ROBUST MULTIGRID METHODS FOR VECTOR-VALUED ALLEN–CAHN EQUATIONS WITH LOGARITHMIC FREE ENERGY

RALF KORNHUBER AND ROLF KRAUSE

Dedicated to Peter Deuflhard on the occasion of his 60th *birthday.*

Abstract. We present efficient and robust multigrid methods for the solution of large, nonlinear, non-smooth systems as resulting from implicit time discretization of vector-valued Allen-Cahn equations with isotropic interfacial energy and logarithmic potential. The algorithms are shown to be robust in the sense that convergence is preserved for arbitrary values of temperature, including the deep quench limit. Moreover, numerical experiments indicate that the convergence speed is independent of temperature, mesh size and time step.

1. Introduction

Since the pioneering paper on grain boundary motion in crystalline solids [1] Allen-Cahn equations have become a prototype model for isothermal phase transitions. Driven by practical relevance and by close connections to fundamental geometric problems [20], Allen-Cahn equations and their discretization have been extensively studied during the last decades (see, e.g., [7, 10, 13] and the literature cited therein). Recently, vector valued versions attracted increasing interest [9, 15, 16, 18, 28].

Numerical computations for Allen-Cahn equations are often based on explicit schemes leading to severe stability constraints limiting the time step in terms of the mesh size. On the other hand, such kind of models describing, e.g., grain growth or other coarsening phenomena naturally call for implicit, unconditionally stable discretizations allowing for the adaptive selection of possibly large time steps in combination with highly non-uniform meshes. In this paper, we present efficient and robust multigrid methods for the solution of the resulting large, nonlinear spatial problems. We concentrate on vector valued Allen-Cahn equations with logarithmic potential. As the smoothness of the nonlinearity deteriorates with decreasing temperature $\theta > 0$ and is lost completely in the deep quench limit $\theta = 0$, we strive for robustness in the sense that convergence properties should be preferably independent of the discretization parameters and of the temperature for all $\theta \geq 0$. For simplicity, we concentrate on isotropic interfacial energy. Extensions to the anisotropic case will be considered elsewhere.

Our starting point is a reformulation of the spatial problems in terms of convex minimization. We present a polygonal Gauß-Seidel method $\mathcal{M}_J$ which is based on successive one-dimensional line search in the direction of the edges of the Gibbs simplex at each node. In contrast to a straightforward block Gauß-Seidel approach this method avoids the solution of local systems with N unknowns reflecting the

N phases. Convergence is proved for all $\theta \geq 0$. Note that such kind of robustness has to be payed by suboptimal complexity, i.e., $\mathcal{O}(N^2 n_J)$ point operations for each iteration step (n_J stands for the number of nodes).

The convergence speed of Gauß-Seidel type methods typically deteriorates with decreasing mesh size. Following [25, 26], polygonal Gauß-Seidel relaxation is accelerated by additional coarse grid corrections $\mathcal{C}_J$, taking the low frequency components of the error into account. The construction of $\mathcal{C}_J$ extends *constrained Newton linearization* [26] and *monotone multigrid* [22] for obstacle problems from the scalar case with box constraints to the vector valued case with Gibbs constraints. The basic idea is to restrict the coarse grid corrections to a neighborhood of an actual intermediate iterate in which the Newton linearization can be controlled by local Lipschitz constants. This procedure can be interpreted as some kind of a priori damping. Monotonically decreasing energy is enforced by additional a posteriori damping which is applied locally to each correction at each node on each refinement level. In the deep quench limit $\theta = 0$, the whole approach collapses to a multilevel version of polygonal Gauß-Seidel relaxation [27]. The resulting multigrid methods have the same order of complexity as the Gauß-Seidel scheme and we prove convergence of a standard and a truncated version for all $\theta \geq 0$. Asymptotic or even global bounds for the convergence rates will be the subject of future research. In our numerical experiments we consider 5 or 3 phases in 2 or 3 space dimensions, respectively. In both examples, our multigrid methods showed uniformly bounded convergence rates with respect to temperature θ. Comparisons with an undamped version indicate that there is still some potential for even more sophisticated damping strategies.

The paper is organized as follows. In the next section, we compile some basic information about vector-valued Allen-Cahn equations and their discretization. The resulting spatial problems can be regarded as special cases of a convex minimization problem as stated in Section 3. In the following section, we introduce polygonal Gauß-Seidel relaxation and give a convergence proof based on geometrical properties of the Gibbs simplex. In Section 5 we provide a general framework for the accelaration of this scheme by additional coarse grid correction. The resulting convergence results are then applied to monotone multigrid methods as derived in Section 6. Numerical experiments finally illustrate the efficiency and robustness of these solvers.

2. Vector–valued Allen–Cahn equations

We consider the evolution 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 $i = 1, \ldots, N$ at each particular point $(x, t) \in Q = \Omega \times [0, T_0]$, $T_0 > 0$, are represented by the components $u_i(x, t)$ of the order parameter $u = (u_1, \ldots, u_N)^T$. As concentrations are non-negative and add up to unity, the order parameter satisfies the constraints

$$u(x,t) \in G = \{v \in \mathbb{R}^N \mid v_i \geq 0,\ \textstyle\sum_{i=1}^N v_i = 1\} \qquad \forall (x,t) \in Q.$$

The closed convex set $G \subset \mathbb{R}^N$ is often called Gibbs simplex. The Ginsburg–Landau total free energy of our system is assumed to take the form

$$\mathcal{E}(u) = \int_\Omega \frac{\varepsilon}{2} \sum_{i=1}^N |\nabla u_i|^2 + \frac{1}{\varepsilon} \Psi(u)\, dx \tag{2.1}$$

with fixed $\varepsilon > 0$. The quadratic interfacial energy is balanced by a free energy Ψ which potentially gives rise to phase separation. We concentrate on a multi-phase version of the well-known logarithmic free energy (cf., e.g., Garcke et al. [15], Barrett & Blowey [4]). More precisely, $\Psi = \Psi_\theta$ is given by

$$\Psi_\theta(u) = \theta\Phi(u) + \Psi_0(u), \qquad u \in G, \tag{2.2}$$

with the convex function

$$\Phi(u) = \sum_{i=1}^{N} u_i \ln(u_i), \qquad u \in G, \tag{2.3}$$

and the quadratic term

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

which is induced by a symmetric interaction matrix $C = (c_{ij})_{i,j=1}^N$ (cf. De Fontaine [14]). Here, θ, θ_c are denoting absolute and critical temperature, respectively. For $\theta < \theta_c$, we assume that Ψ_θ has exact N distinct local minima on G corresponding to the pure phases $i = 1, \dots, N$. For example, this is true for the choice

$$C = (1 - \delta_{ij})_{i,j=1}^N \qquad \text{(Kronecker-}\delta\text{)} \tag{2.4}$$

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. Garcke et al. [16, 17, 18])

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

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

$$\Psi_\theta(w) = \tfrac{1}{2}\theta[(1+w)\ln(\tfrac{1+w}{2}) + (1-w)\ln(\tfrac{1-w}{2})] + \tfrac{1}{2}\theta_c(1-w^2)$$

of the scalar order parameter $w := u_2 - u_1$ is recovered in this way. In the shallow quench, i.e. for $\theta \approx \theta_c$, polynomial free energies generalizing the quartic potential $(1-w^2)^2$ provide good approximations of the logarithmic free energy Ψ_θ (cf. Steinbach et al. [30]). 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 $\theta \to 0$ we obviously have $\Psi_\theta(u) \to \Psi_0(u)$ uniformly on G.

The vector-valued Allen-Cahn equation

$$\varepsilon u_t = \varepsilon \Delta u - \frac{1}{\varepsilon} P\Psi'(u) \tag{2.5}$$

is the projected L^2-gradient flow of the total free energy $\mathcal{E}$ defined in (2.1). The orthogonal projection $Pv = v - \frac{1}{N}(\mathbf{1} \cdot v)\mathbf{1}$ with $\mathbf{1} := (1, \dots, 1)^T \in \mathbb{R}^N$ maps $\mathbb{R}^N$ on the linear subspace

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

It accounts for the fact that admissible variations of $u(x,t) \in G$ must be in H. In addition, we prescribe suitable initial conditions $u(x,0) \in G$ on Ω and we impose Neumann conditions on the boundary $\partial\Omega$ of Ω. Concerning existence, uniqueness and sharp interface limits of (2.5), we refer to Bronsard & Reitich [9] and Garcke et al. [15].

