MAX-PLANCK-GESELLSCHAFT

**Max Planck Institute Magdeburg
Preprints**

Jessica Bosch   Martin Stoll

Preconditioning for vector-valued Cahn–Hilliard equations

MAX–PLANCK–INSTITUT
FÜR DYNAMIK KOMPLEXER
TECHNISCHER SYSTEME
MAGDEBURG

The solution of vector-valued Cahn–Hilliard systems is of interest in many applications. We discuss strategies for the handling of smooth and nonsmooth potentials as well as for different types of constant mobilities. Concerning the nonsmooth systems, the necessary bound constraints are incorporated via the Moreau–Yosida regularization technique. We develop effective preconditioners for the efficient solution of the linear systems in saddle point form. Numerical results illustrate the efficiency of our approach. In particular, we numerically show mesh and phase independence of the developed preconditioner in the smooth case. The results in the nonsmooth case are also satisfying and the preconditioned version always outperforms the unpreconditioned one.

1 Introduction

The Cahn–Hilliard equation is a partial differential equation of fourth order which is used in materials science [47, 30], image processing [19] or chemistry [59]. It was originally introduced to model phase separation in binary alloys [39, 16] that occurs when the temperature of a homogeneous mixture is rapidly quenched below a critical temperature. In practice, often more than two phases occur, see e.g. [45, 25, 22, 21, 6, 42, 32], and the phase field model has been extended to deal with multi-component systems. A vector-valued order parameter $\mathbf{u} = (u_1, \ldots, u_N)^T \colon \Omega \times (0, T) \to \mathbb{R}^N$ is introduced, where $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$) is a bounded domain, $T > 0$ is an arbitrary but fixed time and N is the number of phases. Each u_i describes the fraction of one phase, i.e. if $u_i = 0$ then phase i is absent in that region and if $u_i = 1$ only phase i is present in that region. Hence

$$\sum_{i=1}^{N} u_i = 1 \tag{1}$$

and $u_i \geq 0$ is required, so that admissible states belong to the Gibbs simplex

$$\mathcal{G}^N := \left\{ \mathbf{v} \in \mathbb{R}^N \;\middle|\; \sum_{i=1}^{N} v_i = 1, \; v_i \geq 0 \text{ for } i = 1, \ldots, N \right\}. \tag{2}$$

We study a diffuse phase transition, i.e. the region between the phases has a certain width b, the so-called interface (phase field model). There is also the limit case $b \downarrow 0$ which gives the sharp interface model [31, 30, 49]. The motion of the interfaces separating N bulk regions can be modeled with the Ginzburg–Landau energy

$$\mathcal{E}(\mathbf{u}) = \int_{\Omega} \frac{\varepsilon^2}{2} \sum_{i=1}^{N} |\nabla u_i|^2 + \psi(\mathbf{u}) \, d\mathbf{x},$$

where $\varepsilon > 0$ is the gradient energy coefficient. The potential function $\psi \colon \mathbb{R}^N \to \mathbb{R}_0^+ \cup \{\infty\}$ gives rise to phase separation. It can be modeled by a smooth free energy, e.g. using multi-well potentials [20] such as

$$\psi(\mathbf{u}) := \frac{1}{4} \sum_{i=1}^{N} u_i^2 (1 - u_i)^2, \tag{3}$$

or by a nonsmooth multi obstacle potential [4]

$$\psi(\mathbf{u}) := \begin{cases} \psi_0(\mathbf{u}) = -\frac{1}{2} \mathbf{u} \cdot A\mathbf{u} & \mathbf{u} \in \mathcal{G}^N, \\ \infty & \text{otherwise}, \end{cases} \tag{4}$$

where the symmetric matrix $A \in \mathbb{R}^{N \times N}$ contains constant interaction parameters A_{ij}. From physical considerations, A must have at least one positive eigenvalue. A typical choice is $A = I - \mathbf{1}\mathbf{1}^T$ with $\mathbf{1} = (1, \ldots, 1)^T$ and the identity matrix $I \in \mathbb{R}^{N \times N}$, which means

that the interaction between all different components is equal and no self-interaction occurs. Other possible potentials are logarithmic ones, see e.g. [3]. This work deals with the two types of potential (3) and (4). The smooth potential (3) is used for shallow temperature quenches. It has the disadvantage that physically non-admissible values $u_i < 0$ or $u_i > 1$ can be attained during the evolution, see section 8.1. The consideration of the deep quench limit, i.e. a very rapid cooling of the mixture, leads to the multi obstacle potential (4). It omits the disadvantage of (3) but leads to a system of variational inequalities. Motivated by the work of Hintermüller et al. [40] as well as our previous studies [10, 9], all of them considering scalar, nonsmooth Cahn–Hilliard systems, we incorporate the bound constraints via the Moreau–Yosida regularization technique and solve the resulting subproblems by a semismooth Newton (SSN) method.

As we show in the course of this paper the solution of a linear system $\mathcal{K}x = b$ with a real nonsymmetric matrix $\mathcal{K}$ is at the heart of this method. The sparse linear systems are usually of very large dimension and in combination with three-dimensional experiments the application of direct solvers such as UMFPACK [18] becomes infeasible. As a result iterative methods have to be employed (see e.g. [36, 52] for introductions to this field). We propose the use of a Krylov subspace solver. The convergence behavior of the iterative scheme typically depends on the conditioning of the problem and the clustering of the eigenvalues. These properties can be enhanced using preconditioning techniques $\mathcal{P}^{-1}\mathcal{K}x = \mathcal{P}^{-1}b$, where $\mathcal{P}$ is an invertible matrix that is easy to invert and resembles $\mathcal{K}$. In this paper, we provide efficient preconditioners $\mathcal{P}$ for the solution of Cahn–Hilliard variational (in-)equalities using an effective Schur complement approximation and (algebraic) multigrid developed for elliptic systems [27, 52, 51].

The paper is organized as follows. In section 2, we derive the vector-valued Cahn–Hilliard systems for the use of the smooth potential (3). Section 3 presents the systems with the nonsmooth potential (4) and their handling with the Moreau–Yosida regularization technique. The problems are discretized in time in section 4. Section 5 shortly introduces the SSN method to solve the regularized subproblems. The linear systems arising from the discretization using finite elements are derived in section 6. In section 7, we analyze the linear systems and propose preconditioning strategies for the saddle point problems. Section 8 illustrates the efficiency of our approach. In section 9, we discuss alternative approaches and finally, section 10 summarizes our findings.

2 Derivation

The evolution of $\mathbf{u}$ is governed by the H^{-1}-gradient of the Ginzburg–Landau energy under the constraint (1), which has to hold everywhere at any time. Using the smooth

potential (3), the vector-valued Cahn–Hilliard equations read

$$\partial_t u_i = (L\Delta \mathbf{w})_i, \tag{5}$$

$$w_i = f(u_i) + \beta(\mathbf{u}) - \varepsilon^2 \Delta u_i, \tag{6}$$

$$\nabla u_i \cdot \mathbf{n} = (L\nabla \mathbf{w})_i \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega, \tag{7}$$

for $i = 1, \ldots, N$. The matrix $L = (L_{ij})_{i,j=1,\ldots,N} \in \mathbb{R}^{N \times N}$ is the mobility matrix and

$$\mathbf{f}(\mathbf{u}) = (f(u_1), \ldots, f(u_N))^T := \left(\frac{\partial \psi}{\partial u_1}, \ldots, \frac{\partial \psi}{\partial u_N} \right)^T = \frac{\partial \psi}{\partial \mathbf{u}},$$

in which $f(u_i) = u_i^3 - \frac{3}{2}u_i^2 + \frac{1}{2}u_i$, and $\beta(\mathbf{u}) := -\frac{1}{N}\sum_{i=1}^N f(u_i)$. In the process, the chemical potentials $\mathbf{w} = (w_1, \ldots, w_N)^T$ result from the variational derivative of the energy $\mathcal{E}$. In doing so, admissible variations $\mathbf{d} = (d_1, \ldots, d_N)^T$ of $\mathbf{u}$ have to fulfill $\sum_{i=1}^N d_i = 0$ in order to ensure (1). This explains the presence of the term $\beta(\mathbf{u})$, see also [44] for a more detailed calculation. Equation (7) contains the natural zero Neumann boundary condition $\nabla u_i \cdot \mathbf{n} = 0$ on $\partial\Omega$ as well as the mass conserving boundary condition $(L\nabla \mathbf{w})_i = 0$ on $\partial\Omega$, $i = 1, \ldots, N$. Since using the latter in (5) yields together with Gauss's theorem $\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u_i \, \mathrm{d}\mathbf{x} = 0$, i.e. the total mass of each phase is conserved. We refer to [22] for a detailed development of the vector-valued Cahn–Hilliard equations.

The coefficients L_{ij} may depend on $\mathbf{u}$ (see e.g. [21]) but this work deals with constant L_{ij}. In order to ensure the constraint (1), a common way in the literature is to assume that L is symmetric and $L\mathbf{1} = \mathbf{0}$, since summing (5) over $i = 1, \ldots, N$ leads then to

$$\frac{\partial}{\partial t}\sum_{i=1}^N u_i = \sum_{i=1}^N \frac{\partial u_i}{\partial t} = \sum_{i=1}^N \nabla \cdot (L\nabla \mathbf{w})_i = \nabla \cdot \sum_{i,j=1}^N L_{ij}\nabla w_j = \nabla \cdot \sum_{j=1}^N \nabla w_j \sum_{i=1}^N L_{ij} = 0.$$

Therefore, (1) is fulfilled during the evolution if $\sum_{i=1}^N u_i = 1$ at time 0. It is further assumed that L is positive semidefinite as differentiating the energy $\mathcal{E}$ gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\mathbf{u}) = -\int_\Omega \sum_{i=1}^N \nabla w_i \cdot (L\nabla \mathbf{w})_i \, \mathrm{d}\mathbf{x} \leq 0,$$

where we have used Green's first identity. Therefore, the total energy is non-increasing in time.

