 \leq v = v^{(1)} + v^{(2)} \leq 2$ and thus

$$\mathcal{L}(v, z) \leq C_1 + (u^{\text{old}} - v, z) - C_2(\nabla Pz, \nabla Pz).$$

In order to show coercivity we decompose z according to

$$z = \underbrace{z - |\Omega|^{-1} \int_{\Omega} z}_{z^{(0)}} + \underbrace{|\Omega|^{-1} \int_{\Omega} z}_{z^{(c)}}.$$

Then we have for

$$\begin{aligned} (u^{\text{old}} - v, z) &= (u^{\text{old}}, Pz^{(0)}) + (u^{\text{old}}, (I - P)z^{(0)}) + (u^{\text{old}} - v^{(1)}, z^{(c)}) - (v^{(2)}, z) \\ &= (u^{\text{old}}, Pz^{(0)}) - (\rho \mathbf{1}, z^{(0)}) + (u^{\text{old}} - v^{(1)}, z^{(c)}) - (v^{(2)}, z) \\ &= (u^{\text{old}}, Pz^{(0)}) + (u^{\text{old}} - \rho \mathbf{1} - v^{(1)}, z^{(c)}) + (\rho \mathbf{1} - v^{(2)}, z). \end{aligned}$$

Using the Poincaré inequality the first term can be estimated by

$$(u^{\text{old}}, Pz^{(0)}) \leq \|u^{\text{old}}\|_0 \|Pz^{(0)}\|_0 \leq C_3 \sqrt{(\nabla Pz^{(0)}, \nabla Pz^{(0)})} = C_3 \sqrt{(\nabla Pz, \nabla Pz)}.$$

Before estimating the second term we fix

$$\rho = \frac{1}{2|\Omega|} \min_i \int_{\Omega} u_i^{\text{old}} > 0$$

and set $\mu_i = |\Omega|^{-1} \int_{\Omega} (u_i^{\text{old}} - \rho)$. Then $0 < \mu_i < 1$, the equivalence of $|\cdot|_1$ and $|\cdot|$ in \mathbb{R}^N , and the fact that P is a projection imply

$$\begin{aligned} (u^{\text{old}} - \rho \mathbf{1} - v, z^{(c)}) &= \sum_{i=1}^N (\mu_i - v_i) \int_{\Omega} z_i = - \sum_{i=1}^N |\mu_i - v_i| \left| \int_{\Omega} z_i \right| \\ &\leq -C_4 \left| \int_{\Omega} z \right|_1 \leq -C_5 \left| \int_{\Omega} z \right|_2 \leq -C_5 \left| \int_{\Omega} Pz \right|_2. \end{aligned}$$

To estimate the third term we note that

$$(\rho \mathbf{1} - v^{(2)}(p)) \cdot z(p) = -\rho \text{sgn}[\mathbf{1} \cdot z(p)] \mathbf{1} \cdot (I - P)z(p) = -\rho |\mathbf{1} \cdot z(p)|.$$

Hence we get

$$(\rho \mathbf{1} - v^{(2)}, z) = -\rho \int_{\Omega} |\mathbf{1} \cdot z| = -\rho \sqrt{N} \int_{\Omega} |(I - P)z|.$$

Finally we get

$$\mathcal{L}(v, z) \leq C_6 - C_5 \left| \int_{\Omega} Pz \right| - C_7(\nabla Pz, \nabla Pz) - C_8 \int_{\Omega} |(I - P)z|$$

Hence the infimum over all v must also be bounded uniformly in $\|z\|$. \square

Remark 2.1. *Due to the mass conservation for each component of u the assumption in Theorem 2.1 does only mean that the initial configuration u^0 does neither consist of one phase only nor does a phase completely vanish. Hence we only exclude trivial initial conditions.*

Theorem 2.2. *Under the assumptions of Theorem 2.1 the order parameter u and $\nabla P\tilde{w}$ of the solution to problem (SI5) are unique.*