Let $\mathcal{T}_J$ denote a given partition of $\overline{\Omega}$ into simplices with the minimal diameter $h_J = \mathcal{O}(2^{-J})$. We assume that the intersection of two simplices is either a common lower-dimensional simplex or empty. The sets of edges and vertices are denoted by $\mathcal{E}_J$ and $\mathcal{N}_J$, respectively. Then

$$\mathcal{S}_J = \{v \in C(\overline{\Omega}) \mid v|_t \text{ is linear } \forall t \in \mathcal{T}_J\}$$

is the space of piecewise linear finite elements associated with $\mathcal{T}_J$ and

$$\Lambda_J = \{\lambda_p^{(J)} \mid p \in \mathcal{N}_J\}$$

is the nodal basis of $\mathcal{S}_J$. It is convenient to introduce the subspace

$$\mathcal{H}_J = \{v \in \mathcal{S}_J^N \mid v(p) \in H \ \forall p \in \mathcal{N}_J\} \subset \mathcal{S}_J^N$$

and the subset

$$\mathcal{G}_J^\circ = \{v \in \mathcal{S}_J^N \mid v(p) \in G \text{ and } v_i(p) > 0,\ \forall i = 1, \dots, N,\ p \in \mathcal{N}_J\}.$$

Discretizing the Allen-Cahn equation (2.5) in time by the backward Euler scheme with step size $\tau > 0$ and in space by piecewise linear finite elements with respect to $\mathcal{T}_J$, we obtain the nonlinear problem

$$u_J^k \in \mathcal{G}_J^\circ : \qquad a(u_J^k, v) + \frac{\tau}{\varepsilon^2}\langle \Phi'(u_J^k), v\rangle = \ell^k(v) \qquad \forall v \in \mathcal{H}_J \tag{2.6}$$

which has to be solved in the k–th time step. Here,

$$a(v, w) = \langle (I + \theta_c N \tfrac{\tau}{\varepsilon^2} C)v, w\rangle + \tau(\nabla v, \nabla w), \qquad v, w \in \mathcal{S}_J^N,$$

is a symmetric bilinear form and

$$\ell^k(v) = \langle u_J^{k-1}, v\rangle, \qquad v \in \mathcal{S}_J^N,$$

is a linear functional. We have used the notation $(\nabla v, \nabla w) = \int_\Omega \nabla v(x) : \nabla w(x)\, dx$ and

$$\langle v, w\rangle = \sum_{p \in \mathcal{N}_J} v(p) \cdot w(p)\, h_p, \qquad h_p = \int_\Omega \lambda_p^{(J)}(x)\, dx,$$

stands for the lumped L^2-product on $\mathcal{S}_J^N$. Note that we used the test space $\mathcal{H}_J \subset \mathcal{S}_J^N$ in (2.6) instead of the projection P. For sufficiently small τ, the bilinear form $a(\cdot, \cdot)$ is positive definite on $\mathcal{S}_J^N$. For example, τ has to satisfy $\tau\theta_c N < \varepsilon^2$ for C taken from (2.4). In order to avoid such severe stability restrictions on the time step, the expanding linear part of Ψ' is often discretized explicitly (cf. Blowey & Elliott [6], Eyre [12], Barrett et al. [3], and others). This approach leads to an unconditionally stable scheme of the form (2.6) with the symmetric, positive definite bilinear form

$$a(v, w) = \langle v, w\rangle + \tau(\nabla v, \nabla w), \qquad v, w \in \mathcal{S}_J^N, \tag{2.7}$$

and the right hand side

$$\ell^k(v) = \langle (I - \theta_c N \tfrac{\tau}{\varepsilon^2} C)u_J^{k-1}, v\rangle, \qquad v \in \mathcal{S}_J^N. \tag{2.8}$$

A priori error estimates have been derived mainly for the scalar case and polynomial free energies or for the deep quench limit (see, however, Barrett & Blowey [4]). Such estimates typically degenerate with decreasing ε. For detailed information, we refer, e.g., to recent work of Feng & Prohl [13] and the literature cited therein. A posteriori error estimates for the quartic free energy have been derived recently by Kessler et al. [21].

3. Discrete minimization

We consider the discrete variational problem

$$u_J \in \mathcal{G}_J^\circ : \qquad a(u_J, v) + \phi'_{J,\theta}(u)(v) = \ell(v) \qquad \forall v \in \mathcal{H}_J, \tag{3.1}$$

where $a(\cdot,\cdot)$ denotes a symmetric, positive definite bilinear form with the associated energy norm $\|\cdot\|$ on $\mathcal{S}_J^N$, $\ell \in (\mathcal{S}_J^N)'$, and $\phi_{J,\theta}$ is a convex functional of the form

$$\phi_{J,\theta}(v) = \theta \sum_{p\in\mathcal{N}_J} \Phi(v(p))\; h_p, \qquad v \in \mathcal{G}_J,$$

with a sufficiently smooth convex function $\Phi : G \to \mathbb{R}$. To fix the ideas, we choose Φ according to (2.3) so that the discretized Allen-Cahn equation (2.6) is a special case of (3.1).

The variational problem (3.1) is equivalent to the minimization problem

$$u_J \in \mathcal{G}_J : \quad \mathcal{J}(u_J) + \phi_{J,\theta}(u_J) \le \mathcal{J}(v) + \phi_{J,\theta}(v) \quad \forall v \in \mathcal{G}_J \tag{3.2}$$

with the strictly convex quadratic energy

$$\mathcal{J}(v) = \tfrac{1}{2}a(v,v) - \ell(v)$$

and the closed convex subset $\mathcal{G}_J$,

$$\mathcal{G}_J = \{v \in \mathcal{S}_J^N \mid v(p) \in G,\ p \in \mathcal{N}_J\},$$

which is the closure of $\mathcal{G}_J^\circ$. Utilizing the reformulation (3.2), existence and uniqueness of a solution of (3.1) follows from Proposition 1.2. of Ekeland & Temam [11].

For $\theta \to 0$ the solutions $u_J = u_{J,\theta}$ of (3.2) converge to the unique solution of the quadratic obstacle problem

$$u_J \in \mathcal{G}_J : \quad \mathcal{J}(u_J) \le \mathcal{J}(v) \quad \forall v \in \mathcal{G}_J \tag{3.3}$$

which is obtained from (3.2) for $\theta = 0$. Observe that the variational reformulation of (3.3) leads to the variational inequality

$$u_{J,0} \in \mathcal{G}_J : \quad a(u_{J,0}, v - u_{J,0}) \ge \ell(v - u_{J,0}) \quad \forall v \in \mathcal{G}_J. \tag{3.4}$$

in contrast to the equality (3.1) that holds for $\theta > 0$. In order to obtain robust solvers for (3.1), we shall now derive algorithms for (3.2) that converge for all $\theta \ge 0$. Note that standard Gauß-Newton methods for (3.1) are not robust in this sense.

4. Polygonal Gauss-Seidel relaxation

We consider the splitting

$$\mathcal{H}_J = \sum_{l=1}^{m_J} \mathcal{V}_l, \quad \mathcal{V}_l = \operatorname{span}\{\mu_l\}, \qquad \mu_{l(n,m)} = \lambda_{p_n}^{(J)} E_m,\ \ l = 1,\dots,m_J, \tag{4.1}$$

where $E_m \in \mathbb{R}^N$, $m = 1,\dots,M := \frac{1}{2}N(N-1)$ denote the edges of the Gibbs simplex G, $l = l(n,m)$ is some enumeration and $m_J := n_J M$. Note that $E_m = E_{m(i,j)} = e_i - e_j$, $1 \le i < j \le N$, where e_i is denoting the i-th unit vector in $\mathbb{R}^N$ and $m = m(i,j)$ is another enumeration. As each subspace $V_l = \operatorname{span}\{\lambda_{p_n}^{(J)} E_m\}$ is generated by a nodal basis function $\lambda_{p_n}^{(J)}$ and an edge E_m of the polygon G, the nonlinear successive subspace correction method (cf. Xu [31], Kornhuber [22, 26]) generated by the splitting (4.1) can be regarded as a *polygonal Gauß-Seidel relaxation*. It reads as follows.

Let $\theta \geq 0$ be fixed. For a given ν-th iterate $u_J^\nu \in \mathcal{G}_J$ we compute a sequence of intermediate iterates $w_l^\nu = w_{l-1}^\nu + v_l^\nu$, $l = 1, \ldots, m_J$, starting with $w_0^\nu = u_J^\nu$. The corrections v_l^ν are the unique solutions of the local subproblems

$$\begin{aligned} v_l^\nu \in \mathcal{D}_l^\nu : \quad & \mathcal{J}(w_{l-1}^\nu + v_l^\nu) + \phi_{J,\theta}(w_{l-1}^\nu + v_l^\nu) \\ & \leq \mathcal{J}(w_{l-1}^\nu + v) + \phi_{J,\theta}(w_{l-1}^\nu + v) \quad \forall v \in \mathcal{D}_l^\nu \end{aligned} \tag{4.2}$$

where the closed convex subsets $\mathcal{D}_l^\nu$ are defined by

$$\mathcal{D}_l = \mathcal{D}_l(w_{l-1}) = \{v \in \mathcal{V}_l \mid w_{l-1}^\nu + v \in \mathcal{G}_J\}. \tag{4.3}$$

Finally, the new iterate of the polygonal Gauß-Seidel relaxation is given by

$$u_J^{\nu+1} = \mathcal{M}_{J,\theta}(u_J^\nu) := w_{m_J}^\nu = u_J^\nu + \sum_{l=1}^{m_J} v_l^\nu. \tag{4.4}$$

For ease of presentation the indices θ and ν will be frequently skipped in the sequel. By construction, we have monotonically decreasing energy,

$$\mathcal{J}(w_l) + \phi_J(w_l) \leq \mathcal{J}(w_{l-1}) + \phi_J(w_{l-1}), \quad l = 1, \ldots m_J, \tag{4.5}$$

and the uniqueness of the solutions v_l of (4.2) implies

$$\mathcal{J}(\mathcal{M}_J(w)) + \phi_J(\mathcal{M}_J(w)) = \mathcal{J}(w) + \phi_J(w) \Leftrightarrow \mathcal{M}_J(w) = w. \tag{4.6}$$

For $\theta > 0$ and given w_{l-1}, each correction $v_{l(n,m(i,j))} = z_l \lambda_{p_n}^{(J)}(e_i - e_j)$, can be equivalently computed from the scalar nonlinear equation

$$z_l \in (-\alpha_l, \beta_l) : \quad d_l z_l - r_{l-1} + \theta[\ln(\alpha_l + z_l) - \ln(\beta_l - z_l)]h_{p_n} = 0 \tag{4.7}$$

where $\alpha_l = (w_{l-1}(p_n))_i$, $\beta_l = (w_{l-1}(p_n))_j$, $d_l = a(\mu_l, \mu_l)$, $r_{l-1} = \ell(\mu_l) - a(w_{l-1}, \mu_l)$. In the deep quench limit $\theta = 0$, we directly obtain

$$z_l = \max\{-\alpha_l, \min\{\beta_l, r_{l-1}/d_l\}\}. \tag{4.8}$$

Assuming that each local problem (4.7) can be solved with $\mathcal{O}(1)$ complexity, each overall iteration step (4.4) requires $\mathcal{O}(N^2 n_J)$ point operations. The suboptimal quadratic growth in N will turn out to be the price to be paid for robustness.