Remark 1. *Although it is explained that L should fulfill L$\mathbf{1}$ = $\mathbf{0}$ (see also e.g. [22, 8]), Lee et al. [43] work with L = I for convenience and obtain at least visually correct results. Therefore, we also allow this case in our work and our numerical solver simplifies.*

This work studies the two cases $L = I$ (used e.g. in [43]) and $L = I - \frac{1}{N}\mathbf{1}\mathbf{1}^T$ (used e.g. in [21, 35]).

3 Moreau–Yosida regularization

If we now use the nonsmooth multi obstacle potential (4) the gradient of $\mathcal{E}$ employs subdifferentials, which results in the following vector-valued Cahn–Hilliard variational inequalities

$$\langle \partial_t u_i, v \rangle + ((L\nabla \mathbf{w})_i, \nabla v) = 0 \quad \forall v \in H^1(\Omega), \tag{8}$$

$$\varepsilon^2 (\nabla u_i, \nabla(v_i - u_i)) - \left(w_i + (A\mathbf{u})_i - \frac{1}{N} \sum_{j=1}^{N} (A\mathbf{u})_j, v_i - u_i \right) \geq 0 \quad \forall \mathbf{v} \in \mathcal{G}^N \cap H^1(\Omega)^N, \tag{9}$$

$$\mathbf{u} \in \mathcal{G}^N \cap H^1(\Omega)^N \quad \text{a.e. in } \Omega, \tag{10}$$

for $i = 1, \dots, N$. Here $(\cdot, \cdot)$ and $\langle \cdot, \cdot \rangle$ stand for the $L^2(\Omega)$-inner product and the duality pairing of $H^1(\Omega)$ and $H^1(\Omega)^*$, respectively. As motivated in [40, 10, 9], we handle the pointwise constraints in (10) with the Moreau–Yosida regularization technique. Instead of the energy functional $\mathcal{E}$ we consider

$$\mathcal{E}_\nu(\mathbf{u}_\nu) = \int_\Omega \frac{\varepsilon^2}{2} \sum_{i=1}^{N} |\nabla u_{\nu,i}|^2 + \psi_0(\mathbf{u}_\nu) + \frac{1}{2\nu} \sum_{i=1}^{N} |\min(0, u_{\nu,i})|^2 \, d\mathbf{x}.$$

where $0 < \nu \ll 1$ denotes the penalty parameter. As in the case of the smooth potential in the section before, we can now derive the nonsmooth vector-valued Cahn–Hilliard systems

$$\partial_t u_{\nu,i} = (L\Delta \mathbf{w}_\nu)_i, \tag{11}$$

$$w_{\nu,i} = \frac{1}{\nu} \min(0, u_{\nu,i}) - (A\mathbf{u}_\nu)_i - \frac{1}{N} \sum_{j=1}^{N} \left[\frac{1}{\nu} \min(0, u_{\nu,j}) - (A\mathbf{u}_\nu)_j \right] - \varepsilon^2 \Delta u_{\nu,i}, \tag{12}$$

$$\nabla u_{\nu,i} \cdot \mathbf{n} = (L\nabla \mathbf{w}_\nu)_i \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega, \tag{13}$$

for $i = 1, \dots, N$. We now want to discretize the problems (5)–(7) and (11)–(13) (in weak formulation) in time and space. This is demonstrated in the following only with the latter nonsmooth system, the smooth system can be handled analogously.

4 Discretization in time

Concerning the time, fully implicit discretizations are the most accurate, see e.g. [7, 15, 10]. Let $\tau > 0$ denote the time step size and $n \in \mathbb{N}$ the time step. We use the backward Euler discretization for the time derivative $\partial_t u_{\nu,i}$, $i = 1, \dots, N$, and treat all the other terms implicitly. Then, for every time step we have to solve the time-discrete systems

$$\left(u_{\nu,i} - u_{\nu,i}^{(n-1)}, v \right) + \tau \left((L\nabla \mathbf{w}_\nu)_i, \nabla v \right) = 0 \quad \forall v \in H^1(\Omega), \tag{14}$$

$$(w_{\nu,i}, v) - \varepsilon^2 (\nabla u_{\nu,i}, \nabla v) + ((A\mathbf{u}_\nu)_i, v) - \frac{1}{\nu}(\min(0, u_{\nu,i}), v)$$

$$+ \frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{\nu}(\min(0, u_{\nu,j}), v) - ((A\mathbf{u}_\nu)_j, v)\right] = 0 \quad \forall v \in H^1(\Omega), \quad (15)$$

$i = 1, \ldots, N$, for a sequence $\nu \to 0$, where we write $\mathbf{u}_\nu^{(n)} = \mathbf{u}_\nu$ and $\mathbf{w}_\nu^{(n)} = \mathbf{w}_\nu$. Now, in order to get the linear systems of equation, we solve (14)–(15) by an SSN method which is motivated in [40, 10, 9] and shortly summarized in the following section. The smooth nonlinear time-discrete Cahn–Hilliard equations for (5)–(7) are solved by the standard Newton method.

5 Semismooth Newton method

For a specified sequence $\nu \to 0$, we solve the system (14)–(15), which can be compactly written as $\mathbf{F}_\nu(\mathbf{u}_\nu, \mathbf{w}_\nu) = \mathbf{0}$, for every ν by an SSN algorithm, see also [41]. Due to the presence of the minimum operator, $\mathbf{F}_\nu$ is not Fréchet-differentiable. However, the minimum operator satisfies the weaker notion of Newton differentiability, see [41, 40].

Definition 1 (Definition 5.1 in [40]). *Let X and Z be Banach spaces, $D \subset X$ an open subset. A mapping $F \colon D \subset X \to Z$ is called Newton-differentiable in $U \subset D$ if there exists a family of mappings $G \colon U \to Z$ such that*

$$\lim_{d \to 0} \frac{\|F(x + d) - F(x) - G(x + d)d\|_Z}{\|d\|_X} = 0 \quad \forall x \in U.$$

The operator G is called a Newton derivative of F on U.

In general, for a Newton-differentiable mapping F with a Newton derivative G, the SSN iteration is given as

$$x^{(k+1)} = x^{(k)} - G\left(x^{(k)}\right)^{-1} F(x^{(k)}), \quad k = 0, 1, \ldots \quad (16)$$

Given a sufficiently close initial guess $x^{(0)}$, [41, Theorem 1.1] shows superlinear convergence of the sequence $\{x^{(k)}\}_{k\in\mathbb{N}}$, generated by (16), to the solution of $F(x) = 0$.

Regarding the scalar two-phase Cahn–Hilliard equation, the Newton differentiability of the arising mapping F_ν as well as the superlinear convergence of the corresponding SSN iteration are proven in [40] for the semi-implicit time-discrete system and extended to the implicit time-discrete system in [10]. Both works are based on the Newton derivative of the minimum mapping $\min(0, \cdot) \colon H^1(\Omega) \to H^1(\Omega)^*$, which is given as

$$G_{\min}(y)(x) = \begin{cases} 1 & \text{if } y(x) \le 0, \\ 0 & \text{if } y(x) > 0, \end{cases}$$

see [41, Proposition 4.1] and [40, Lemma 5.3]. We use this result for the Newton derivative of every minimum operator in our mapping $\mathbf{F}_\nu$ given in (14)–(15). Regarding the i-th component, this leads us to the Newton derivative

$$
\left\langle G_\nu^{(i)}(\mathbf{u}, \mathbf{w})(\delta\mathbf{u}, \delta\mathbf{w}), (\phi, \psi) \right\rangle
$$
$$
= \left(\begin{array}{c} \tau((L\nabla\delta\mathbf{w})_i, \nabla\phi) + (\delta u_i, \phi) \\ (\delta w_i, \psi) - \varepsilon^2(\nabla\delta u_i, \nabla\psi) + ((A\delta\mathbf{u})_i, \psi) - \frac{1}{\nu}(\chi_{\mathcal{A}(u_i)}\delta u_i, \psi) + \frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{\nu}(\chi_{\mathcal{A}(u_j)}\delta u_j, \psi) - ((A\delta\mathbf{u})_j, \psi)\right] \end{array} \right),
$$

where $\chi_{\mathcal{A}(u_i)}$ is the characteristic function of the set

$$
\mathcal{A}(u_i) := \{x \in \Omega : u_i(x) < 0\}.
$$

We now want to discretize the time-discrete problem (14)–(15) in space and then discuss its efficient solution.

6 Finite element approximation

For the discretization in space we use finite elements [55]. In the following we assume for simplicity that Ω is a polyhedral domain. Generalizations to curved domains are possible using boundary finite elements with curved faces. Let $\{\mathcal{R}_h\}_{h>0}$ be a triangulation of Ω into disjoint open rectangular elements with maximal element size h, J_h be the set of nodes of $\mathcal{R}_h$ and $p_j \in J_h$ be the coordinates of these nodes. The use of rectangles is motivated by performing the implementation with deal.II [2]. We approximate the infinite-dimensional space $H^1(\Omega)$ by the finite-dimensional space

$$
S_h := \left\{ \phi \in C^0(\overline{\Omega}) : \phi|_R \in Q_1(R) \quad \forall R \in \mathcal{R}_h \right\} \subset H^1(\Omega)
$$

of continuous, piecewise multilinear functions. We denote the standard nodal basis functions of S_h by φ_j for all $j \in J_h$. The discretized version of the penalized problem (14)–(15) consists in finding $(\mathbf{u}_{\nu,h}, \mathbf{w}_{\nu,h}) \in S_h^N \times S_h^N$ such that

$$
\langle \mathbf{F}_{\nu,h}(\mathbf{u}_{\nu,h}, \mathbf{w}_{\nu,h}), \mathbf{v}_h \rangle = \mathbf{0} \quad \forall \mathbf{v}_h \in S_h^N, \tag{17}
$$

where the components are

$$
\left\langle F_{\nu,h}^{(1,i)}(\mathbf{u}_{\nu,h}, \mathbf{w}_{\nu,h}), v_h \right\rangle = \left(u_{\nu,h,i} - u_{h,i}^{(n-1)}, v_h \right)_h + \tau \left((L\nabla\mathbf{w}_{\nu,h})_i, \nabla v_h \right),
$$

$$
\left\langle F_{\nu,h}^{(2,i)}(\mathbf{u}_{\nu,h}, \mathbf{w}_{\nu,h}), v_h \right\rangle = (w_{\nu,h,i}, v_h)_h - \varepsilon^2(\nabla u_{\nu,h,i}, \nabla v_h) + ((A\mathbf{u}_{\nu,h})_i, v_h)_h - \frac{1}{\nu}(\min(0, u_{\nu,h,i}), v_h)_h
$$
$$
+ \frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{\nu}(\min(0, u_{\nu,h,j}), v_h)_h - ((A\mathbf{u}_{\nu,h})_j, v_h)_h\right],
$$