Proof. Let u^1, \tilde{w}^1 and u^2, \tilde{w}^2 be two solutions of (SI5). Testing the variational inequality for u^1 with u^2 and vice versa and adding both inequalities we get

$$\varepsilon^2 (\nabla(u^1 - u^2), \nabla(u^1 - u^2)) + \varepsilon^2 \int_{\Omega} (u^1 - u^2) \cdot \int_{\Omega} (u^1 - u^2) \leq (\tilde{w}^1 - \tilde{w}^2, u^1 - u^2).$$

Conversely testing the variational equation for both solutions with $w^1 - w^2$ and adding them gives

$$(\tilde{w}^1 - \tilde{w}^2, u^1 - u^2) = -\tau (L\nabla(\tilde{w}^1 - \tilde{w}^2), \nabla(\tilde{w}^1 - \tilde{w}^2)).$$

Inserting this in the inequality yields the assertion. \square

For uniqueness of \tilde{w} we need an extra assumptions on the solution.

Theorem 2.3. *Additionally to the assumptions of Theorem 2.2 assume that there is a basis*

$$B \subset \{\eta_{i,j} = e^i - e^j \in \mathbb{R}^N \mid 1 \leq i, j \leq N, i \neq j\}$$

of $H = \{x \in \mathbb{R}^N \mid \mathbf{1} \cdot x = 0\}$ such that for each $\eta_{i,j} \in B$ there is a vertex $p \in \mathcal{N}(\mathcal{T})$ with $u_i(p), u_j(p) > 0$. Then the solution (u, \tilde{w}) of (SI5) is unique.

Proof. For simplicity we denote the primal bilinear form in the variational inequality by a and the right hand side by f .

Let \tilde{w}^1, \tilde{w}^2 to solutions. First we show $P\tilde{w}^1 = P\tilde{w}^2$. By Theorem 2.2 we already know $\nabla P(\tilde{w}^1 - \tilde{w}^2) = 0$. Hence $P(\tilde{w}^1 - \tilde{w}^2) = c$ for a fixed constant vector $c \in H$. Now let $\eta_{i,j} \in B$ and $p \in \mathcal{N}(\mathcal{T})$ such that $u_i(p), u_j(p) > 0$. Then $v_{\pm} = u \pm \delta \eta_{i,j} \lambda_p \geq 0$ holds true for $\delta > 0$ small enough. Testing the variational inequality for \tilde{w}^1 with v_+ and for \tilde{w}^2 with v_- , adding the results, and deviding by δ yields

$$\frac{\phi_{\theta}^T(u + \delta \eta_{i,j} \lambda_p) - \phi_{\theta}^T(u)}{\delta} - \frac{\phi_{\theta}^T(u) - \phi_{\theta}^T(u - \delta \eta_{i,j} \lambda_p)}{\delta} - (\tilde{w}^1 - \tilde{w}^2, \eta_{i,j} \lambda_p) \geq 0.$$

Due to $u_i(p), u_j(p) > 0$ the function $r \mapsto \phi_{\theta}^T(u + \delta \eta_{i,j} r p)$ is differentiable. Thus we can take the limit $\delta \rightarrow 0$ to get

$$-(\tilde{w}^1 - \tilde{w}^2, \eta_{i,j} \lambda_p) \geq 0.$$

Exchanging the role of \tilde{w}^1 and \tilde{w}^2 we get.

$$0 = (\tilde{w}^1 - \tilde{w}^2, \eta_{i,j} \lambda_p) = (P(\tilde{w}^1 - \tilde{w}^2), \eta_{i,j} \lambda_p) = c \cdot \eta_{i,j} \int_{\Omega} \lambda_p = c \cdot \eta_{i,j}$$

Since this is true for all basis vectors $\eta_{i,j}$ of H we have $c = 0$ and thus $P\tilde{w}^1 = P\tilde{w}^2$.