The following representation lemma which was already stated without proof by Kornhuber & Krause [27] will be crucial for the convergence of the polygonal Gauß-Seidel relaxation (4.4) in the limit case $\theta = 0$.

Lemma 4.1. *For any given $u^*, v \in G$ there is a decomposition*

$$v = u^* + \sum_{e=1}^{N-1} \eta_e, \qquad \eta_e \in \mathit{span}\{E_{m_e}\}, \tag{4.9}$$

with the property

$$u^* + \eta_e \in G \qquad \forall e = 1, \ldots, N-1. \tag{4.10}$$

Proof. We present a proof due to Ziegler [32]. Let $w \in H$ with exactly $n^* \leq N$ non-zero components w_{i_k}, $k = 1, \ldots, n^*$. Assume for the moment that for such w there is a decomposition

$$w = \sum_{e=1}^{n^*-1} \eta_e, \qquad \eta_e \in \mathrm{span}\{E_{m_e}\} \tag{4.11}$$

with the properties

$$(\eta_e)_{i_k} w_{i_k} \geq 0, \quad k = 1, \dots, n^*, \qquad (\eta_e)_{i_k} = 0, \quad k = n^*+1, \dots, N, \tag{4.12}$$

for all $e = 1, \dots, n^* - 1$. Then the assertion can be shown in the following way.

For given $u^*, v \in G$, let

$$v - u^* = \sum_{e=1}^{n^*-1} \eta_e$$

be a decomposition (4.11) of $w := v - u^* \in H$ satisfying (4.12). It remains to show (4.10) or, equivalently,

$$u_i^* + (\eta_{e_0})_i \geq 0 \qquad \forall i = 1, \dots, N$$

for each fixed $e_0 = 1, \dots, n^* - 1$. As a consequence of (4.12) $v_i \geq u_i^*$ leads to $(\eta_{e_0})_i \geq 0$ so that

$$u_i^* + (\eta_{e_0})_i \geq u_i^* \geq 0.$$

In the case $v_i \leq u_i^*$ we get $(\eta_e)_i \leq 0$ $\forall e = 1, \dots, n^* - 1$ and therefore

$$u_i^* + (\eta_{e_0})_i \geq u_i^* + \sum_{e=1}^{n^*-1} (\eta_e)_i = v_i \geq 0.$$

This proves the assertion.

We still have to show existence of a decomposition (4.11) with the properties (4.12). For given $w \in H$ let

$$I^+(w) = \{i \mid w_i > 0\}, \;\; I^-(w) = \{i \mid w_i < 0\}, \;\; I^*(w) = I^+(w) \cup I^-(w).$$

We proceed by induction on the number of elements $|I^*(w)|$ of $I^*(w)$. The first non-trivial case is $|I^*(w)| = 2$, i.e., $I^+(w) = \{i_1\}$, $I^-(w) = \{j_1\}$. Then $w \in H$ implies the representation

$$w = w_{i_1}(e_{i_1} - e_{j_1}) =: \eta_1 \in \operatorname{span}\{E_{m(i_1,j_1)}\}$$

which clearly has the form (4.11). It is easily checked that the conditions (4.12) are satisfied. Now assume that the assertion holds for all $w \in H$ with $|I^*(w)| \leq n^*$ and fixed $n^* \geq 2$. We consider some $v \in H$ with $|I^*(v)| = n^* + 1$. First we select $i_0 \in I^*(v)$ such that

$$0 < |v_{i_0}| = \min_{i \in I^*(v)} |v_i|.$$

Without loss of generality let $i_0 \in I^+(v)$. Then we set

$$w = v - \eta_{n^*}, \qquad \eta_{n^*} = v_{i_0}(e_{i_0} - e_{j_0}). \tag{4.13}$$

with arbitrary $j_0 \in I^-(v)$. By construction, we have

$$I^+(w) = I^+(v) \setminus \{i_0\}, \quad I^+(w) \cap I^-(v) = \emptyset, \quad i_0 \notin I^*(w), \quad |I^*(w)| \leq n^*. \tag{4.14}$$

Hence, there is a decomposition (4.11) of w with the properties (4.12). Inserting it in (4.13) we obtain $v = \sum_{e=1}^{n^*} \eta_e$ and

$$(\eta_e)_i v_i \geq 0 \quad \forall i \in I^*(v), \qquad (\eta_e)_i = 0 \quad \forall i \notin I^*(v) \qquad \forall e = 1, \dots, n^*$$

follows directly from related properties of w and (4.14). This concludes the proof. □

Now we are ready to prove convergence.

Theorem 4.1. *For each fixed $\theta \geq 0$ and arbitrary initial iterate $u_J^0 \in \mathcal{G}_J$, the polygonal Gauß-Seidel relaxation (4.4) converges to the solution u_J of (3.2).*

Proof. We first state that any subsequence of $(u_J^\nu)_{\nu\in\mathbb{N}}$ has a convergent subsequence $(u_J^{\nu_k})_{k\in\mathbb{N}}$ with limit $u_J^* \in \mathcal{G}_J$ satisfying

$$\mathcal{M}_J(u_J^*) = u_J^*. \tag{4.15}$$

Based on the strict monotonicity (4.5), (4.6) of the energy $\mathcal{J} + \phi_J$, this assertion follows literally as in the proof of Theorem 2.1 in [24].

We now show that u_J^* is the solution u_J of (3.2). First let $\theta > 0$. Then, rewriting (4.2) in variational form and utilizing (4.15), we obtain

$$a(u_J^*, \mu_l) + \phi_J'(u_J^*)(\mu_l) = \ell(\mu_l) \qquad \forall l = 1, \ldots, m_J.$$

As $\mathcal{H}_J = \operatorname{span}\{\mu_l \mid l = 1, \ldots, m_J\}$, this variational equality holds for all $v \in \mathcal{H}_J$. Hence, u_J^* solves (3.1) and therefore $u_J^* = u_J$.

Now let $\theta = 0$. In this case, the variational reformulation of (4.2) takes the form

$$v_l \in \mathcal{D}_l(w_{l-1}): \quad a(v_l, \eta - v_l) \geq \ell(\eta - v_l) - a(w_{l-1}, \eta - v_l) \quad \forall \eta \in \mathcal{D}_l(w_{l-1})$$

for all $l = 1, \ldots, m_J$. Hence, (4.15) implies

$$0 \geq \ell(\eta) - a(u_J^*, \eta) \qquad \forall \eta \in \mathcal{D}_l(u_J^*) \quad \forall l = 1, \ldots, m_J. \tag{4.16}$$

Let $v \in \mathcal{G}_J$ be arbitrarily chosen. Then, utilizing Lemma 4.1, we find $\eta_e(p) \in \operatorname{span}\{E_{m_e}\}$ with the property

$$v(p) - u_J^*(p) = \sum_{e=1}^{N-1} \eta_e(p), \qquad u_J^*(p) + \eta_e(p) \in G,$$

for each $p \in \mathcal{N}_J$. Hence, $\eta = \lambda_{p_n}^{(J)} \eta_e(p_n) \in \mathcal{D}_{l(n,m_e)}(u_J^*)$ can be inserted in (4.16) for each $p_n \in \mathcal{N}_J$ and $e = 1, \ldots, N-1$. Summing up all the resulting inequalities, we obtain

$$a(u_J^*, v - u_J^*) \geq \ell(v - u_J^*)$$

and therefore $u_J^* = u_J$.

We have shown that any subsequence of $(u_J^\nu)_{\nu\in\mathbb{N}}$ has a subsequence converging to u_J. Hence, the whole sequence must converge to u_J. □

Observe that for $\theta > 0$ a splitting of the form (4.1) into $n_J(N-1) < m_J$ subspaces associated with any selection E_{m_k}, $k = 1, \ldots, N-1$, of edges for each node $p \in \mathcal{N}_J$ would lead to a $\mathcal{O}(Nn_J)$ algorithm which still converges for $\theta > 0$. However, the convergence speed typically deteriorates for $\theta \to 0$, because in general no convergence occurs for $\theta = 0$.

As another alternative to (4.1) one might consider the splitting of $\mathcal{H}_J$ into the $(N-1)$-dimensional subspaces

$$V_i = \operatorname{span}\{\lambda_{p_n}^{(J)} \mid v \in H\}, \qquad n = 1, \ldots, n_J.$$

which leads to a natural block version of the well-known scalar Gauß-Seidel relaxation. Though such an algorithm is clearly convergent, it seems to be less attractive, because the solution of the resulting $(N-1)$-dimensional nonlinear subproblems becomes more and more complicated for increasing N.

In general, the local problems (4.7) have to be solved iteratively, e.g., by a bisection scheme. The convergence of inexact variants of (4.4) is preserved, if

approximations of z_l are sufficiently accurate. Suitable stopping criteria can be derived according to Kornhuber [26].