6

for $i = 1, \ldots, N$. The semi-inner product $(\cdot, \cdot)_h$ on $C_0(\overline{\Omega})$ is defined by

$$(f, g)_h := \int_\Omega \pi_h(f(\mathbf{x})g(\mathbf{x})) \, d\mathbf{x} = \sum_{i=1}^m (1, \varphi_i) f(p_i) g(p_i) \quad \forall f, g \in C_0(\overline{\Omega}),$$

where $\pi_h \colon C_0(\overline{\Omega}) \to S_h$ is the Lagrange interpolation operator. Within our finite element framework, for a given $(\mathbf{u}_h, \mathbf{w}_h) \in S_h^N \times S_h^N$, every step of the SSN method for solving (17) requires to compute $(\delta\mathbf{u}_h, \delta\mathbf{w}_h) \in S_h^N \times S_h^N$ satisfying

$$(\delta u_{h,i}, v_h)_h + \tau \left((L \nabla \delta\mathbf{w}_h)_i, \nabla v_h\right) = -F_{v,h}^{(1,i)}(\mathbf{u}_h, \mathbf{w}_h),$$

$$(\delta w_{h,i}, v_h)_h - \varepsilon^2(\nabla \delta u_{h,i}, \nabla v_h) + ((A\delta\mathbf{u}_h)_i, v_h)_h - \frac{1}{\nu}(\chi^h_{\mathcal{A}(u_{h,i})} \delta u_{h,i}, v_h)_h$$
$$+ \frac{1}{N} \sum_{j=1}^N \left[\frac{1}{\nu}(\chi^h_{\mathcal{A}(u_{h,j})} \delta u_{h,j}, v_h)_h - ((A\delta\mathbf{u})_{h,j}, v_h)_h \right] = -F_{v,h}^{(2,i)}(\mathbf{u}_h, \mathbf{w}_h),$$

for all $v_h \in S_h$ and $i = 1, \ldots, N$, where $\chi^h_{\mathcal{A}(u_{h,i})} := \sum_{j=1}^m \chi^h_{\mathcal{A}(u_{h,i})}(p_j)\varphi_j$ with $\chi^h_{\mathcal{A}(u_{h,i})}(p_j) = 0$ if $u_{h,i}(p_j) \geq 0$ and $\chi^h_{\mathcal{A}(u_{h,i})}(p_j) = 1$ otherwise. If we now write a function $v_h \in S_h$ by $v_h = \sum_{j \in J_h} v_{h,j}\varphi_j$ and denote the vector of coefficients by $\mathbf{v}$, the fully discrete linear systems (smooth and nonsmooth) read in matrix form as

$$\begin{bmatrix} I \otimes M & -\mathcal{B} \\ \tau L \otimes K & I \otimes M \end{bmatrix} \begin{bmatrix} \mathbf{w}^{(k+1)} \\ \mathbf{u}^{(k+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ (I \otimes M)\,\mathbf{u}^{(n-1)} \end{bmatrix}, \tag{18}$$

where k denotes the Newton step. The first right hand side is

$$\mathbf{b} = (I \otimes M)\left(-2\left(\mathbf{u}^{(k)}\right)^3 + \frac{3}{2}\left(\mathbf{u}^{(k)}\right)^2\right) + \frac{1}{N}(I \otimes M)\left(\sum_{j=1}^N 2\left(u_j^{(k)}\right)^3 - \frac{3}{2}\left(u_j^{(k)}\right)^2\right)$$

for the use of the smooth potential and

$$\mathbf{b} = 0$$

for the use of the nonsmooth potential. Further, $K := ((\nabla\chi_i, \nabla\chi_j))_{i,j=1,\ldots,m} \in \mathbb{R}^{m \times m}$ is the stiffness matrix, $M := ((\chi_i, \chi_j))_{i,j=1,\ldots,m} \in \mathbb{R}^{m \times m}$ is the lumped mass matrix and $I \in \mathbb{R}^{N \times N}$ is the identity matrix. M is a symmetric positive definite diagonal matrix and K is symmetric and positive semidefinite. For $N = 3$, the block $\mathcal{B}$ is given as

$$\mathcal{B} = \begin{bmatrix} B_{11} & B_2 & B_3 \\ B_1 & B_{22} & B_3 \\ B_1 & B_2 & B_{33} \end{bmatrix},$$

7

where for $i = 1, \ldots, N$

$$
\begin{aligned}
B_{ii} &= \varepsilon^2 K + \left(1 - \frac{1}{N}\right) F_i M F_i, \\
B_i &= -\frac{1}{N} F_i M F_i \\
F_i &= \operatorname{diag}\left(3(u_i^{(k)}(p_j))^2 - 3u_i^{(k)}(p_j) + \frac{1}{2}\right),
\end{aligned}
\tag{19}
$$

in the smooth system and

$$
\begin{aligned}
B_{ii} &= \varepsilon^2 K + \left(1 - \frac{1}{N}\right)\left(\frac{1}{\nu} G_i M G_i - M\right), \\
B_i &= -\frac{1}{N}\left(\frac{1}{\nu} G_i M G_i - M\right), \\
G_i &= \operatorname{diag}\left(\begin{matrix} 1 & u_i^{(k)}(p_j) < 0, \\ 0 & \text{otherwise,} \end{matrix}\right),
\end{aligned}
\tag{20}
$$

in the nonsmooth system. Again, this work uses $A = I - \mathbf{1}\mathbf{1}^T$ as well as $L = I$ and $L = I - \frac{1}{N}\mathbf{1}\mathbf{1}^T$. The system matrix in (18) is denoted by $\mathcal{K}$ for the remainder of the paper.

7 Preconditioning

In both cases, smooth and nonsmooth, a linear nonsymmetric system in saddle point form is at the heart of the computation. We propose the block-triangular preconditioner

$$
\mathcal{P} = \left[\begin{matrix} I \otimes M & 0 \\ \tau L \otimes K & -\hat{\mathcal{S}} \end{matrix}\right],
\tag{21}
$$

motivated by [23, 46], where $\hat{\mathcal{S}}$ is an approximation of the Schur complement $\mathcal{S} = I \otimes M + \tau(L \otimes K)(I \otimes M)^{-1}\mathcal{B}$. The preconditioned matrix becomes

$$
\mathcal{P}^{-1}\mathcal{K} = \left[\begin{matrix} I & -(I \otimes M)^{-1}\mathcal{B} \\ 0 & -\hat{\mathcal{S}}^{-1}\mathcal{S} \end{matrix}\right].
$$

It has Nm eigenvalues at 1 and the remaining ones are characterized as the eigenvalues of the matrix $\hat{\mathcal{S}}^{-1}\mathcal{S}$, which has for $\hat{\mathcal{S}}$ being a good approximation only a small number of different eigenvalue clusters. This in turn is known to result in only a few iterations of suitable Krylov subspace solvers until convergence [23, 46]. Therefore, the theoretical ideal choice is $\hat{\mathcal{S}} = \mathcal{S}$, since the generalized eigenvalue problem $\mathcal{K}x = \lambda \mathcal{P}x$ has in this case only two distinct eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$. But the application of the preconditioner $\mathcal{P}$ requires the action of the inverses of $I \otimes M$ and of $\hat{\mathcal{S}}$. From this

point of view, the ideal choice $\hat{S} = S$ is not practical since this is a full matrix. Inverting the block $I \otimes M$ is cheap as M is a nonsingular diagonal matrix[1]. The remaining task is now to create a Schur complement approximation $\hat{S}$ that is easy to invert and resembles S. The two difficult points thereby are the nondiagonal block matrices $\mathcal{B}$ and $L \otimes K$ (with the nondiagonal matrix L) which couple N equations, respectively. Concerning the latter, note that $L = I - \frac{1}{N}\mathbf{1}\mathbf{1}^T$ is a circulant matrix and can therefore be diagonalized using the Fourier matrix, see [17]. We will see in the next section how we can take advantage of this property for the construction of preconditioners. Besides the gradient energy parts (that only arise in the diagonal blocks) the block matrix $\mathcal{B}$ contains the interacting terms coming from the potential. This includes in the case of the nonsmooth potential the coupling of all penalization terms. In fact, the latter poses the most challenging part, see section 7.2 for details. Regarding the use of the smooth potential, we present in the following an efficient way how to simplify $\mathcal{B}$ to a matrix of block-diagonal structure and how to create an efficient Schur complement approximation based on this simplification.