Now $\tilde{w}^1 - \tilde{w}^2$ must satisfy $\tilde{w}^1 - \tilde{w}^2 = (I - P)(\tilde{w}^1 - \tilde{w}^2) = \mathbf{1} \cdot \bar{w}$ for some scalar function $\bar{w} \in \mathcal{S}^N(\mathcal{T})$. To show $\bar{w} = 0$ let $p \in \mathcal{N}(\mathcal{T})$ and select $1 \leq i \leq N$ such that $u_i(p) > 0$. Then $v_{\pm} = u \pm \delta e^i \lambda_p \in \mathcal{S}^N(\mathcal{T})$ satisfies $v_{\pm} \geq 0$ for $\delta > 0$ small enough. Proceeding as above we get

$$0 = (\tilde{w}^1 - \tilde{w}^2, e^i \lambda_p) = (\mathbf{1} \cdot \bar{w}, e^i \lambda_p) = (\bar{w}, \lambda_p).$$

Since we can do this for all $p \in \mathcal{N}(\mathcal{T})$ we have $\bar{w} = 0$. \square

Remark 2.2. *The extra assumption on u in Theorem 2.3 does essentially mean that for each phase has an interface to another phase.*

2.3. Algebraic formulation. Now we rewrite the discrete problems (SI4) and (SI5) in algebraic form to simplify the presentation of the nonlinear algebraic in the next section. To this end we assume an enumeration

$$\mathcal{N}(\mathcal{T}) = \{p_1, \dots, p_m\}$$

with $m := |\mathcal{N}(\mathcal{T})|$ and define the bijective index map π by

$$\pi : \{1, \dots, N\} \times \mathcal{N}(\mathcal{T}) \rightarrow \{1, \dots, n\}, \quad \pi(i, p_k) = i + N(k-1)$$

where $n := mN$. Then the basis of $\mathcal{S}(\mathcal{T})^N$ takes the form $\mathcal{B}^N(\mathcal{T}) = \{\lambda^1 \dots \lambda^n\}$ with $\lambda^{\pi(i,p)} = e^i \lambda_p$. Using these indices we define the matrices $\tilde{A}, B, C \in \mathbb{R}^{n,n}$ and the vectors $\tilde{f}, g \in \mathbb{R}^n$ by

$$\begin{aligned} \tilde{A}_{i,j} &:= \varepsilon^2 (\nabla \lambda^j, \nabla \lambda^i), & \tilde{f}_i &:= -(\Psi'_0(u^{\text{old}}), \lambda^i), \\ B_{i,j} &:= -(\lambda^j, \lambda^i), & g_i &:= -(u^{\text{old}}, \lambda^i), \\ C_{i,j} &:= \tau (L \nabla \lambda^j, \nabla \lambda^i), \end{aligned}$$

for all $i, j = 1, \dots, n$. Furthermore we define the algebraic representation $\varphi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ of the nonlinearity ϕ_θ by

$$\varphi(V) = \phi_\theta \left(\sum_{i=1}^n V_i \lambda^i \right) = \sum_{i=1}^n \varphi_i(V_i), \quad \varphi_i(\xi) := (\lambda^i, \mathbf{1})^T \Phi_\theta(\xi).$$

Then (SI4) takes the form

$$\begin{aligned} \tilde{A}U \cdot (V - U^*) + \varphi(U^*) - \varphi(V) + BU \cdot (V - U^*) &\geq \tilde{f} \cdot (V - U^*), \quad \forall V \in \mathbb{R}^n, \\ BU^* - CW^* &= g \end{aligned}$$

where U^*, W^* are the coefficient vectors of u, w , respectively. If we also write the variational inequality in operator form using the subdifferential of φ we get the following algebraic problem:

(SP4) Find $U^*, W^* \in \mathbb{R}^n$ such that

$$\begin{pmatrix} \tilde{A} + \partial\varphi & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} U^* \\ W^* \end{pmatrix} \ni \begin{pmatrix} \tilde{f} \\ g \end{pmatrix}.$$

Analogously to (SI4) the matrix \tilde{A} is only positive semi-definite. In order to write (SI5) algebraically we also define the matrix $M \in \mathbb{R}^{N,n}$ by

$$M_{i,j} = \varepsilon \left(\int_{\Omega} \lambda^j \right)_i.$$

Then we have

$$MV = \varepsilon \int_{\Omega} \sum_{i=1}^n V_i \lambda^i, \quad M^T MU \cdot V = \varepsilon^2 \int_{\Omega} \sum_{i=1}^n U_i \lambda^i \cdot \int_{\Omega} \sum_{i=1}^n V_i \lambda^i.$$