5. Constrained Newton linearization and coarse grid correction

Due to a poor representation of low frequency components of the error, the convergence speed of the polygonal Gauß-Seidel relaxation $\mathcal{M}_J$ rapidly degenerates with decreasing mesh size. The convergence speed can be increased by an additional *coarse grid correction* $\mathcal{C}_J$. On the other hand, application of $\mathcal{C}_J$ should not increase the complexity $\mathcal{O}(N^2 n_J)$ of $\mathcal{M}_J$. Successive minimization based on a straightforward extension of the splitting (4.1) by additional coarse grid spaces would not have this property, because the evaluation of nonlinear functionals on such spaces requires additional prolongations and restrictions. We now provide a construction of C_J based on *constrained Newton linearization*. It can be regarded as an extension of classical Newton multigrid methods (cf., e.g., Hackbusch [19]) to piecewise smooth $\phi_{J,\theta}$. The underlying ideas have been discussed in some detail by Kornhuber [26].

Additional coarse correction $\mathcal{C}_{J,\theta}$ leads to the two-stage iteration

$$u_J^{\nu+1} = \mathcal{C}_{J,\theta}\bar{u}_J^\nu, \qquad \bar{u}_J^\nu = \mathcal{M}_{J,\theta}u_J^\nu, \qquad \theta \geq 0. \tag{5.1}$$

Adopting multigrid terminology, $\mathcal{M}_{J,\theta}$ is sometimes called fine grid smoother and $\bar{u}_J^\nu$ is the smoothed iterate. As usual, the subscript θ is mostly skipped in the sequel. The following convergence result is an immediate consequence of Theorem 4.1 and Theorem 3.1 in [26].

Theorem 5.1. *Assume that $\mathcal{C}_J$ is monotone in the sense that*

$$\mathcal{J}(\mathcal{C}_J w) + \phi_J(\mathcal{C}_J w) \leq \mathcal{J}(w) + \phi_J(w) \tag{5.2}$$

holds for all w in the range of $\mathcal{M}_J$. Then, for each initial iterate $u_J^0 \in \mathcal{G}_J$ the iteration (5.1) converges to the solution u_J of (3.2).

Let us proceed with constructing $C_{J,\theta}$ in the case $\theta > 0$. As a starting point we determine a neighborhood $\mathcal{G}_{\bar{u}_J^\nu}$ of a given smoothed iterate $\bar{u}_J^\nu \in \mathcal{G}_J^\circ$ on which the Fréchèt derivative $\phi'_{J,\theta}$ can be controlled by local Lipschitz constants. This requires some preparation. For each node $p \in \mathcal{N}_J$ we define the closed convex subset $G_{\bar{u}_J^\nu(p)} \subset G$,

$$G_{\bar{u}_J^\nu(p)} = \{v \in G \mid v_i \geq \epsilon_i \text{ if } (\bar{u}_J^\nu(p))_i > \epsilon_0 \text{ else } v_i = (\bar{u}_J^\nu(p))_i\}, \tag{5.3}$$

with

$$\epsilon_i = \tfrac{1}{1+\theta}\epsilon_0 + \tfrac{\theta}{1+\theta}(\bar{u}_J^\nu(p))_i, \qquad \epsilon_0 = \sqrt{\tfrac{(\theta(1+\theta))^{\frac{1}{2}}N}{L}} < 1 \tag{5.4}$$

and given sufficiently large $L > 0$. Note that $G_{\bar{u}_J^\nu(p)}$ is a proper subset of G which might even reduce to $G_{\bar{u}_J^\nu(p)} = \{\bar{u}_J^\nu(p)\}$ if $\bar{u}_J^\nu(p)$ is close to a vertex of G.

For an illustration of the case $N = 3$, $\theta > 0$ we refer to Figure 1 showing $G_{\bar{u}_J^\nu(p)}$ for $(\bar{u}_J^\nu(p))_i \geq \varepsilon_0$, $i = 1, 2, 3$, on the left (triangle), for $(\bar{u}_J^\nu(p))_i \geq \varepsilon_0$, $i = 1, 2$, but $(\bar{u}_J^\nu(p))_3 < \varepsilon_0$ in the middle (interval), and for $(\bar{u}_J^\nu(p))_i < \varepsilon_0$, $i = 1, 3$, on the right (point). We set

$$N^\bullet_{\bar{u}_J^\nu(p)} = \{i \in \{1, \dots, N\} | v_i = (\bar{u}_J^\nu(p))_i \; \forall v \in G_{\bar{u}_J^\nu(p)}\}, \; N^\circ_{\bar{u}_J^\nu(p)} = \{1, \dots, N\} \setminus N^\bullet_{\bar{u}_J^\nu(p)}.$$

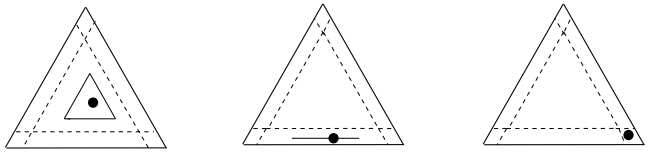

FIGURE 1. The subset $G_{\bar{u}_J^\nu(p)}$ depending on the location $\bullet$ of $\bar{u}_J^\nu(p)$

The definition (5.3) is motivated by the following lemma which can be proved by straightforward computation.

Lemma 5.1. *For each node $p \in \mathcal{N}_J$ and all v, $w \in \mathbb{R}^N$ the Lipschitz continuity*

$$|\Phi''(z_1)(v,w) - \Phi''(z_2)(v,w)| \le L_p |z_1 - z_2|_\infty |v|_\infty |w|_\infty \quad \forall z_1, z_2 \in G_{\bar{u}_J^\nu(p)} \tag{5.5}$$

holds with the local Lipschitz constant

$$L_p = \theta \sum_{i \in N^\circ_{\bar{u}_J^\nu(p)}} \tfrac{1}{\epsilon_i^2} \le (\tfrac{\theta}{1+\theta})^{\frac{1}{2}} L \tag{5.6}$$

denoting $|v|_\infty = \max_{i=1,\dots,N} |v_i|$, $v \in \mathbb{R}^N$.

The Lipschitz continuity (5.5) of the Jacobian Φ'' is inherited by the Fréchèt derivative ϕ_J' on the closed convex subset $\mathcal{G}_{\bar{u}_J^\nu} \subset \mathcal{G}_J$ defined by

$$\mathcal{G}_{\bar{u}_J^\nu} = \{ v \in \mathcal{S}_J^N \mid v(p) \in G_{\bar{u}_J^\nu(p)} \ \forall p \in \mathcal{N}_J \}.$$

Observe that

$$\phi_J(w) = \phi_{\bar{u}_J^\nu}(w) + \text{const.} \qquad \forall w \in \mathcal{G}_{\bar{u}_J^\nu}$$

holds with

$$\phi_{\bar{u}_J^\nu}(w) = \theta \sum_{p \in \mathcal{N}_J} \sum_{i \in N^\circ_{\bar{u}_J^\nu(p)}} w_i(p) \ln(w_i(p))\, h_p. \quad w \in \mathcal{G}_{\bar{u}_J^\nu} \tag{5.7}$$

Hence, for fixed $\bar{u}_J^\nu$, minimization of $\mathcal{J} + \phi_{\bar{u}_J^\nu}$ is equivalent to minimization of $\mathcal{J} + \phi_J$ on $\mathcal{G}_{\bar{u}_J^\nu}$. Using Taylor's expansion

$$\phi_{\bar{u}_J^\nu}(w) \approx \phi_{\bar{u}_J^\nu}(\bar{u}_J^\nu) + \phi'_{\bar{u}_J^\nu}(\bar{u}_J^\nu)(w - \bar{u}_J^\nu) + \tfrac{1}{2}\phi''_{\bar{u}_J^\nu}(\bar{u}_J^\nu)(w - \bar{u}_J^\nu, w - \bar{u}_J^\nu)$$

we approximate $\mathcal{J} + \phi_{\bar{u}_J^\nu}$ by the quadratic energy functional $\mathcal{J}_{\bar{u}_J^\nu}$ defined by

$$\mathcal{J}_{\bar{u}_J^\nu}(w) = \tfrac{1}{2} a_{\bar{u}_J^\nu}(w,w) - \ell_{\bar{u}_J^\nu}(w) \approx \mathcal{J}(w) + \phi_J(w)$$

with the bilinear form

$$a_{\bar{u}_J^\nu}(w,w) = a(w,w) + \phi''_{\bar{u}_J^\nu}(\bar{u}_J^\nu)(w,w)$$

and the linear functional

$$\ell_{\bar{u}_J^\nu}(w) = \ell(w) - \phi'_{\bar{u}_J^\nu}(\bar{u}_J^\nu)(w) + \phi''_{\bar{u}_J^\nu}(\bar{u}_J^\nu)(\bar{u}_J^\nu, w).$$

The resulting quadratic minimization problem

$$w_{\bar{u}_J^\nu} \in \mathcal{G}_{\bar{u}_J^\nu} : \qquad \mathcal{J}_{\bar{u}_J^\nu}(w_{\bar{u}_J^\nu}) \le \mathcal{J}_{\bar{u}_J^\nu}(v) \qquad \forall v \in \mathcal{G}_{\bar{u}_J^\nu} \tag{5.8}$$

can be regarded as a *constrained Newton linearization* of (3.2). The effect of the constraints $\mathcal{G}_{\bar{u}_J^\nu}$ is an *a priori damping* of usual Newton corrections. Note that the linearity of $\mathcal{J}'_{\bar{u}_J^\nu}$ will be crucial for the optimal complexity of each step of the monotone multigrid methods to be constructed later on.