7.1 Schur complement preconditioner in the smooth case

The first step for the construction of a practical Schur complement preconditioner consists of the approximation of the nondiagonal block matrix $\mathcal{B}$. An ideal case would be a nonsingular block-diagonal approximation in which every diagonal block is equal, i.e. a block matrix of the form $I \otimes B$ as we have it in the (1,1)-block of the system matrix in (18). We already now that such types of matrices are easier to handle and cheap to invert. The specification of $\mathcal{B}$ given in (19) shows that the matrices F_i, $i = 1, \ldots, N$, only depend on the known solution $\mathbf{u}^{(k)}$ from the previous Newton step. From the constraints of the Gibbs simplex (2) we know $\mathbf{0} \lesssim \mathbf{u}^{(k)} \lesssim \mathbf{1}$. Therefore, every diagonal entry of F_i, $i = 1, \ldots, N$, approximately ranges from -2.5 up to 3.5. Together with the estimated order of entries $O(M) = h^2$ for the mass matrix we propose the following approximation of B_{ii} and B_i:

$$\hat{B}_{ii} = \varepsilon^2 K + \left(1 - \frac{1}{N}\right)M, \tag{22}$$

$$\hat{B}_i = 0, \tag{23}$$

for $i = 1, \ldots, N$. In other words, we set all nondiagonal blocks B_i, $i = 1, \ldots, N$, to zero matrices, as their estimated order of entries ranges from $-\frac{3.5h^2}{N}$ up to $\frac{2.5h^2}{N}$, i.e. a small interval around zero. Remember, the first requirement of an ideal approximation mentioned in the beginning of this section is fulfilled now. The second requirement was the nonsingularity and equality of the diagonal blocks. If we would approximate the diagonal blocks B_{ii}, $i = 1, \ldots, N$, with the same strategy as done for the nondiagonal blocks, we would end with the singular matrix $\varepsilon^2 K$ on the diagonals. But remember, every diagonal entry of F_i, $i = 1, \ldots, N$, approximately ranges from -2.5 up to 3.5,

---

[1]For consistent mass matrices the Chebyshev semi-iteration [34] provides a powerful preconditioner [57, 50].

9

so we can also do the simplification $F_i \approx I$ for $i = 1, \ldots, N$. Substituting this into the definition of B_{ii} in (19) gives us exactly our approximation $\hat{B}_{ii}$ in (22). And we see, these approximated diagonal blocks are all equal and nonsingular. All in all, our approximation of $\mathcal{B}$ can be written as

$$\hat{\mathcal{B}} = I \otimes \left(\varepsilon^2 K + \left(1 - \frac{1}{N} \right) M \right).$$

After we have found an ideal approximation $\hat{\mathcal{B}}$, we can go over to the construction of the preconditioner $\hat{\mathcal{S}}$ for the Schur complement $\mathcal{S} = I \otimes M + \tau(L \otimes K)(I \otimes M)^{-1}\mathcal{B}$. Simply replacing $\mathcal{B}$ by $\hat{\mathcal{B}}$ is still not practical since it remains a full matrix. However, what we can efficiently invert is a matrix product of the form AB, with e.g. A and B symmetric positive definite. This forms exactly the main idea for the derivation of an efficient approximation $\hat{\mathcal{S}}$ (see also [48]): Construct a preconditioner of the form $\hat{\mathcal{S}} = ABC$, with A, B and C symmetric positive definite, such that the exact Schur complement is captured as close as possible. Therefore, we propose the following Schur complement preconditioner

$$\hat{\mathcal{S}} = \hat{\mathcal{S}}_1 (I \otimes M)^{-1} \hat{\mathcal{S}}_2 = \left(\frac{N}{N-1}(I \otimes M) + \tau(L \otimes K) \right)(I \otimes M)^{-1}\hat{\mathcal{B}}$$

$$= I \otimes M + \tau(L \otimes K)(I \otimes M)^{-1}\hat{\mathcal{B}} + \frac{\varepsilon^2 N}{N-1}(I \otimes K). \tag{24}$$

The first two terms in (24) match the exact Schur complement very close and the remainder is kept relative small due to the small interfacial parameter ε. Let us discuss now the action of the inverse of $\hat{\mathcal{S}}$ which consists of the action of the inverses of the block matrix $\hat{\mathcal{S}}_1$ and of the diagonal block matrix $\hat{\mathcal{S}}_2$ as well as cheap multiplications with the mass matrix M. The efficiency of $\hat{\mathcal{S}}_2 = \hat{\mathcal{B}}$ is already discussed at the beginning of this section. It is of block-diagonal form and contains the same elliptic operator on each diagonal block. Therefore, we approximate the inverse of each diagonal block with one and the same algebraic multigrid (AMG) preconditioner[2]. Regarding $\hat{\mathcal{S}}_1$, it has the same ideal properties as $\hat{\mathcal{S}}_2$ if the mobility matrix $L = I$ is used. Therefore, we take a second AMG preconditioner that approximates the inverse of each diagonal block of $\hat{\mathcal{S}}_1$. In total, for preconditioning the system matrix $\mathcal{K}$, which is of size $2Nm$, we need to apply only two AMG preconditioners of size m each plus multiplications with the diagonal mass matrix. As long as the mesh does not change we do not have to recompute them. In this sense, the application of the preconditioner $\mathcal{P}$ is independent of the number of phases.

Now, let us study the case if $L = I - \frac{1}{N}\mathbf{1}\mathbf{1}^T$ and therefore $\hat{\mathcal{S}}_1$ is not of block-diagonal

---

[2] Algebraic multigrid methods typically exhibit geometric-like properties for positive definite elliptic type operators, but use only algebraic information. This has the advantage that AMG can work well even for complicated geometries and meshes. We refer to [51, 27] for more information on AMG. We also want to emphasize that geometric multigrid (GMG see e.g. [58, 37]) approximations are also well suited to approximate $\hat{\mathcal{S}}_1$ and $\hat{\mathcal{S}}_2$ provided they can be readily applied.

form anymore. As already observed, L is a circulant matrix and can therefore be diagonalized using the Fourier matrix F, i.e.

$$L = F \operatorname{diag}(\lambda_1, \ldots, \lambda_N) F^H,$$

see [17]. This property forms the basis of an efficient Fast Fourier Transform (FFT) based preconditioner which is used e.g. by Stoll [54] and briefly reviewed in the following. The idea is to diagonalize not only L but also the whole block matrix $\hat{S}_1$ (which contains L) since the latter is the matrix whose inverse we have to apply. More precisely, if we apply the FFT to the system $\hat{S}_1 y = g$ we get an equivalent system with the block-diagonal system matrix

$$(F^H \otimes I)\hat{S}_1(F \otimes I) = \frac{N}{N-1}(I \otimes M) + \tau \operatorname{diag}(\lambda_1, \ldots, \lambda_N) \otimes K. \tag{25}$$

Inserting the eigenvalues of L, which are $\lambda_1 = 0$ and $\lambda_2 = \ldots = \lambda_N = 1$, we see that the resulting approximation in (25) (almost) fulfills the two requirements of an ideal approximation: It is of block-diagonal form and almost all diagonal blocks are equal. In fact, only two different diagonal blocks occur, $\frac{N}{N-1}M$ for $\lambda_1 = 0$ and $\frac{N}{N-1}M + \tau K$ for all remaining eigenvalues $\lambda_j = 1$. Typically, the extra effort is neglectable. As the application of the Fourier transform will in general result in complex valued systems, we formulate the blocks in (25) to 2×2 real valued block systems. In detail, we have to solve two types of systems

$$\begin{bmatrix} \frac{N}{N-1}M & 0 \\ 0 & \frac{N}{N-1}M \end{bmatrix} \begin{bmatrix} \tilde{y}_r \\ \tilde{y}_c \end{bmatrix} = \begin{bmatrix} \tilde{g}_r \\ \tilde{g}_c \end{bmatrix}$$

and

$$\begin{bmatrix} \frac{N}{N-1}M + \tau K & 0 \\ 0 & \frac{N}{N-1}M + \tau K \end{bmatrix} \begin{bmatrix} \tilde{y}_r \\ \tilde{y}_c \end{bmatrix} = \begin{bmatrix} \tilde{g}_r \\ \tilde{g}_c \end{bmatrix}.$$

Again, the first of the above systems arises for the diagonal block with $\lambda_1 = 0$ and the second one for all the remaining eigenvalues λ_j. As in [54], we solve these real valued systems with a fixed number of steps of an inexact Uzawa-type method

$$\tilde{y}^{(l+1)} = \tilde{y}^{(l)} + \omega \, \mathcal{P}_1^{-1} \tilde{r}^{(l)},$$

where $r^{(l)}$ denotes the residual and ω is the relaxation parameter. $\mathcal{P}_1$ is a block-diagonal preconditioner whose inverse is applied by inverting the diagonal, nonsingular block $\frac{N}{N-1}M$ or by using an AMG approximation of the block $\frac{N}{N-1}M + \tau K$. Again, independent of the number of phases, this is the one and the same AMG preconditioner every time and in this sense, the circulant approach also leads to a phase-independent preconditioner $\mathcal{P}$. Section 8.2 shows the efficiency of the proposed preconditioning strategy for both cases of L. In particular, we numerically illustrate the independence of $\mathcal{P}$ with respect to the parameters h and N.

Let us turn now to the more challenging case with a nonsmooth potential.

11

7.2 Schur complement preconditioner in the nonsmooth case

As already mentioned in section 7, the nondiagonal block $\mathcal{B}$ within the Schur complement $\mathcal{S} = I \otimes M + \tau(L \otimes K)(I \otimes M)^{-1}\mathcal{B}$ complicates the construction of a Schur complement approximation whose inverse can be applied in an efficient and easy way. Taking only the smooth potential into consideration, we have seen that we can approximate $\mathcal{B}$ with a block-diagonal matrix by using only the bound constraints from the Gibbs simplex (2). However, the inclusion of nonsmoothness involves additional severe penalizations into the system matrix. As can be seen from (20), penalized entries are scattered throughout the diagonals of every block of $\mathcal{B}$. The intensity of the penalization can be controlled by the penalty parameter ν. The smaller ν the stronger the penalization and the more accurate the numerical approximation of the nonsmoothness. In particular, regarding the nondiagonal blocks of $\mathcal{B}$, the estimated order of those penalized entries is $-\frac{h^2}{N}\left(\frac{1}{\nu}-1\right)$, whereas it is $\frac{h^2}{N}$ for nonpenalized entries. The estimated difference of both types of entries is then of order $\frac{h^2}{N\nu}$, which indicates a severe dependency between h, N and ν. This implies e.g. that this difference decreases with decreasing mesh size. However, we have in mind that we want to go over to adaptive mesh strategies in the future. Therefore, the estimated order of penalized entries is usually of large size and highly differs to the order of the remaining nonpenalized entries. So they should not be neglected. That is why an approximation of the matrix $\mathcal{B}$ in block-diagonal form (as it was done in the smooth case) seems not to be of good quality and our experiences also confirm this observation.