Hence we can represent the primal bilinear form and the right hand side in (SI5) by the matrix and the vector

$$A := \tilde{A} + M^T M, \quad f := \tilde{f} + M^T \left(\varepsilon \int_{\Omega} u^{\text{old}} \right)$$

leading to the algebraic problem:

(SP5) Find $U^*, W^* \in \mathbb{R}^n$ such that

$$\begin{pmatrix} A + \partial\varphi & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} U^* \\ W^* \end{pmatrix} \ni \begin{pmatrix} f \\ g \end{pmatrix}.$$

3. NONSMOOTH SCHUR–NEWTON METHODS

In this section we consider the efficient algebraic solution of the nonsmooth nonlinear saddle point problem (SP5).

For saddle point problems resulting from time and space discretization of a binary Cahn–Hilliard equation with obstacle potential the Schur nonsmooth Newton method was introduced in [14]. Numerical results show that the method behaves mesh independent and outperforms former approaches. In [17, 12] the algorithm was extended to more general nonsmooth nonlinearities and applied to a binary Cahn–Hilliard equation with logarithmic potential.

In the following we will present the method and discuss how its application to multi-component Cahn–Hilliard equations with obstacle and logarithmic potential.

3.1. Nonlinear Schur complement and unconstrained minimization. The starting point for the development of the Schur nonsmooth Newton method is a dual unconstrained minimization problem.

First we note that (SP5) is indeed a saddle point problem whose solutions satisfy

$$\mathcal{L}(U^*, \mu) \leq \mathcal{L}(U^*, W^*) \leq \mathcal{L}(V, W^*) \quad \forall \mu, V \in \mathbb{R}^n$$

for the associated Lagrangian functional given by

$$\mathcal{L}(U, W) = \frac{1}{2}AU \cdot U - f \cdot U + \varphi(U) + (BU - g) \cdot W - \frac{1}{2}CW \cdot W.$$

In order to derive the dual problem we eliminate U to get the equation $H(W^*) = 0$ for the nonsmooth nonlinear Schur complement

$$(3.1) \quad H(W) := -B(A + \partial\varphi)^{-1}(f - B^T W) + CW + g.$$

Lemma 3.1. *The nonlinear saddle point problem (SP5) is equivalent to $H(W^*) = 0$ in the sense that U^*, W^* is a solution of (SP5) iff $H(W^*) = 0$ and $U^* = (A + \partial\varphi)^{-1}(f - B^T W^*)$.*

□

Proposition 3.1. *The operator H has the following properties:*

- (1) H is single valued and Lipschitz continuous.
- (2) H is a monotone operator.
- (3) There is a Fréchet-differentiable convex functional $h : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $H = \nabla h$.

Proof. (1) By the Poincaré inequality the primal bilinear form in (SI5) is coercive and thus A is symmetric and positive definite. Since φ is convex, proper, lower semi-continuous the operator $A + \partial\varphi$ has a single valued Lipschitz continuous inverse with Lipschitz constant given by the reciprocal of the coercivity constant of A .

(2) Monotonicity of H follows from symmetry and positive semi-definiteness of C and monotonicity of $\partial\varphi$.

(3) Using Corollary 5.2 in [7, p. 22] it can be seen, that the subdifferential of

$$h(W) = - \inf_{V \in \mathbb{R}^n} \mathcal{L}(V, W) = -\mathcal{L}((A + \partial\varphi)^{-1}(f - B^T W), W)$$

is given by $\partial h(W) = H(W)$. Single valuedness and continuity of H imply that it is actually a Fréchet-derivative. \square

A direct consequence of Proposition 3.1 is that (SP5) is equivalent to minimizing h over \mathbb{R}^n . This equivalence to an unconstrained convex minimization problem for an LC^1 function is the basis for the development of the iterative solver which is by construction a descent method for h .