In the next step, we approximate (5.8) by a relaxation scheme generated by additional search directions μ_l with the properties

$$\mu_l \in \mathcal{S}_J, \quad \max_{p\in\mathcal{N}_J} |\mu_l(p)|_\infty = 1, \quad l = m_J+1,\dots,M_J. \tag{5.9}$$

More precisely, we compute intermediate iterates w_l^ν according to

$$w_{m_J}^\nu = \bar{u}_J^\nu, \qquad w_l^\nu = w_{l-1}^\nu + \omega_l^\nu v_l^\nu, \quad l = m_J+1,\dots,M_J. \tag{5.10}$$

Each correction v_l^ν is the solution of the one-dimensional quadratic obstacle problem

$$v_l^\nu \in \mathcal{D}_l^\nu : \quad \mathcal{J}_{\bar{u}_J^\nu}(w_{l-1}^\nu + v_l^\nu) \le \mathcal{J}_{\bar{u}_J^\nu}(w_{l-1}^\nu + v) \quad \forall v \in \mathcal{D}_l^\nu \tag{5.11}$$

with constraints of the form

$$\mathcal{D}_l^\nu = \{v \in \mathcal{V}_l \mid \underline{\psi}_l^\nu \le v \le \overline{\psi}_l^\nu\} \subset \mathcal{V}_l := \operatorname{span}\{\mu_l\} \tag{5.12}$$

satisfying

$$0 \in \mathcal{D}_l^\nu \subset \{v \in \mathcal{V}_l \mid w_{l-1}^\nu + v \in \mathcal{G}_{\bar{u}_J^\nu}\}. \tag{5.13}$$

As usual, the index ν is often skipped in the sequel.

The local damping parameters ω_l in (5.10) are used to ensure the monotonicity (5.2) *a posteriori*. Inductive computation of ω_l can be performed according to the following lemma.

Proposition 5.1. *Let $v_l = z_l\mu_l \neq 0$ be the solution of (5.11). Define*

$$\omega_l = \min\left\{1, 2\left\{\frac{|\ell_{\bar{u}_J^\nu}(\mu_l) - a_{\bar{u}_J^\nu}(w_{l-1},\mu_l)| - L_l B_l^2}{|z_l|(a_{\bar{u}_J^\nu}(\mu_l,\mu_l) + L_l(B_l + |z_l|))}\right\}_+\right\} \tag{5.14}$$

with the local Lipschitz constant L_l and the sum of preceding corrections B_l,

$$L_l = \sum_{p\in\mathcal{N}_J} L_p|\mu_l(p)|_\infty\, h_p, \qquad B_l = \sum_{k=m_J+1}^{l-1} \|v_k\|_\infty \tag{5.15}$$

and $z_+ = \max\{0, z\}$. Then the next intermediate iterate $w_{l+1} = w_l + \omega_l v_l$ fulfills the monotonicity condition

$$\mathcal{J}(w_{l+1}) + \phi_{\bar{u}_J^\nu}(w_{l+1}) \le \mathcal{J}(w_l) + \phi_{\bar{u}_J^\nu}(w_l). \tag{5.16}$$

Proof. Using $B_l \ge \|\bar{u}_J^\nu - w_{l-1}\|_\infty$ the assertion can be shown in the same way as Proposition 4.1 by Kornhuber [26]. □

Observe that

$$|z_l|a(\mu_l,\mu_l)_{\bar{u}_J^\nu} \le |\ell_{\bar{u}_J^\nu}(\mu_l) - a_{\bar{u}_J^\nu}(w_{l-1},\mu_l)| \tag{5.17}$$

holds because $v_l = z_l\mu_l$ solves (5.11). Hence, we obtain $\omega_l = 1$, if L_l is sufficiently small, which, in turn, occurs for sufficiently small θ (cf. Lemma 5.1).

Now we are ready to define the coarse grid correction

$$\mathcal{C}_{J,\theta}\bar{u}_J^\nu = w_{M_J}^\nu = \bar{u}_J^\nu + \sum_{l=m_J+1}^{M_J} \omega_l^\nu v_l^\nu. \tag{5.18}$$

By construction, C_J is monotone in the sense of (5.2).

We still have to consider the case $\theta = 0$. Inserting $\theta = 0$ in (5.4) and (5.7) we directly obtain $G_{\bar{u}_J^\nu(p)} = G$, $L_p = 0$, $\mathcal{G}_{\bar{u}_J^\nu} = \mathcal{G}_J$, $\mathcal{J}_{\bar{u}_J^\nu} = \mathcal{J}$ and $\omega_l = 1$ As a consequence, the coarse grid correction $\mathcal{C}_{J,0}$ resulting from (5.18) for $\theta = 0$ reduces to the successive minimization of the energy $\mathcal{J}$ in the direction of the additional

search directions μ_l (cf. Kornhuber & Krause [27]). Obviously Lemma 5.1 and Proposition 5.1 still remain valid in this case.

Suitable multilevel search directions μ_l and corresponding defect constraints $\underline{\psi}_l$, $\overline{\psi}_l$ will be specified in the next section.

6. Monotone Multigrid

Assume that $\mathcal{T}_J$ is resulting from J refinements of an intentionally coarse partition $\mathcal{T}_0$. In this way, we obtain a sequence of partitions $\mathcal{T}_0, \ldots, \mathcal{T}_J$ and corresponding nested finite element spaces $\mathcal{S}_0 \subset \cdots \subset \mathcal{S}_J$. Though the algorithms and convergence results to be presented can be easily generalized to nonuniform grids, we assume for convenience that the partitions are uniformly refined. For example, in the case of two space dimensions each triangle $t \in \mathcal{T}_k$ is subdivided into four congruent subtriangles in order to produce the next triangulation $\mathcal{T}_{k+1}$. Note that refinement is more delicate for $\Omega \in \mathbb{R}^3$. We refer, e.g., to [5, 8] for further information. Collecting all nodal basis functions from all refinement levels, we obtain the standard multilevel nodal basis $\Lambda_{\mathcal{S}}$,

$$\Lambda_{\mathcal{S}} = \left(\lambda_{p_1}^{(J)}, \lambda_{p_2}^{(J)} \ldots, \lambda_{p_{n_J}}^{(J)}, \ldots, \lambda_{p_1}^{(0)}, \ldots, \lambda_{p_{n_0}}^{(0)}\right). \tag{6.1}$$

The elements $\lambda_{s(k,n)} = \lambda_{p_n}^{(k)}$, $s = 1, \ldots, n_{\mathcal{S}} := n_J + \cdots + n_0$ are ordered from fine to coarse, i.e. $s(k,n) \le s(k',n')$ implies $k \ge k'$.

6.1. **Standard monotone multigrid methods.** In order to apply the abstract framework of the preceding section, we first select the search directions

$$\mu_{l(s,m)} = \lambda_s E_m, \qquad s = 1, \ldots n_{\mathcal{S}},\ m = 1, \ldots, M, \tag{6.2}$$

which obviously have the properties (5.9). The enumeration $l = m_J + 1, \ldots, M_J := n_{\mathcal{S}} M$ first counts all edges E_m before proceeding to the next multilevel nodal basis function λ_s, i.e. $l(s,m) \le l(s',m')$ implies $s \le s'$. The definition of suitable local obstacles $\underline{\psi}_l$, $\overline{\psi}_l$ as occuring in (5.12) requires some preparation. Let $p \in \mathcal{N}_J$ and $E_m = E_{m(i,j)} = e_i - e_j$. With ϵ_i, ϵ_j and ϵ_0 taken from (5.4), we then define

$$\underline{\psi}_{E_m}^{(J)}(p) = \tfrac{1}{N-1}(\epsilon_i - (\bar{u}_J^\nu(p))_i), \quad \overline{\psi}_{E_m}^{(J)}(p) = \tfrac{1}{N-1}((\bar{u}_J^\nu(p))_j - \epsilon_j) \tag{6.3}$$

if $(\bar{u}_J^\nu(p))_i$, $(\bar{u}_J^\nu(p))_j > \epsilon_0$ and

$$\underline{\psi}_{E_m}^{(J)}(p) = \overline{\psi}_{E_m}^{(J)}(p) = 0 \tag{6.4}$$

otherwise. As exactly $N-1$ edges have a non-zero entry in each fixed component i, successive corrections of the form $v_{E_m} E_m$ with $v_{E_m} \in \mathcal{S}_J$ satisfying the conditions

$$\underline{\psi}_{E_m}^{(J)} \ \le\ v_{E_m} \le \overline{\psi}_{E_m}^{(J)} \qquad \forall m = 1, \ldots, M, \tag{6.5}$$

are feasible in the sense that

$$\bar{u}_J^\nu + \sum_{m=1}^{M} v_{E_m} E_m \in \mathcal{G}_{\bar{u}_J^\nu}. \tag{6.6}$$

Observe that the polygonal constraints (6.5) can be checked *separately* for each edge E_m. This is in analogy to box constraints. For $N = 3$, $\theta = 0$ and some $\bar{u}_J^\nu(p)$ the nodal values $v_{E_m}(p) E_m$ with v_{E_m} satisfying (6.5) are illustrated in Figure 2. Here, the star of intervals intersecting at $\bar{u}_J^\nu(p)$ represents the polygonal constraints.

Linear combinations of constrained corrections are feasible in the sense of (6.6) as depicted, e.g., by the dotted lines.