On the other hand, the position of penalized entries is changing with every Newton step. In detail, the places with the penalty parameter in the blocks B_i or B_{ii} depend on the phase u_i. Since all phases are separated in the domain (at least after a few time steps), one cannot expect the penalty parameter to act in the same regions for all phases. Therefore, a common distribution of penalized entries to all phases is not satisfying. That is why an approximation of the matrix $\mathcal{B}$ in form of equal blocks for all phases (as it was also done in the smooth case) seems not to be of good quality and our experiences also confirm this observation. All in all, we end up with keeping the whole block $\mathcal{B}$ within the Schur complement preconditioner. Regarding its construction, we proceed as in the previous section, i.e. we construct a preconditioner in matrix product form such that the exact Schur complement is captured as close as possible. The proposed Schur complement preconditioner is then

$$
\begin{aligned}
\hat{\mathcal{S}} &= \hat{\mathcal{S}}_1 (I \otimes M)^{-1} \hat{\mathcal{S}}_2 \\
&= \left(\frac{N}{N-1}(I \otimes M) + \sqrt{\tau}(L \otimes K) \right)(I \otimes M)^{-1}\left(\frac{N-1}{N}(I \otimes M) + \sqrt{\tau}\mathcal{B} \right) \\
&= I \otimes M + \tau(L \otimes K)(I \otimes M)^{-1}\mathcal{B} + \frac{\sqrt{\tau}\,N}{N-1}\mathcal{B} + \sqrt{\tau}\,\frac{N-1}{N}(L \otimes K), \quad\quad (26)
\end{aligned}
$$

which is similar to the corresponding approximation in the smooth case. The first two terms in (26) match the exact Schur complement. Due to the balanced distribution of τ in form of $\sqrt{\tau}$ in both factors, $\hat{\mathcal{S}}_1$ and $\hat{\mathcal{S}}_2$, the influence of both remainder terms in (26) is reduced. Let us discuss the action of the inverses of $\hat{\mathcal{S}}_1$ and $\hat{\mathcal{S}}_2$. The former was already

studied in the previous section. Therefore, let us concentrate on the latter now. The factor $\hat{S}_2$ still contains the complicated, nondiagonal block $\mathcal{B}$ but its diagonal blocks are now shifted by mass matrices. These shifts were not just products of randomness. Without them, the diagonal blocks of $\hat{S}_2$ would be indefinite. However, we want to apply the action of the inverse of the diagonal blocks of $\hat{S}_2$ (see below). Solving indefinite systems typically causes problems, see e.g. [24] that describes the difficulty of solving Helmholtz problems with classical iterative methods. However, shifting the diagonal blocks as proposed in $\hat{S}_2$ makes them positive definite whenever $\tau < 1$ which is the case for our time discretization scheme. The proposed strategy concerning the solution of the system $\hat{S}_2 y = g$ is the use of a block Jacobi method with a fixed number of steps:

$$y^{(l+1)} = y^{(l)} + \omega \mathcal{P}_2^{-1} r^{(l)}.$$

$\mathcal{P}_2$ is a block-diagonal preconditioner whose inverse is applied by using AMG approximations of the positive definite diagonal blocks of $\hat{S}_2$. Unfortunately, this implies the need of N different AMG preconditioners, which have to be recomputed in every Newton step.

All in all, due to the structure of $\mathcal{B}$, preconditioning the nonsmooth system is more complicated than for the smooth one. But on the other hand, the results obtained with the nonsmooth system are much more accurate than the results obtained with the smooth system, see section 8.1. Additionally, section 8.3 presents the performance of the presented preconditioner and shows promising results.

8 Numerical Results

In this section, we show results for the vector-valued Cahn–Hilliard problems. Concerning the regularized subproblems in the case of the nonsmooth potential, we choose the sequence $\nu_1 = 10^{-1} \geq \nu_2 = 10^{-2} \geq \ldots \geq \nu_{\max} = 10^{-7}$ of penalty parameters and solve each corresponding subproblem $\mathbf{F}_{\nu_i}(\mathbf{u}_h^{(k)}, \mathbf{w}_h^{(k)})$ by the SSN method. In doing so, each Newton method is initialized by the approximate solution of the previous one. After the first time step we fix $\nu = \nu_{\max}$, i.e. from then on it suffices to solve only one SSN method per time step. This is because the initial solution at the beginning might not be a good starting point for the SSN methods. For the (smooth and nonsmooth) Newton method we use the stopping criterion in [40], given by

$$\|\mathbf{F}_\nu(\mathbf{u}_h^{(k)}, \mathbf{w}_h^{(k)})\|_2 \leq \epsilon_{\mathrm{rel}} \, \|\mathbf{F}_\nu(\mathbf{u}_h^{(0)}, \mathbf{w}_h^{(0)})\|_2 + \epsilon_{\mathrm{abs}}, \quad k = 1, \ldots, k_{\max},$$

where we set $k_{\max} = 20$, $\epsilon_{\mathrm{rel}} = 10^{-12}$ and $\epsilon_{\mathrm{abs}} = 10^{-6}$ in all examples. In each Newton step, we solve the linear system (18) by a Krylov subspace solver. The left preconditioners we have presented can be embedded into various of such iterative solvers. For our nonsymmetric system matrix $\mathcal{K}$ we propose the use of a nonsymmetric short-term recurrence method, namely BiCG [28], but note that also other solvers such as QMR [29], BiCGSTAB [56] or GMRES [53] can be used with this preconditioner. We set the BiCG tolerance to be 10^{-7} for the preconditioned relative residual in all examples. The

FFT based preconditioner uses three steps of the inexact Uzawa method and the block Jacobi preconditioner uses five steps. For the multilevel approximations we choose Trilinos AMG approximations [38]. For one application of the preconditioner we take in general 10 steps of a Chebyshev smoother and two V-cycles. The discretization is performed with deal.II [2], which allows the use of the Trilinos library. All numerical experiments listed here are generated with finite elements on rectangles. Experiments show that it is essential to ensure that at least eight vertices lie on the interfaces to avoid mesh effects. Therefore, in all examples we set $\varepsilon \approx \frac{9h}{\pi}$. Concerning the time step, existence and uniqueness of corresponding discrete solutions of the nonsmooth system has been shown in [8, Theorem 2.4] under the condition $\tau < \frac{4\varepsilon^2}{\lambda_A^2 \|L\|}$, where λ_A is the largest positive eigenvalue of A and $\|L\|$ denotes the spectral norm of L. For our choice of matrices A and L this leads to the bound $\tau < 4\varepsilon^2$. For the smooth system, we use the same time step sizes. If not mentioned otherwise, the domain is set to be $[0,1]^2$ and for the initial condition, 100 circles with radius 0.0457–0.0525 are randomly distributed over Ω and randomly assigned to the different components.

8.1 Comparison of the smooth and nonsmooth model

In Figure 1, we compare the performance of the smooth and nonsmooth model. It shows the evolution of five phases over 100 time steps for a mesh with size $h = 2^{-8}$. Table 1 illustrates the minimum and maximum value of the order parameter u_1.
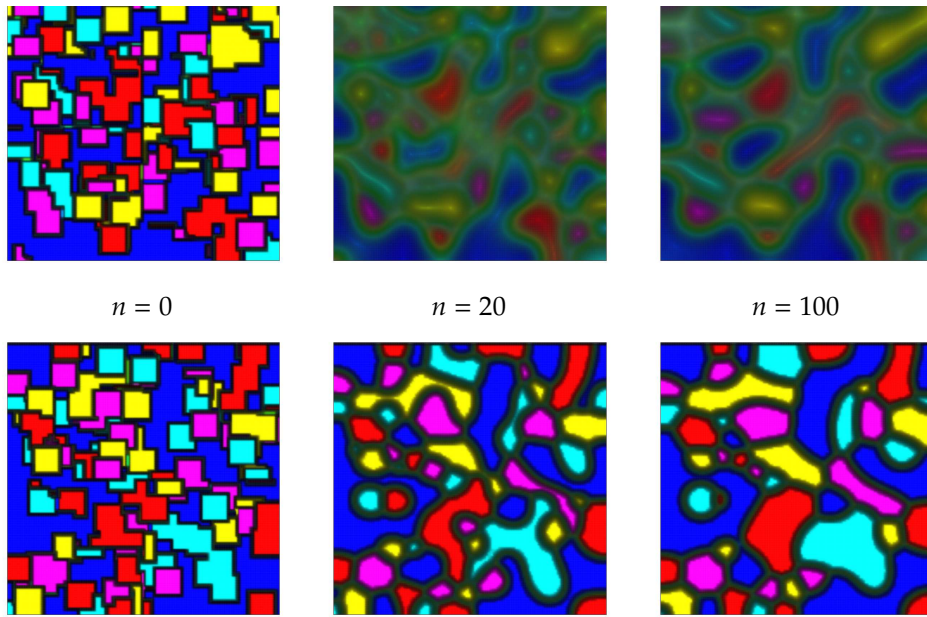

$n = 0$	$n = 20$	$n = 100$

Figure 1: Smooth (above) and nonsmooth (below) computation for five phases.

		time step				
		20	40	60	80	100
min	smooth	−0.02771	−0.02151	−0.02439	−0.02143	−0.02627
	nonsmooth	$-1.186 \cdot 10^{-7}$	$-1.174 \cdot 10^{-7}$	$-1.172 \cdot 10^{-7}$	$-1.204 \cdot 10^{-7}$	$-1.178 \cdot 10^{-7}$
max	smooth	0.9764	0.9803	1.001	0.9845	0.9972
	nonsmooth	1	1	1	1	1

Table 1: Minimum and maximum values of the order parameter u_1 in the smooth and nonsmooth model.

We see that the concentrations are closer to 0 and 1 in the nonsmooth model. Moreover, the sharper interface is given by this model. Comparing both results, the more accurate ones are obtained with the nonsmooth model. This verifies our preference for using a nonsmooth potential.