3.2. Newton like gradient-Related Methods. We now give a short summary of gradient-related descent methods for h skipping all proofs. For a detailed presentation we refer to text books like, e.g., [18]. Adapted proofs that exactly fit into the setting considered here can be found in [11].

We consider iterative methods of the form

$$(3.2) \quad W^{\nu+1} = W^\nu + \rho_\nu D^\nu.$$

In the simplest case D^ν can be chosen to be $D^\nu = -\nabla h(W^\nu)$ leading to the classical gradient method. This method is in general very slow for the problems we consider. Thus we will look at so called gradient-related directions instead. More specifically we skip the definition of such directions and look at Newton-like directions of the form

$$D^\nu = -S_\nu^{-1} \nabla h(W^\nu)$$

for a sequence of symmetric positive definite preconditioners S_ν . Such directions turn out to be gradient-related if the sequence S_ν is bounden from above and below.

While it is essential that the directions allow for sufficient descent of the functional h the step sizes ρ_ν have to guarantee that this descent is actually realized. This is guaranteed by so called efficient step sizes that satisfy

$$h(W^\nu + \rho_\nu D^\nu) \leq h(W^\nu) - c_S \left(\frac{\nabla h(W^\nu) \cdot D^\nu}{|D^\nu|} \right)^2$$

for a constant $c_S > 0$ independent of ν .

Theorem 3.1. *Assume that the sublevel set $\{W \in \mathbb{R}^n \mid h(W) \leq h(W^0)\}$ is compact, that there are constants $\gamma, \Gamma > 0$ such that*

$$\gamma |v|^2 \leq S_\nu V \cdot V \leq \Gamma |V|^2,$$

and that the step sizes ρ_ν are efficient. Then W^ν converges to the minimizer W^ .*

Proof. See, e.g., Theorem 5.2 and Theorem 5.7 in [11]. \square

The following lemma guarantees that Theorem 3.1 can be applied to the problem at hand.

Lemma 3.2. *If h is convex, continuous, and has a unique minimizer, then the sublevel set $\{W \in \mathbb{R}^n \mid h(W) \leq h(W^0)\}$ are compact for all $W^0 \in \mathbb{R}^n$.*

Proof. See Lemma 5.1 in [11]. \square

While the selection of directions essentially determines the properties of the algorithm there are a lot of generic strategies to select efficient step sizes, for example the widely known Armijo rule. As an alternative we propose a strategy that approximates the minimizer of h along $W^\nu + \rho D^\nu$.

Proposition 3.2. *Assume that $D^\nu \cdot \nabla h(W^\nu) < 0$ and that*

$$\nabla h(W^\nu + \rho_\nu D^\nu) \cdot D^\nu \in [\alpha \nabla h(W^\nu) \cdot D^\nu, 0]$$

for all ν and some fixed $\alpha \in [0, 1)$. Then the step sizes ρ_ν are efficient.

Proof. See Proposition 5.4 in [11]. □

While the exact minimizers obtained for $\alpha = 0$ are in general not available, such step sizes can be computed by a simple bisection for any fixed $\alpha \in (0, 1)$.

Remark 3.1. *The above result remains valid for certain inexact variants of the algorithm. On the one hand the directions D^ν can be replaced by inexact versions \tilde{D}^ν . In this case convergence is still guaranteed as long as these are still descent directions and the relative error decreases, i.e.,*

$$\tilde{D}^\nu \cdot \nabla h(W^\nu) < 0, \quad |D^\nu - \tilde{D}^\nu|/|\tilde{D}^\nu| \rightarrow 0.$$

On the other hand the step sizes $\rho_\nu = 1$ can be used for all steps where the contraction $|D^\nu| \leq \sigma |D^{\nu-1}|$ holds true for some fixed $\sigma < 1$. For a detailed analysis of these inexact versions we refer to [11].

3.3. Nonsmooth Schur–Newton directions. For a sufficiently smooth functional h the most efficient choice of preconditioners would clearly be $S_\nu = \nabla^2 h(W^\nu)$ leading to the classical Newton method.