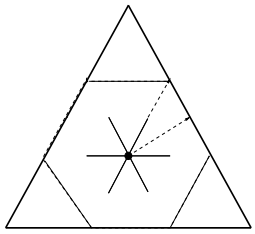

FIGURE 2. Polygonal constraints for $N = 3$, $\theta = 0$ and the location $\bullet$ of $\bar{u}_J^\nu(p)$

Additional prolongations would be necessary to check whether coarse grid corrections $v_{E_m}^{(k)} \in \mathcal{S}_k$ satisfy the fine grid constraints (6.5) or not. This would increase the order of complexity by a factor ranging from $\log n_J$ in case of uniform refinement to n_J in the highly non-uniform case. Following Mandel [29] and Kornhuber [22] we therefore introduce the monotone restriction operators $\underline{r}_{k+1}^k$, $\overline{r}_{k+1}^k : \mathcal{S}_{k+1} \to \mathcal{S}_k$ defined by

$$\begin{aligned} \underline{r}_{k+1}^k v(p) &= \max\{v(q) \mid q \in \mathcal{N}_{k+1} \cap \text{ int supp } \lambda_p^{(k)}\} \\ \overline{r}_{k+1}^k v(p) &= \min\{v(q) \mid q \in \mathcal{N}_{k+1} \cap \text{ int supp } \lambda_p^{(k)}\} \end{aligned} \qquad p \in \mathcal{N}_k,\ v \in \mathcal{S}_{k+1}. \tag{6.7}$$

Starting with $\underline{\psi}_{E_m}^{(J)}$, $\overline{\psi}_{E_m}^{(J)}$ and $v_{E_m}^{(J)} = 0$, coarse grid constraints $\underline{\psi}_{E_m}^{(k)}$, $\overline{\psi}_{E_m}^{(k)}$ are obtained by successive update and restriction

$$\underline{\psi}_{E_m}^{(k)} = \underline{r}_{k+1}^k(\underline{\psi}_{E_m}^{(k+1)} - v_{E_m}^{(k+1)}), \quad \overline{\psi}_{E_m}^{(k)} = \overline{r}_{k+1}^k(\overline{\psi}_{E_m}^{(k+1)} - v_{E_m}^{(k+1)}), \tag{6.8}$$

where $k = J-1, \dots, 0$ and

$$v_{E_m}^{(k+1)} E_m = \sum_{n=1}^{n_{k+1}} \omega_{l(s(k+1,n),m)} v_{l(s(k+1,n),m)}$$

denotes the sum of all local corrections in the direction E_m on level $k+1$. Finally, we define the local obstacles $\underline{\psi}_l$, $\overline{\psi}_l$ according to

$$\underline{\psi}_{l(s(k,n),m)} = \underline{\psi}_{E_m}^{(k)}(p_n)\lambda_{p_n}^{(k)} E_m, \quad \overline{\psi}_{l(s(k,n),m)} = \overline{\psi}_{E_m}^{(k)}(p_n)\lambda_{p_n}^{(k)} E_m. \tag{6.9}$$

The conditions (5.13) are satisfied by construction. For a more detailed discussion of monotone restriction, we refer to Kornhuber [22] or [24, pp. 74].

Inserting the search directions μ_l and the local defect obstacles $\underline{\psi}_l$, $\overline{\psi}_l$ as specified in (6.2) and (6.9) into (5.9) and (5.12), respectively, the *standard coarse grid correction* $\mathcal{C}_{J,\theta}^{\text{std}}$, $\theta \geq 0$, is obtained from (5.18).

Theorem 6.1. *For each fixed $\theta \geq 0$ and arbitrary initial iterate $u_J^0 \in \mathcal{G}_J$ the standard monotone multigrid method*

$$u_J^{\nu+1} = \mathcal{C}_{J,\theta}^{\text{std}} \bar{u}_J^\nu, \qquad \bar{u}_J^\nu = \mathcal{M}_{J,\theta} u_J^\nu \tag{6.10}$$

converges to the solution u_J of (3.2).

Proof. In the light of Proposition 5.1 the proof is an immediate consequence of Theorem 5.1. □

For $N = 2$, algorithm (6.10) turns out to be equivalent to the scalar version introduced by Kornhuber [22, 26]. In this case, asymptotic multigrid convergence rates are available for each fixed $\theta \geq 0$. The proofs are essentially based on linear convergence theory which can be applied to non-degenerate problems after a finite number of iteration steps ($\theta = 0$) and on higher order approximation of Newton linearization ($\theta > 0$). Extensions to $N > 2$ could rely on similar arguments and will be considered elsewhere. Even for scalar problems, upper bounds for the convergence rates which hold uniformly in J, still seem to be an open problem (see, however, Badea et al. [2] for results concerning a related approach).

The computational cost of the overall multigrid method (6.10) is of the same order as Gauß-Seidel smoothing $\mathcal{M}_J$, because the coarse grid correction $C_J^{\rm std}$ can be implemented as a classical V-cycle with complexity $\mathcal{O}(N^2 n_J)$: The successive solution of the local subproblems (5.11) amounts to projected polygonal Gauß-Seidel smoothing on each refinement level $k = 1, \ldots, J$. As local corrections on all levels have the form $z\lambda_p^{(k)} E_m$, it is convenient to choose the generating system

$$\Lambda_E^{(k)} = \{\lambda_p E_m \mid p \in \mathcal{N}_k,\ m = 0, \ldots, M\} \tag{6.11}$$

of $\mathcal{S}_k$ for the representation of bilinear forms, residuals and corrections in terms of stiffness matrices and vectors on each level k. Here, $E_0 \in \mathbb{R}^N$ is a vector not contained in the subspace spanned by the edges $E_1, \ldots, E_N$. For instance, one might choose $E_0 = e_1$. Stiffness matrices and residuals are then restricted in the canonical way. The coarse grid obstacles (6.9) are resulting from successive update and monotone restriction (6.8). Computation of local damping parameters ω_l according to (5.14) requires local Lipschitz constants L_l that can restricted like linear functionals, e.g., like residuals, and the sum of preceding corrections B_l are obtained by straightforward updates. Finally all corrections represented in terms of $\Lambda_E^{(k)}$ are prolongated in the canonical way. Multiple pre- or post-smoothing or W-cycles are carried out as usual. Within the framework of Section 5 those variants can be formulated by multiple occurrence of the same local coarse grid space $\mathcal{V}_l$.

6.2. **Truncated monotone multigrid methods.** The representation of low frequencies as incorporated in $C_J^{\rm std}$ is still suboptimal as compared with the linear unconstrained case. As an example, we consider a phase transition from pure phase i to pure phase j for $\theta = 0$. Assume that the corresponding diffuse interface is resolved by $\mathcal{T}_J$ but not by $\mathcal{T}_k$, i.e., that all triangles $t \in \mathcal{T}_k$ contain nodes from both phases i, j. Assume further that $\bar{u}_J^\nu$ is sufficiently accurate, i.e., $(\bar{u}_J^\nu)_j = 0$ on phase i and $(\bar{u}_J^\nu)_i = 0$ on phase j. Then any coarse grid correction of the form $z\lambda_p^{(k)} E_m$, $E_m = (e_i - e_j)$, applied to $\bar{u}_J^\nu$ must be zero, because it would violate either the constraint $(\bar{u}_J^\nu)_j \geq 0$ or $(\bar{u}_J^\nu)_i \geq 0$ otherwise. As a well-known remedy [22, 24, 26], we adapt the search directions $\mu_{l(s,m)} = \lambda_s E_m$ to the approximate discrete phases

$$\mathcal{N}^\bullet_{J,E_m}(\bar{u}_J^\nu) = \{p \in \mathcal{N}_J \mid \underline{\psi}_{E_m}^{(J)}(p) \cdot \overline{\psi}_{E_m}^{(J)}(p) = 0\}$$

by so-called *truncation*. More precisely, we set

$$\tilde{\mu}_{l(s(k,n),m)}^\nu = I_{\mathcal{S}^\nu_{J,E_m}} \circ \cdots \circ I_{\mathcal{S}^\nu_{k+1,E_m}} \lambda_{p_n}^{(k)} E_m, \qquad k = J-1, \ldots, 0,$$

where $I_{\mathcal{S}^\nu_{k,E_m}} : \mathcal{S}_J \to \mathcal{S}^\nu_{k,E_m}$ denotes nodal interpolation to the subspace

$$\mathcal{S}^\nu_{k,E_m} = \{v \in \mathcal{S}_k \mid v(p) = 0\ \forall p \in \mathcal{N}^\nu_{k,E_m}\} \subset \mathcal{S}_k$$

and $\mathcal{N}^\nu_{k,E_m} = \mathcal{N}_k \cap \mathcal{N}^\bullet_{J,m}(\bar{u}^\nu_J)$. This construction is motivated by the fact that the resulting search directions $\tilde{\mu}^\nu_l$ vanish on $\mathcal{N}^\bullet_{J,E_m}(\bar{u}^\nu_J)$, nevertheless have large support and are easy to implement (see below). Corresponding local obstacles $\underset{\sim}{\psi}^\nu_l$, $\overline{\widetilde{\psi}}^\nu_l$ are constructed in a similar way as mentioned above. We refer to Kornhuber [22] or [24, p. 81] for details.

Inserting $\tilde{\mu}^\nu_l$ and corresponding local obstacles $\underset{\sim}{\psi}^\nu_l$, $\overline{\widetilde{\psi}}^\nu_l$ into (5.9) and (5.12), respectively, the resulting *truncated coarse grid correction* $\mathcal{C}^{\mathrm{trc}}_{J,\theta}$, $\theta \geq 0$, is obtained from (5.18).

Theorem 6.2. *For each fixed $\theta \geq 0$ and arbitrary initial iterate $u^0_{J,\theta} \in \mathcal{G}_J$ the truncated monotone multigrid method*

$$u^{\nu+1}_J = \mathcal{C}^{\mathrm{trc}}_{J,\theta}\bar{u}^\nu_J, \qquad \bar{u}^\nu_J = \mathcal{M}_{J,\theta}u^\nu_J \tag{6.12}$$

converges to the solution u_J of (3.2).