8.2 Iteration numbers with the smooth model

Next, we consider various uniform mesh sizes and compare the average number of BiCG iterations needed per Newton step over 50 time steps. Moreover, we test the robustness with respect to the number of phases. Figure 2 and 3 show the results for the smooth model. In the legend of Figure 2 the number of degrees of freedom m is listed. The computations are done for $N = 7$ phases. The legend of Figure 3 shows the number of phases N. Here, the computations are done for the mesh size $h = 2^{-8}$. In all calculations, the number of BiCG iterations does not exceed 16. The iteration numbers for the cases $L = I$ and $L = I - \frac{1}{N}\mathbf{1}\mathbf{1}^T$ are almost the same, whereby the results in Figure 2 and 3 are obtained with the latter circulant matrix L. All in all, the results show the robustness of our preconditioner for both, the mesh size and the number of phases.
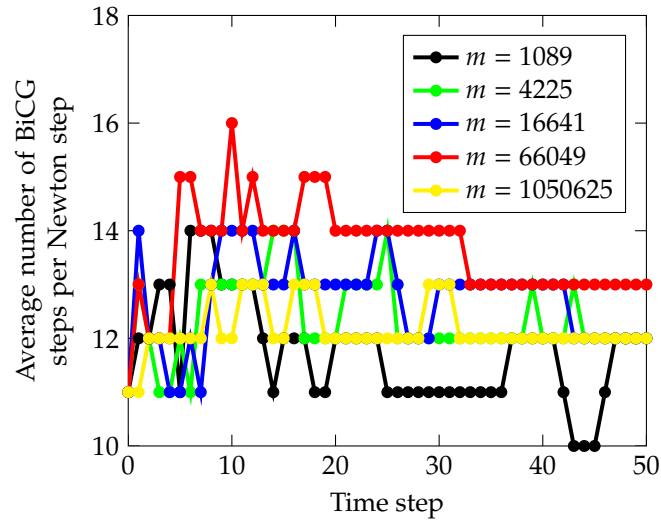
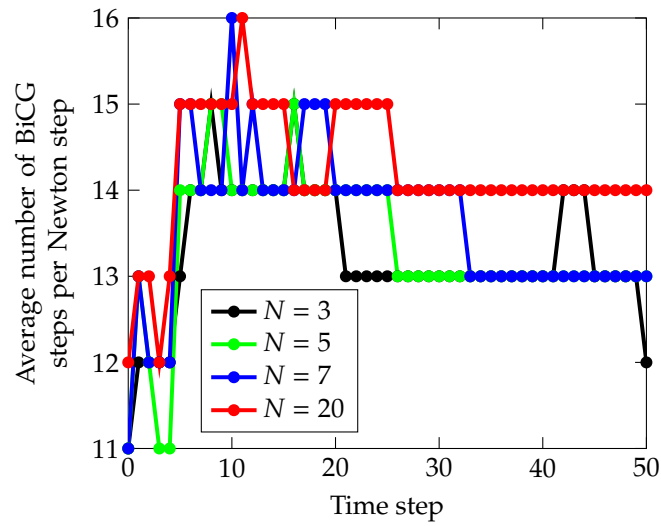Figure 2: Results for 50 time steps of the smooth model with $N = 7$.

Figure 3: Results for 50 time steps of the smooth model with $h = 2^{-8}$.

8.3 Iteration numbers with the nonsmooth model

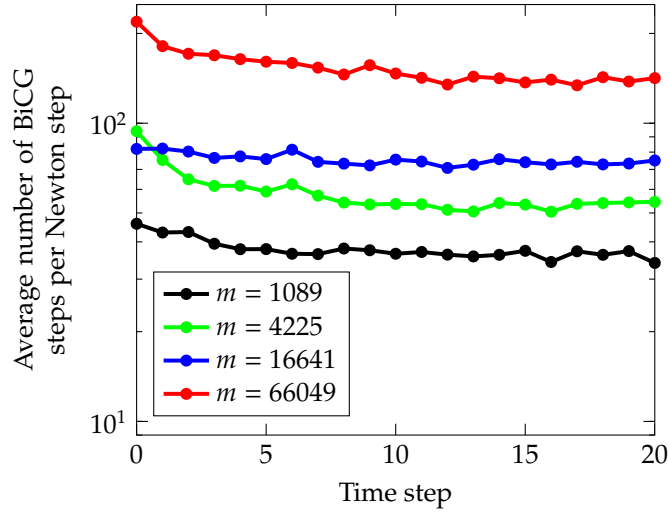Similar computations are done with the nonsmooth model.

Figure 4: Results for 20 time steps of the nonsmooth model with $N = 5$ and $L = I$.
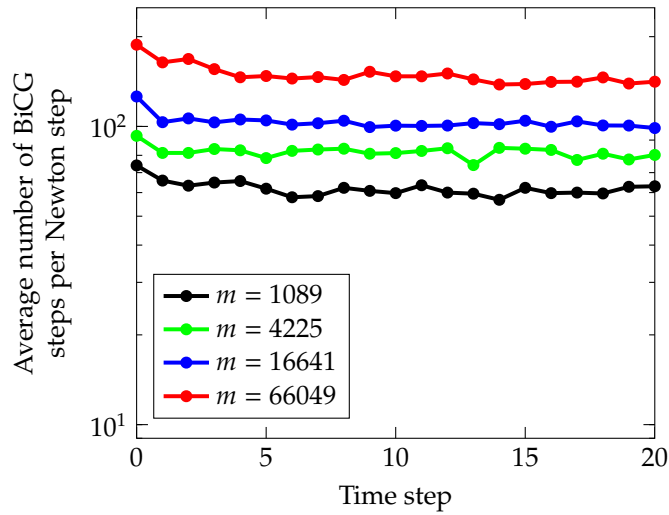

Figure 5: Results for 20 time steps of the nonsmooth model with $N = 20$ and $L = I$.

Again, we consider various uniform mesh sizes and compare the average number of BiCG iterations needed per Newton step over 20 time steps. Thereby, we use five phases in Figure 4 and 20 phases in Figure 5. The mobility matrix here is $L = I$. Results for the circulant L are illustrated in Figure 6 and 7 and are quite similar to the ones with $L = I$. Unfortunately, the numerical mesh independence of our preconditioner has been lost. Nevertheless, in consideration of the complexity of the nonsmooth problem, the gaps in the iteration numbers between two sequent mesh sizes are

satisfying. Note, that the whole system size is in fact $2Nm$. The maximum (maximum average) BiCG iteration numbers are 255 (219) for five phases and 212 (188) for twenty phases when $L = I$ is used as well as 202 (158) for five phases and 304 (235) for twenty phases when $L = I - \frac{1}{N}\mathbf{11}^T$ is used. But we are still thinking about improvements. An interesting observation arises while comparing the iteration numbers between Figure 4 and 5 for a fixed mesh size. While the numbers for the first three mesh sizes up to 16641 degrees of freedom are better in the five-phase model, the numbers for the smallest mesh size (red line) are better in the 20-phase model or at least they are of the same size as the ones in the five-phase model. We expect this effect to be more significant for finer meshes or for higher phase numbers.
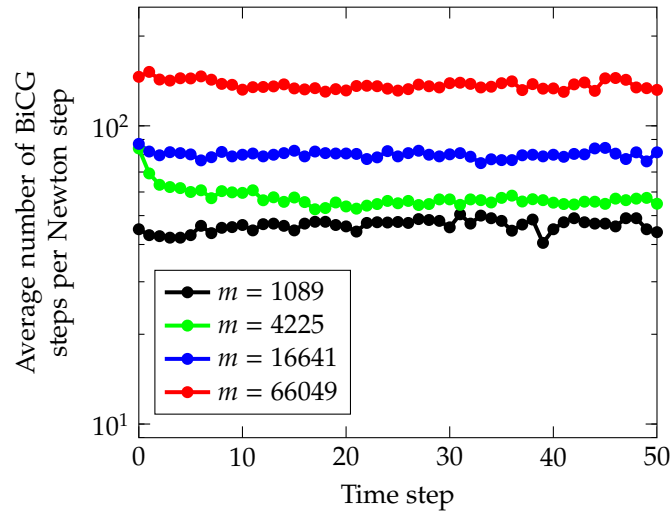

Figure 6: Results for 50 time steps of the nonsmooth model with $N = 5$ and $L = I - \frac{1}{N}\mathbf{11}^T$.
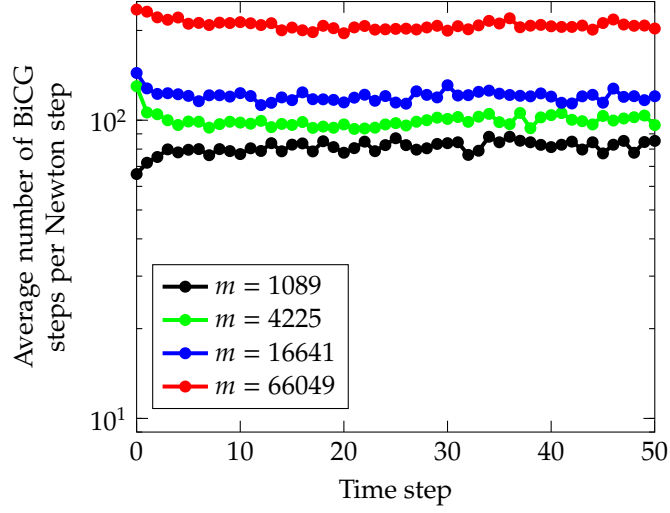
18

Figure 7: Results for 50 time steps of the nonsmooth model with $N = 20$ and $L = I - \frac{1}{N}\mathbf{1}\mathbf{1}^T$.