Since h is only once differentiable and its derivative $\nabla h = H$ is only Lipschitz continuous but not differentiable itself this is not applicable here. For problems of this type one could in principle apply a semi-smooth Newton method by selecting $S_\nu \in \partial_C H(W^\nu)$ where ∂_C denoted the generalized Jacobian in the sense of Clarke [6]. However, it is in general very hard to compute elements of this set, since the chain rule does not hold for ∂_C in general.

As a remedy preconditioner that are in general not contained in $\partial_C H(W^\nu)$ are constructed by assuming a chain rule in [11]. We follow the same approach here omitting the details construction. Instead we will only present the main ideas and the resulting preconditioner for the present case.

In principle we want to construct linearizations of $\nabla h = H$ at some W by

$$(3.3) \quad BQB^T + C$$

where Q is a linearization of the nonsmooth operator $(A + \partial\varphi)^{-1}$ at $f - B^T W$. The main idea is that $(A + \partial\varphi)^{-1}$ is smooth in certain components and essentially constant in all the others. To specify this we introduce the inactive set

$$\mathcal{I}(U) := \{i \mid \partial\varphi_i(U_i) \text{ is single valued}\}$$

of all components where $A + \partial\varphi$ is smooth. For the present case this reduces to

$$\mathcal{I}(U) = \{i \mid U_i > 0\}.$$

It turns out that for all $i \in \mathcal{I}((A + \partial\varphi)^{-1}Y)$ the i -th component of the operator is differentiable while it is constant (at least in one direction) for all the others. Using this idea we can define a linearization of $(A + \partial\varphi)^{-1}$ at some Y with $X = (A + \partial\varphi)^{-1}Y$ by

$$(A + \varphi''(X))_{\mathcal{I}(X)}^+$$

where $\varphi''(X)$ denotes the diagonal matrix with the diagonal entries $\varphi''_i(X_i)$. Here we used the notation R^+ for the Moore–Penrose pseudoinverse of R and $R_{\mathcal{I}}$ for the truncated matrix

$$(R_{\mathcal{I}})_{i,j} = \begin{cases} R_{i,j} & \text{if } i, j \in \mathcal{I}, \\ 0 & \text{else.} \end{cases}$$

Note that we only need to evaluate φ''_i for components where it is differentiable. Plugging this into (3.3) we get the linearization

$$\partial H(W) := B(A + \varphi''(U))_{\mathcal{I}(U)}^+ B^T + C$$

of H at W where $U = (A + \partial\varphi)^{-1}(f - B^T W)$.

Unfortunately even without truncation the matrices $\partial H(W)$ are in general not uniformly bounded from below for the logarithmic potential. This is because the diagonal entries $\varphi''_i(U_i)$ might be arbitrarily large if U_i is close to zero. To avoid this problem we will also truncate indices i where $\Phi''_{\theta}(U_i) \geq c_T$ for some fixed large constant c_T . The latter is equivalent to using the smaller inactive set

$$\bar{\mathcal{I}}(U) := \{i \mid U_i > \delta\}$$

for some small $\delta > 0$ depending on c_T leading to reduced linearizations

$$\bar{\partial} H(W) := B(A + \varphi''(U))_{\bar{\mathcal{I}}(U)}^+ B^T + C.$$

It is furthermore possible that $\bar{\partial} H(W)$ has a nontrivial kernel. In this case we regularize it by adding some scaled symmetric positive definite matrix to it, e.g., a scaled mass matrix βB , ending up with the following preconditioner

$$S_{\nu} = \begin{cases} \bar{\partial} H(W^{\nu}) & \text{if } \bar{\partial} H(W^{\nu}) \text{ is regular,} \\ \bar{\partial} H(W^{\nu}) + \beta B & \text{else.} \end{cases}$$

It is obvious that these matrices are symmetric positive definite. But even more they are sufficient to apply the convergence result in Theorem 3.1.