Proof. The proof of Theorem 5.1 easily extends to coarse grid corrections $\mathcal{C}_J = \mathcal{C}^\nu_J$ varying in each iteration step (cf. [26] or [24, pp. 53]). Hence, the assertion follows from Theorem 5.1 and Proposition 5.1. □

As for the standard version (6.10), asymptotic multigrid convergence rates are available in the scalar case [22, 26].

The implementation of $\mathcal{C}^{\mathrm{trc}}_{J,\theta}$ can be obtained as a simple variant of the standard version: Let $p \in \mathcal{N}^\bullet_{J,E_m}(\bar{u}^\nu_J)$. Then the E_m-component at node p is set to zero in the restriction of the stiffness matrix and the residual. In turn the prolongation contributes zero to the E_m-component at node p in this case. Finally, the monotone restrictions (6.7) of the upper or lower defect obstacles are modified in such a way that all entries $v(p)$ are treated as ∞ or $-\infty$, respectively.

7. Numerical Experiments

We consider the vector-valued Allen-Cahn equation (2.5) with the interaction matrix C given in (2.4), the critical temperature $\theta_c = 1$ and $\varepsilon = 0.05$ for $N = 5$, 3 components in $d = 2$, 3 space dimensions. After discretization in time by the semi-implicit Euler method, the discrete spatial problems take the form (2.6) with bilinear form and right hand side given in (2.7) and (2.8), respectively. For simplicity, we use constant time steps and an heuristic strategy for adaptive mesh refinement. The refinement indicators η_T,

$$\eta_T = \sum_{p,q \in T \cap \mathcal{N}_J} |u(p) - u(q)|, \qquad T \in \mathcal{T}_J,$$

are intended to detect the steep gradients in the diffuse interface. A triangle T is refined or coarsened, if η_T is above or below certain thresholds. Starting from a given coarse triangulation $\mathcal{T}_0$ we obtain the refined triangulation $\mathcal{T}_J$ for the first time step by performing adaptive refinement geared to the initial condition $u(\cdot, 0)$. Refinement is stopped as soon as the maximal refinement depth J is reached. The subsequent triangulations are produced by adaptive refinement and coarsening in a similar way.

The spatial problems arising in each time step are solved iteratively using a V-cycle with 3 pre- and 3 post-smoothing steps of the truncated monotone multigrid method (MMG) introduced in Theorem 6.2. We select the threshold $L = 10^5$ for

the local Lipschitz constants (cf. (5.4) and Lemma 5.1). Similar behavior of our multigrid solvers was observed for $L = 10^i$, $i = 4, \dots, 8$. The initial iterate on the finest grid $\mathcal{T}_J$ is taken directly from the previous time step. In each time step we require an algebraic accuracy of 5%. More precisely, we use the stopping criterion

$$\|u_J^\nu - u_J^{\nu-1}\| \leq \sigma \, 0.05 \|u_J^\nu\| \tag{7.1}$$

with the (quite pessimistic) safety factor $\sigma = 10^{-3}$ accounting for the approximation of the exact algebraic error $\|u_J - u_J^{\nu-1}\|$ by the increment $\|u_J^\nu - u_J^{\nu-1}\|$.

For detailed investigations of the convergence speed we consider the averaged convergence rates

$$\rho = \sqrt[\nu_0]{\|u_J - u_J^{\nu_0}\| / \|u_J - u_J^0\|} \tag{7.2}$$

where ν_0 is chosen such that $\|u_J - u_J^{\nu_0}\| < 10^{-12}$. The "exact" solution u_J is precomputed by iterative solution up to machine accuracy.

7.1. **Example 1: N = 5 components and d = 2 space dimensions.** In our first example, we consider grain growth of $N = 5$ components in the unit square $\Omega = (0,1)^2$ in $d = 2$ space dimensions. The initial condition $u(\cdot, 0) \in \mathcal{G}$ is a randomly chosen superposition of circular grains each of which corresponds to a pure phase (see the upper left picture of Figure 3).

The coarse triangulation $\mathcal{T}_0$ is obtained by one uniform refinement step applied to Ω. More precisely, Ω is divided into two congruent triangles and each of them is then refined once. Adaptively refined triangulations $\mathcal{T}_J$ with maximal refinement depth $J = 7$ are are obtained as described above. Note that $h_J = 2^{-8} \approx \frac{1}{12}\varepsilon$ so that we can expect that the diffuse interface is well resolved by $\mathcal{T}_J$. We select the uniform time step size $\tau_0 = \frac{1}{2}\varepsilon$ and perform 100 time steps. About 5 - 8 multigrid iterations are required for each spatial problem to meet the stopping criterion (7.1). The evolution for the temperature $\theta = 0.1$ is shown in Figure 3. Comparing the approximate solutions for $t = 0$ with $t = \tau_0$ and for $t = 50\tau_0$ with $t = 100\tau_0$, the uniform time step size τ_0 seems to be too large in the beginning and too small in the end of the evolution. This motivates adaptive time stepping in the future.

In order to investigate the convergence properties of the underlying monotone multigrid solver MMG (cf. Theorem 6.2), we now consider the first spatial problem. The initial iterate on the underlying grid $\mathcal{T}_J$ ($J = 7$, 17293 nodes) is taken directly from the previous time step, i.e. the interpolation of $u(\cdot, 0)$ to $\mathcal{T}_J$ is taken as an initial iterate. In our first experiment, we investigate the robustness of the convergence speed with respect to the time step size τ. This seems to be relevant for a future combination of MMG with adaptivity in time. In the left picture of Figure 4, we show the averaged convergence rates ρ (cf. (7.2)) occurring for $\tau = 10^i \tau_0$, $i = -2, \dots, 5$. Observe that the convergence speed of MMG improves for smaller and also for larger time steps. We compare MMG with an *undamped version* as obtained by simply setting $\omega_l^\nu = 1$ in (5.18). As undamped MMG is always much faster it seems at first sight that damping is not necessary or even counterproductive.

Next we consider the robustness with respect to temperature θ. The right picture of Figure 4 shows the averaged convergence rates ρ over θ^{-1}. Observe that the convergence rates of MMG are uniformly bounded for all values of θ. As expected from theory, the damping parameters tend to one as $\theta \to 0$ so that MMG tends to the undamped version in this case. Though undamped MMG is mostly faster

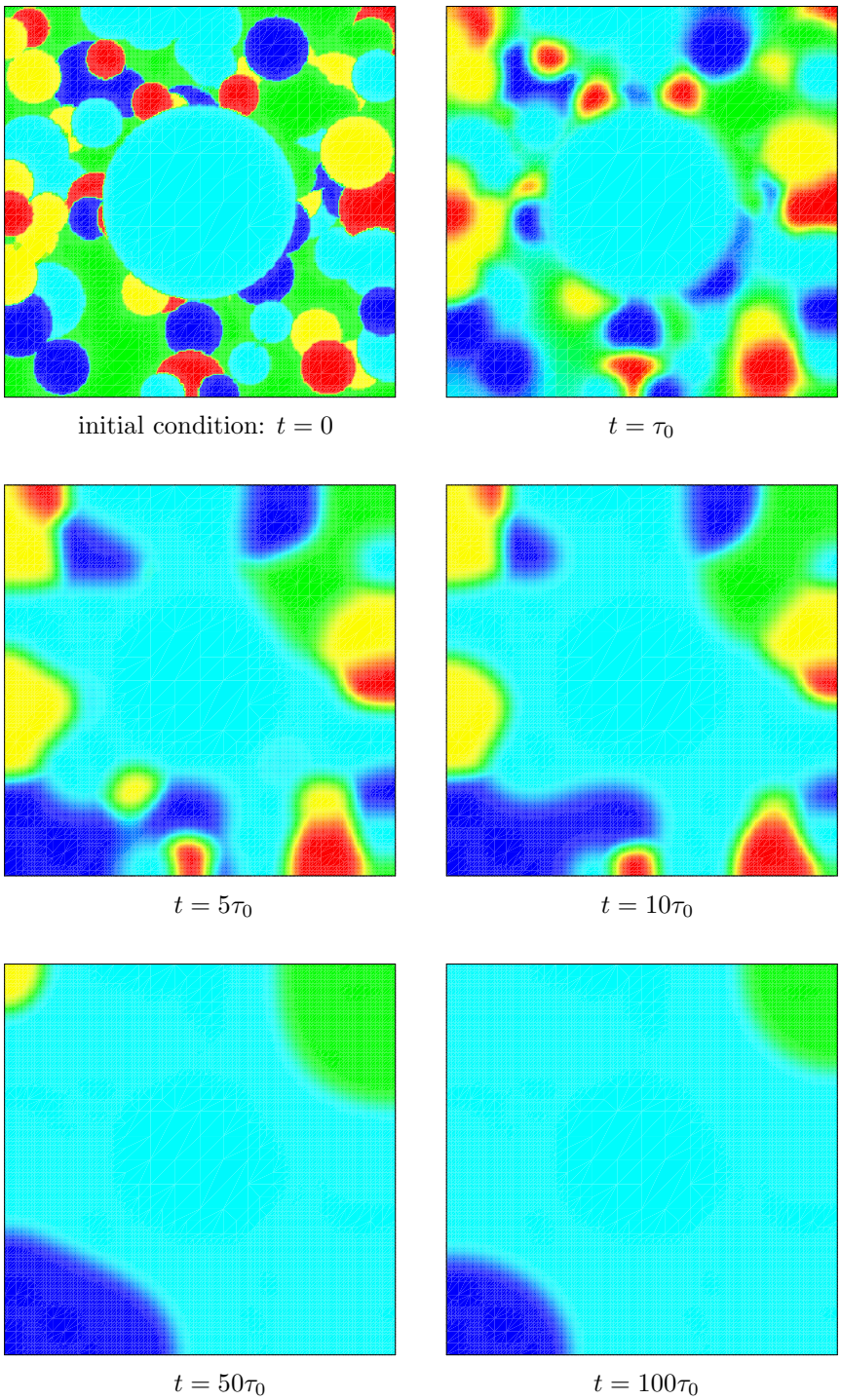

FIGURE 3. Evolution of the phases

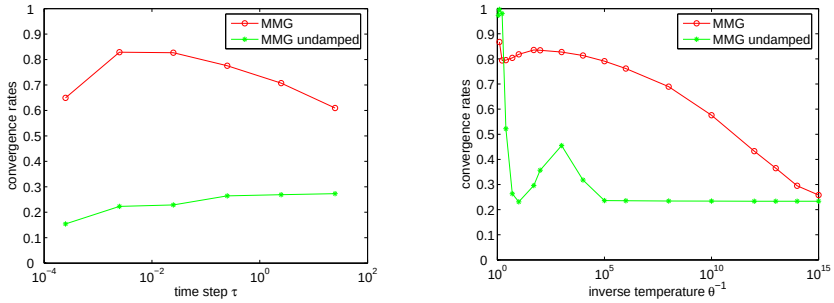

FIGURE 4. Robustness with respect to τ and θ ($N = 5$, $d = 2$)