Finally, we compare the average BiCG iteration numbers with and without preconditioning for the first time step in Figure 8. Here, the number of phases is three and $L = I$. As mentioned in the beginning of this section, we solve within this first time step seven SSN methods for the sequence of penalty parameters $\nu_1 = 10^{-1}, \ldots, \nu_7 = 10^{-7}$, respectively. As can be seen from this, although the preconditioned iteration numbers are considerably worse compared to the one in the smooth model, the preconditioned version always outperforms the unpreconditioned method. A factor of 1500 (3500) for $h = 2^{-5}$ ($h = 2^{-6}$) can be observed and again we would expect this to be even more significant if a larger number of phases or degrees of freedom is used.
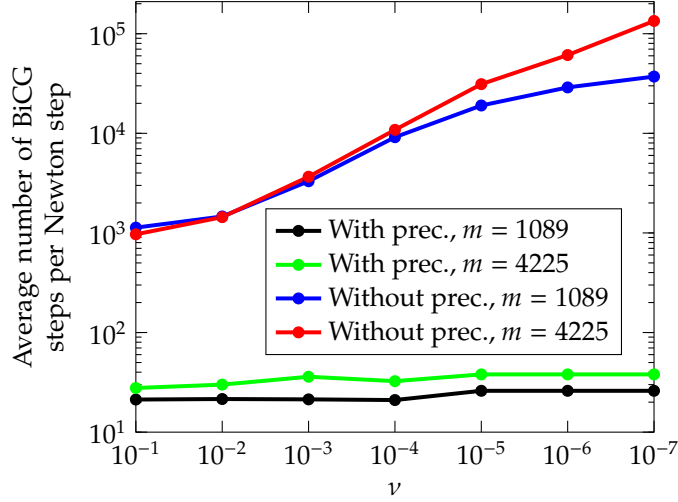
Figure 8: Preconditioning vs. no preconditioning in the nonsmooth model.

8.4 Comparison of different constant mobilities

In this section, we numerically compare various constant mobilities. We test the choices $L_1 = I$ as well as the two different circulant mobilities $L_2 = I - \frac{1}{N}\mathbf{1}\mathbf{1}^T$ and

$$L_3 = \begin{bmatrix} 0.5 & -0.25 & 0 & -0.25 \\ -0.25 & 0.5 & -0.25 & 0 \\ 0 & -0.25 & 0.5 & -0.25 \\ -0.25 & 0 & -0.25 & 0.5 \end{bmatrix}.$$

The test example is taken from [44, Section 4.6] and considers four phases. We use their initial state on the domain $\Omega = [0,1]^2$ with mesh size $h = 2^{-7}$ as illustrated in Figure 9. The expression $u_i^{(0)}$, $u_j^{(0)}$ indicates that the initial phase $u_i^{(0)}$ is set to be randomly between 0.5 and 0.51 and the initial phase $u_j^{(0)}$ is set to be $1 - u_i^{(0)}$ in the considered rectangle. The results after 50 time steps are illustrated in Figure 10. There are no differences in the morphologies, only small time differences seem to occur. The same observations are made with a smooth potential. In [21] and references therein, the requirement of concentration dependent mobilities for many applications is mentioned. For example, if the mobility in the interface is larger than in the pure phases. This motivates us to consider concentration dependent mobilities in the future in order to model other physical situations.
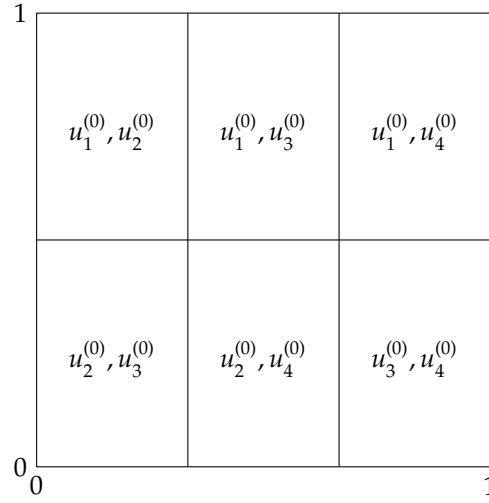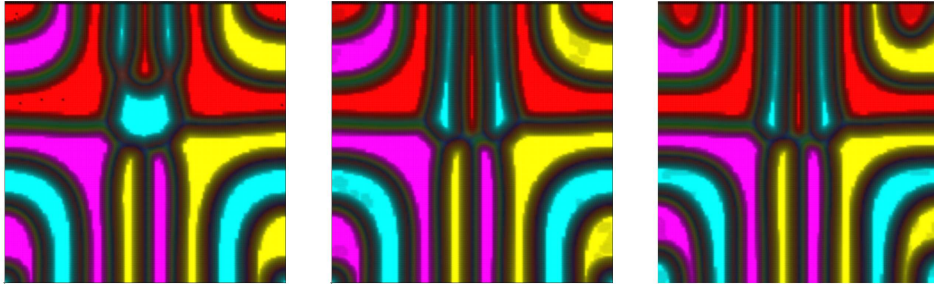
20

Figure 9: Initial state.

Figure 10: Results for L_1, L_2 and L_3 (from left to right) after 50 time steps.

9 Other available solvers

In this section, we compare our solution technique to existing solution methods. A block preconditioning strategy for multi-component Cahn–Hilliard problems is proposed by Boyanova et al. [12]. Their technique is a generalization of the two-phase model from [11, 1]. Both works study the Cahn–Hilliard equations with and without convection. Contrary to us, they do not consider nonsmooth potentials and their matrices in the second line of (18) are all of block-diagonal form, including the additional convection matrix. The nonlinear systems at each time step are solved by a quasi-Newton method. Each Newton step involves the solution of a nonsymmetric linear system with the Jacobian matrix. By simplifying the system matrix they get a preconditioner, which is proven to yield optimal results. In our previous work [10], we

have already tested the method of Boyanova et al. [11] applied to a nonsmooth scalar Cahn–Hilliard system. The idea was based on ignoring the penalization term within the preconditioner and we have shown a severe dependency between the mesh size and the penalty parameter. All in all, this strategy seems to be effective for extremely fine meshes, but typical simulations do not refine that strongly.

A nonlinear multigrid method is proposed by Lee et al. [43, 44] who consider smooth potentials. The first work uses $L = I$ as mobility and a practically unconditionally gradient stable scheme is presented which is based on a nonlinear splitting method. This allows them to decouple the N-component Cahn–Hilliard system into $N-1$ scalar Cahn–Hilliard equations. The efficiency of the approach is shown by means of the average CPU time whose convergence rate is linear with respect to the number of phases. Although they could take arbitrarily large time steps, small time steps have to be used in order to accurately resolve the dynamics. This is an argument for the implicit discretization scheme. The second work [44] uses a concentration dependent mobility matrix and Crank-Nicolson's method for the discretization in time. The authors develop a Full Approximation Storage multigrid method with a pointwise Gauß-Seidel relaxation scheme as a smoother. The nonlinearity is treated using one Newton step. The second-order accuracy of the numerical scheme is demonstrated. They also visually compare phase separation of four phases with a constant and a degenerate concentration dependent mobility and the differences in morphologies and evolution dynamics can be seen.

Gräser et al. [35] propose globally convergent nonsmooth Schur–Newton methods (NSNMG) for the solution of discrete multi-component Cahn–Hilliard systems. They consider logarithmic as well as obstacle potentials. NSNMG can be formulated in primal-dual form and results in a preconditioned Uzawa method. Each step consists first of the update of the primal variable which includes the direct work with the inverse $(A + \partial\varphi)^{-1}$. Here, A is a symmetric positive definite matrix and $\partial\varphi$ is the sub-differential of the nonsmooth part that includes the indicator function $\sum_{i=1}^{N} \chi_{[0,\infty)}(u_i)$. The second step of NSNMG is to compute the dual variable which can be done by solving a truncated linear saddle-point problem and updating the step size for the Uzawa method. The authors solve the linear systems by a preconditioned GMRES method with restart after 50 steps. They numerically investigated local mesh independence of NSNMG as well as a robust convergence speed of NSNMG and TNNMG for different numbers of phases.

In our previous works [10, 9] dealing with scalar smooth and nonsmooth Cahn–Hilliard systems, we have already tested the use of spectral methods based on the FFT against finite element methods (FEM). As the FFT basis functions are eigenvectors of the difference operators, which form the discrete Laplacian, FFT methods rapidly solve diffusion equations on simple domains. For the scalar Cahn–Hilliard equation using a double-well potential, Eyre [26] presents an FFT method for the fast inversion of the preconditioner. He uses a finite difference scheme and suggests unconditionally gradient stable methods. In order to solve the preconditioned system, Eyre proposes the use of a conjugate gradient squared (CGS) method. The overall effort for solving the linear system is dominated by the FFTs and is $m \log(\sqrt{m})$. Another use of FFT

methods appears in [33, 5, 14] for the smooth, scalar Cahn–Hilliard inpainting problem. They propose a two-dimensional FFT method and achieve fast inpainting. In fact, regarding the computational time, FFT methods are hard to beat, but with respect to more complex problems, spectral methods on complicated domains are difficult, see e.g. [13]. Extending the idea of using an FFT based solution scheme to the nonsmooth problem is a challenge. An efficient FFT based implementation employing a nonsmooth potential would typically suffer from the nonconstant, nonsmooth term that originates in the discretization of the penalization term. In general, it holds for spectral methods, the smoother the function, the faster the convergence. It is shown in [10, 9] how the iteration numbers increase with the nonsmoothness obtained by varying the penalty parameter ν. This is the motivation for this work to focus on a discretization via finite elements especially for the nonsmooth problem.

10 Conclusions

In this paper we have analyzed the linear systems arising in smooth and nonsmooth vector-valued Cahn–Hilliard systems. For the latter, we have applied a semismooth Newton method combined with a Moreau–Yosida regularization technique for handling the pointwise constraints. In order to make the semismooth Newton method more efficient we have used a Krylov subspace solver. We have introduced and studied block-triangular preconditioners using an efficient Schur complement approximation. This approximation can be done using multilevel techniques, such as AMG (as in our case), and the numerical results justify this choice.

Acknowledgements

References

[1] O. Axelsson, P. Boyanova, M. Kronbichler, M. Neytcheva, and X. Wu, *Numerical and computational efficiency of solvers for two-phase problems*, Comput. Math. Appl., 65 (2013), pp. 301–314.