Theorem 3.2. *There preconditioners S_{ν} satisfy the assumption of Theorem 3.1.*

Proof. See Theorem 5.7 in [11]. □

Corollary 3.1. *If we use the directions $D^{\nu} = -S_{\nu} \nabla h(W^{\nu})$ and efficient step sizes, then algorithm (3.2) converges to the minimizer W^* of h for arbitrary initial iterate. The same is true for the inexact versions of the algorithm mentioned in Remark 3.1.* □

Remark 3.2. *Although the convergence theory is based on the interpretation as gradient-like method it was observed that the algorithm converges superlinearly in general for binary Cahn–Hilliard equations [14, 11].*

In contrast to the analysis used here, classical convergence analysis for semi-smooth Newton methods leads to local superlinear convergence results [19] with unknown region of convergence. However, this would not give any further insight since the region of convergence could be contained in the region with exact inactive set. In this case the algorithm would be anyway a classical Newton method for the logarithmic potential and terminate in one step for the obstacle potential.

Remark 3.3. We stated the convergence theory in terms of the euclidean norm $|\cdot|$ so that the convergence theory is not mesh-independent. While we could also use a different norm there are more finite dimensional arguments in the convergence proof. In contrast to the theory the algorithm behaves mesh-independent for binary Cahn–Hilliard equations in practice [14, 11]. The theoretical justification for this is still an open problem.

3.4. Algorithmic aspects. Although algorithm was derived in terms of the dual unknown W only it can also be written in the primal–dual form

$$\begin{aligned} U^\nu &= (A + \partial\varphi)^{-1}(f - B^T W^\nu) \\ W^{\nu+1} &= w^\nu + \rho_\nu S_\nu^{-1}(B U^\nu + C W^\nu + g). \end{aligned}$$

This reveals that it is essentially a preconditioned Uzawa method. In each iteration there are least two expensive problems to solve.

The first one is the evaluation of $(A + \partial\varphi)^{-1}$ in the computation of U^ν . It is equivalent to solving a convex minimization problem for an energy functional that is essentially the primal convex part of the Lagrangian functional \mathcal{L} . While there is a multitude of methods for the obstacle problems obtained in case of the obstacle potential, only few efficient methods are available for the case of the logarithmic potential. The *truncated nonsmooth Newton method* [13, 16, 11] has shown to perform very well for both cases. It allows to solve these problems with basically the same multigrid efficiency as for linear elliptic problems. It has to be stressed that it is possible to implement such multigrid methods with $O(n)$ complexity per iteration step although the matrix A is not sparse. The reason is that a careful implementation can exploit the fact that A is the sum of a sparse matrix and a full low-rank matrix with a known product representation [10].

The second problem in each step is the evaluation of S_ν . This can be rewritten as a linear saddle point problem. Again there is a multitude of methods for the solution of these problems. Amongst the most efficient ones are multigrid methods with Vanka smoothers [22] that solve local saddle point problems either Gauß–Seidel like successively or Jacobi like in parallel. While there are only theoretical results for the Jacobi version [20] the Gauß–Seidel version performs even better in general. Unfortunately the algorithm does sometimes not converge for truncated saddle point problems as they appear here. In practice this can be avoided by using the multigrid method as preconditioner in a GMRES method. The development of a convergent multigrid algorithm that also works in these cases is still an open problem.

Besides these two problems the use of a step size rule might require several evaluations of h or $H = \nabla h$. Both incorporate the evaluation of $(A + \partial\varphi)^{-1}$ that has been discussed above. Even if these problems can be solved with multigrid efficiency this is better avoided by using the adaptive step size rule mentioned in Remark 3.1.

REFERENCES

- [1] J. W. Barrett and J. F. Blowey. An improved error bound for a finite element approximation of a model for phase separation of a multi-component alloy with a concentration dependent mobility matrix. *Numer. Math.*, 88:255–297, 2001.

- [2] John W. Barrett and James F. Blowey. An error bound for the finite element approximation of a model for phase separation of a multi-component alloy. *IMA J. Numer. Anal.*, 16:257–287, 1996.
- [3] John W. Barrett and James F. Blowey. Finite element approximation of a model for phase separation of a multi-component alloy with non-smooth free energy. *Numer. Math.*, 77:1–34, 1997.
- [4] J.F. Blowey, M.I.M. Copetti, and C.M. Elliott. Numerical analysis of a model for phase separation of a multi-component alloy. *IMA J. Numer. Anal.*, 16(1):111–139, 1996.
- [5] L. Bronsard, H. Garcke, and B. Stoth. A multi-phase mullins-sekerka system: Matched asymptotic expansions and an implicit time discretisation for the geometric evolution problem. *Proc. R. Soc. Edinb., Sect. A, Math.*, 128(3):481–506, 1998.
- [6] F. H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983.
- [7] I. Ekeland and R. Temam. *Convex Analysis*. Studies in Mathematics and its Applications. North-Holland, Amsterdam New York Oxford, 1976.
- [8] C.M. Elliott and S. Luckhaus. A generalised diffusion equation for phase separation of a multicomponent mixture with interfacial free energy. Preprint 887, IMA, University of Minnesota, 1991.
- [9] D. De Fontaine. An analysis of clustering and ordering in multicomponent solid solutions I. Stability criteria. *J. Phys. Chem. Solids*, 33:287–310, 1972.
- [10] C. Gräser. Analysis und Approximation der Cahn-Hilliard Gleichung mit Hindernispotential. Diplomarbeit, Freie Universität Berlin, 2004.
- [11] C. Gräser. *Convex Minimization and Phase Field Models*. PhD thesis, Freie Universität Berlin, 2011.
- [12] C. Gräser. Schur nonsmooth newton methods for nonsmooth saddle point problems. in preparation, 2012.
- [13] C. Gräser and R. Kornhuber. Multigrid methods for obstacle problems. *J. Comp. Math.*, 27(1):1–44, 2009.
- [14] C. Gräser and R. Kornhuber. Nonsmooth Newton methods for set-valued saddle point problems. *SIAM J. Numer. Anal.*, 47(2):1251–1273, 2009.
- [15] C. Gräser, R. Kornhuber, and U. Sack. Time discretizations of anisotropic allen–cahn equations. Preprint, Matheon Berlin, 2011.
- [16] C. Gräser, U. Sack, and O. Sander. Truncated nonsmooth Newton multigrid methods for convex minimization problems. In M. Bercovier, M. Gander, R. Kornhuber, and O. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XVIII*, LNCSE, pages 129–136. Springer, 2009.
- [17] Carsten Gräser and Oliver Sander. Polyhedral Gauss–Seidel converges. *SIOPT*, submitted.
- [18] J.M. Ortega and W.C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York, 1970.
- [19] L. Qi and J. Sun. A nonsmooth version of Newton’s method. *Math. Prog.*, 58:353–367, 1993.
- [20] J. Schöberl and W. Zulehner. On additive Schwarz-type smoothers for saddle point problems. *Numer. Math.*, 95(2):377–399, 2003.
- [21] I. Steinbach, F. Pezolla, B. Nestler, M. Seeßelberg, R. Prieler, G.J. Schmitz, and J.L.L. Rezende. A phase-field concept for multiphase systems. *Physica D*, 94:135–147, 1996.

- [22] S.P. Vanka. Block-implicit multigrid solution of Navier-Stokes equations in primitive variables. *J. Comput. Phys.*, 65:138–158, 1986.

CARSTEN GRÄSER, FREIE UNIVERSITÄT BERLIN, INSTITUT FÜR MATHEMATIK, ARNIMALLEE 6,
D - 14195 BERLIN, GERMANY
E-mail address: graeser@math.fu-berlin.de

PROF. DR. RALF KORNUBER, FREIE UNIVERSITÄT BERLIN, INSTITUT FÜR MATHEMATIK, ARN-
IMALLEE 6, D - 14195 BERLIN, GERMANY
E-mail address: kornhuber@math.fu-berlin.de

ULI SACK, FREIE UNIVERSITÄT BERLIN, INSTITUT FÜR MATHEMATIK, ARNIMALLEE 6, D -
14195 BERLIN, GERMANY
E-mail address: usack@math.fu-berlin.de