[2] W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II — a general purpose object oriented finite element library*, ACM Trans. Math. Software, 33 (2007), p. 24.

[3] J. W. Barrett and J. 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 (1996), pp. 257–287.

[4] ———, *Finite element approximation of a model for phase separation of a multi-component alloy with non-smooth free energy*, Numer. Math., 77 (1997), pp. 1–34.

23

[5] A. L. Bertozzi, S. Esedoğlu, and A. Gillette, *Inpainting of binary images using the Cahn–Hilliard equation*, IEEE Trans. Image Process., 16 (2007), pp. 285–291.

[6] L. Blank, H. Garcke, L. Sarbu, and V. Styles, *Nonlocal Allen–Cahn systems: Analysis and a primal-dual active set method*, IMA J. Numer. Anal., 33 (2013), pp. 1126–1155.

[7] ——, *Primal-dual active set methods for Allen–Cahn variational inequalities with nonlocal constraints*, Numer. Methods Partial Differential Equations, 29 (2013), pp. 999–1030.

[8] 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 (1996), pp. 111–139.

[9] J. Bosch, D. Kay, M. Stoll, and A. J. Wathen, *Fast solvers for Cahn–Hilliard inpainting*, SIAM J. Imaging Sci., 7 (2014), pp. 67–97.

[10] J. Bosch, M. Stoll, and P. Benner, *Fast solution of Cahn–Hilliard variational inequalities using implicit time discretization and finite elements*, J. Comput. Phys., 262 (2014), pp. 38–57.

[11] P. Boyanova, M. Do-Quang, and M. Neytcheva, *Efficient preconditioners for large scale binary Cahn–Hilliard models*, Comput. Methods Appl. Math., 12 (2012), pp. 1–22.

[12] P. Boyanova and N. Neytcheva, *Efficient numerical solution of discrete multi-component Cahn–Hilliard systems*, Comput. Math. Appl., 67 (2014), pp. 106–121.

[13] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd rev. ed., Dover, Mineola, NY, 2001.

[14] M. Burger, L. He, and C.-B. Schönlieb, *Cahn–Hilliard inpainting and a generalization for grayvalue images*, SIAM J. Imaging Sci., 2 (2009), pp. 1129–1167.

[15] M. Butz, *Computational methods for Cahn–Hilliard variational inequalities*, Ph.D. thesis, University of Regensburg, 2012.

[16] J. W. Cahn and J. E. Hilliard, *Free energy of a nonuniform system. I. Interfacial free energy*, J. Chem. Phys., 28 (1958), pp. 258–267.

[17] M. Chen, *On the solution of circulant linear systems*, SIAM J. Numer. Anal., 24 (1987), pp. 668–683.

[18] T. A. Davis, *UMFPACK Version* 5.6.2 *User Guide*, Technical Report TR-04-003 (revised), Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, 2013.

[19] I. C. Dolcetta, S. F. Vita, and R. March, *Area-preserving curve-shortening flows: From phase separation to image processing*, Interfaces Free Bound., 4 (2002), pp. 325–343.

[20] C. M. Elliott, *The Cahn–Hilliard model for the kinetics of phase separation*, in Mathematical Models for Phase Change Problems, Internat. Ser. Numer. Math. 88, J. F. Rodrigues, ed., Birkhäuser, Basel, 1989, pp. 35–73.

[21] C. M. Elliott and H. Garcke, *Diffusional phase transitions in multicomponent systems with a concentration dependent mobility matrix*, Phys. D, 109 (1997), pp. 242–256.

[22] C. M. Elliott and S. Luckhaus, *A generalised diffusion equation for phase separation of a multi-component mixture with interfacial free energy*, Preprint 195, University of Bonn, 1991.

[23] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Oxford University Press, Oxford, UK, 2005.

[24] O. G. Ernst and M. J. Gander, *Why it is difficult to solve Helmholtz problems with classical iterative methods*, in Numerical Analysis of Multiscale Problems, Lecture Notes in Comput. Sci. Eng. 83, I. G. Graham, T. Y. Hou, O. Lakkis, and R. Scheichl, eds., Springer, Berlin, Heidelberg, 2012, pp. 325–363.

[25] D. J. Eyre, *Systems of Cahn–Hilliard equations*, SIAM J. Appl. Math., 53 (1993), pp. 1686–1712.

[26] ——, *An unconditionally stable one-step scheme for gradient systems*, Technical Report, Department of Mathematics, University of Utah, Salt Lake City, UT, 1997.

[27] R. D. Falgout, *An introduction to algebraic multigrid computing*, Comput. Sci. Eng., 8 (2006), pp. 24–33.

[28] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Numerical Analysis, Lecture Notes in Math. 506, G. A. Watson, ed., Springer, Berlin, Heidelberg, 1976, pp. 73–89.

[29] R. W. Freund and N. M. Nachtigal, *QMR: A quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.

[30] H. Garcke, *Mechanical effects in the Cahn–Hilliard model: A review on mathematical results*, in Mathematical Methods and Models in Phase Transitions, A. Miranville, ed., Nova Science, New York, 2005, pp. 43–77.

[31] H. Garcke, B. Nestler, and B. Stoth, *On anisotropic order parameter models for multi-phase systems and their sharp interface limits*, Phys. D, 115 (1998), pp. 87–108.

[32] ——, *A multi phase field concept: Numerical simulations of moving phasee boundaries and multiple junctions*, SIAM J. Appl. Math., 60 (1999), pp. 295–315.

[33] A. Gillette, *Image inpainting using a modified Cahn–Hilliard equation*, Ph.D. thesis, University of California, Los Angeles, CA, 2006.

[34] G. H. Golub and R. S. Varga, *Chebyshev semi-iterative methods, successive overre-laxation iterative methods, and second order Richardson iterative methods. I, II*, Numer. Math., 3 (1961), pp. 147–156, 157–168.

[35] C. Gräser, R. Kornhuber, and U. Sack, *Nonsmooth Schur–Newton methods for multicomponent Cahn–Hilliard systems*, IMA J. Numer. Anal., (2014). online.

[36] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers Appl. Math. 17, SIAM, Philadelphia, 1997.

[37] W. Hackbusch, *Multigrid Methods and Applications*, Springer Ser. Comput. Math. 4, Springer, Berlin, 1985.

[38] M. A. Heroux, R. A. Bartlett, V. E. Howle, R. J. Hoekstra, J. J. Hu, T. G. Kolda, R. B. Lehoucq, K. R. Long, R. P. Pawlowski, E. T. Phipps, A. G. Salinger, H. K. Thornquist, R. S. Tuminaro, J. M. Willenbring, A. Williams, and K. S. Stanley, *An Overview of Trilinos*, Technical Report SAND2003-2927, Sandia National Laboratories, Albuquerque, NM, 2003.

[39] J. E. Hilliard and J. W. Cahn, *An evaluation of procedures in quantitative metallog-raphy for volume-fraction analysis*, Trans. Am. Inst. Min. Metall. Eng., 221 (1961), pp. 344–352.

[40] M. Hintermüller, M. Hinze, and M. H. Tber, *An adaptive finite-element Moreau–Yosida-based solver for a non-smooth Cahn–Hilliard problem*, Optim. Methods Softw., 26 (2011), pp. 777–811.

[41] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim., 13 (2003), pp. 865–888.

[42] D. A. Kay and A. Tomasi, *Color image segmentation by the vector-valued Allen–Cahn phase-field model: A multigrid solution*, IEEE Trans. Image Process., 18 (2009), pp. 2330–2339.

[43] H. G. Lee, J.-W. Choi, and J. Kim, *A practically unconditionally gradient stable scheme for the N-component Cahn–Hilliard system*, Phys. A, 391 (2012), pp. 1009–1019.

[44] H. G. Lee and J. Kim, *A second-order accurate non-linear difference scheme for the N-component Cahn–Hilliard system*, Phys. A, 387 (2008), pp. 4787–4799.

[45] J. E. Morral and J. W. Cahn, *Spinodal decomposition in ternary systems*, Acta Metallurgica, 19 (1971), pp. 1037–1045.

[46] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput., 21 (2000), pp. 1969–1972.

[47] A. NOVICK-COHEN, *The Cahn–Hilliard equation: Mathematical and modeling perspectives*, Adv. Math. Sci. Appl., 8 (1998), pp. 965–985.

[48] J. W. PEARSON AND A. J. WATHEN, *A new approximation of the Schur complement in preconditioners for PDE-constrained optimization*, Numer. Linear Algebra Appl., 19 (2012), pp. 816–829.

[49] R. L. PEGO, *Front migration in the nonlinear Cahn–Hilliard equation*, Proc. R. Soc. Lond. Ser. A Math. Phys. Sci., 422 (1989), pp. 261–278.

[50] T. REES AND M. STOLL, *Block-triangular preconditioners for PDE-constrained optimization*, Numer. Linear Algebra Appl., 17 (2010), pp. 977–996.

[51] J. W. RUGE AND K. STÜBEN, *Algebraic multigrid*, in Multigrid methods, Frontiers Appl. Math. 3, SIAM, Philadelphia, 1987, pp. 73–130.

[52] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.

[53] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[54] M. STOLL, *One-shot solution of a time-dependent time-periodic PDE-constrained optimization problem*, IMA J. Numer. Anal., (2013). online.

[55] G. STRANG AND G. FIX, *An Analysis of the Finite Element Method*, 2nd ed., Wellesley-Cambridge, Wellesley, MA, 2008.

[56] H. A. VAN DER VORST, *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 631–644.

[57] A. WATHEN AND T. REES, *Chebyshev semi-iteration in preconditioning for problems including the mass matrix*, Electron. Trans. Numer. Anal., 34 (2009), pp. 125–135.

[58] P. WESSELING, *An introduction to multigrid methods*, John Wiley & Sons Ltd., Chichester, UK, 1992.

[59] X.-F. WU AND Y. A. DZENIS, *Phase-field modeling of the formation of lamellar nanostructures in diblock copolymer thin films under inplanar electric fields*, Phys. Rev. E (3), 77 (2008), p. 031807.