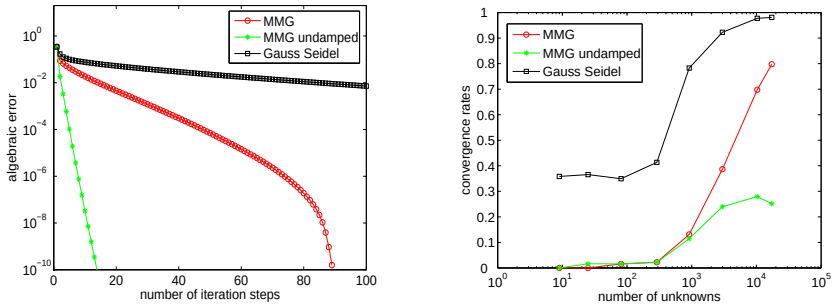

FIGURE 5. MMG and polygonal Gauß-Seidel ($N = 5$, $d = 2$)

than MMG, it fails to converge for sufficiently large θ, e.g., for $\theta = 0.95$. Hence, appropriate damping is necessary for convergence.

In order to get more detailed insight into the convergence behavior, we now consider the iteration history of our multigrid methods. Here we use the default parameters $\tau = \tau_0$ and $\theta = 0.1$ again. The left picture in Figure 5 is showing the algebraic error $\|u_J - u_J^\nu\|$ over the number of iteration steps $\nu = 1, \ldots, 100$. For a comparison, we included the polygonal Gauß-Seidel relaxation (cf. Theorem 4.1) represented by the square markers and the dashed line. The convergence behavior of MMG can be divided into a transient phase and an asymptotic phase. In the transient phase damping slows down the convergence speed while asymptotically the iterates become accurate enough to let the damping parameters tend to one. After 85 iteration steps the convergence speed of MMG and of the undamped version more or less coincide. Such a behavior is typical for monotone multigrid methods [23, 26] and could be remedied by better initial iterates as resulting, e.g., from nested iteration. Once the high frequency components of the error have been eliminated in the first iteration step, polygonal Gauß-Seidel relaxation is clearly outperformed by multigrid.

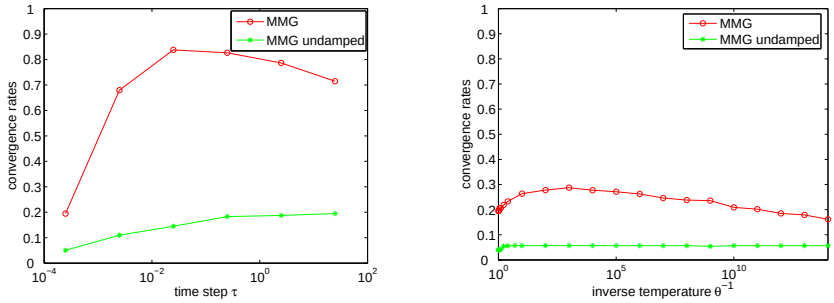

FIGURE 6. Robustness with respect to τ and θ ($N = 3$, $d = 3$)

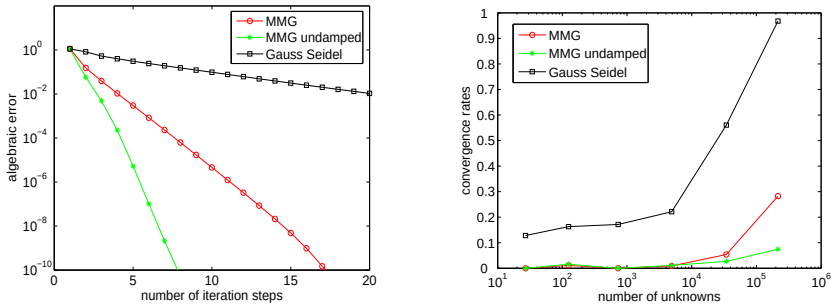

FIGURE 7. MMG and polygonal Gauß-Seidel ($N = 3$, $d = 3$)

The right picture in Figure 5 shows the averaged convergence rates over the number of nodes. As expected, Gauß-Seidel relaxation fails to converge reasonably well as soon as the diffuse interface is resolved, i.e. as soon as low frequency components of the error occur. Recall that the averaged convergence speed of MMG strongly depend on the quality of the initial iterate. Hence, increasing convergence rates with decreasing mesh size indicate a possible mesh dependence of the damping strategy. It seems that the convergence rates of the undamped MMG are about to saturate. This would be in agreement with related results in the scalar case [26].

7.2. **Example 2: $\mathbf{N = 3}$ components and $\mathbf{d = 3}$ space dimensions.** Now we study a related situation for $N = 3$ components in the unit cube $\Omega = (0, 1)^3$ in $d = 3$ space dimensions. As in the previous example, the initial condition $u(\cdot, 0) \in \mathcal{G}$ is resulting from the superposition of randomly chosen spherical grains. The coarse partition $\mathcal{T}_0$ is obtained by one uniform refinement step applied to Ω. Adaptively refined triangulations $\mathcal{T}_J$ with maximal refinement depth $J = 7$ are are obtained as described in the preliminary part of this section. In contrast to the first example, we now select the more realistic initial time step $\tau_0 = \frac{1}{200}\varepsilon$.

In order to illustrate the behavior of MMG (cf. Theorem 6.2) and its undamped version, we again concentrate on the first time step. The underlying partition $\mathcal{T}_J$ has 215502 nodes and the minimal mesh size is $h_J = 2^{-8} \approx \frac{1}{12}\varepsilon$. The initial iterate u_J^0 is taken from the previous time step, i.e. u_J^0 is just the interpolation of $u(\cdot, 0)$ to $\mathcal{T}_J$. The left picture of Figure 6 shows the averaged convergence rates ρ (cf. (7.2)) over the time step size $\tau = 10^i \tau_0$, $i = 0, \ldots, 5$. As in the previous example, the convergence rates of MMG seem to be uniformly bounded in τ and the undamped version of MMG (dotted line) is significantly faster.

Robustness with respect to temperature θ is illustrated in the right picture of Figure 6 showing the averaged convergence rates ρ over θ^{-1}. Both for MMG and the undamped version the averaged convergence rates are uniformly bounded in θ. In accordance with theory, the damping parameters tend to one as $\theta \to 0$ and therefore MMG tends to the undamped version in this case.

The left picture in Figure 7 displays the algebraic error $\|u_J - u_J^\nu\|$ of MMG, of the undamped version and of the polygonal Gauß-Seidel iteration (dashed line). Here we use the default parameters $\tau = \tau_0$ and $\theta = 0.1$ again. Observe that MMG is about to accelerate towards the asymptotic convergence speed as the final accuracy 10^{-10} is reached.

The right picture in Figure 5 shows the averaged convergence rates over the number of nodes. As usual, Gauß-Seidel relaxation deteriorates with increasing refinement. Though the diffuse interface is well resolved by $\mathcal{T}_J$, we do not observe saturation of the multigrid convergence rates. Again, undamped MMG performs best suggesting future research concerning even more sophisticated damping.

Acknowledgment . *Special thanks to G.M. Ziegler for providing a proof of Lemma 4.1 within a couple of days.*

References

[1] *S. Allen and J. Cahn. A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening.* Acta. Metall.*, 27:1084–1095, 1979.*

[2] *L. Badea, X.-Ch. Tai, and J. Wang. Convergence rate analysis of a multiplicative Schwarz method for variational inequalities.* SIAM J. Numer. Anal.*, 41:1052–1073, 2003.*

[3] *J.W. Barret, J.F. Blowey, and H. Garcke. On fully practical finite element approximations of degenerate Cahn-Hilliard systems.* Math. Model. Num. Anal. (M2AN)*, 35:713–748, 2001.*

[4] *J.W. Barrett and J.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.*

[5] *J. Bey.* Finite-Volumen- und Mehrgitterverfahren für elliptische Randwertprobleme*. Teubner, Stuttgart, 1998.*

[6] *J.F. Blowey and C.M. Elliott. The Cahn–Hilliard gradient theory for phase separation with non–smooth free energy II: Numerical analysis.* Euro. J. Appl. Math.*, 3:147–179, 1992.*

[7] *J.F. Blowey and C.M. Elliott. Curvature dependent phase boundary motion and parabolic double obstacle problems. In W.-M. Ni, L.A. Peletier, and J.L. Vazquez, editors,* Degenerate Diffusions*, pages 19–60. Springer, 1993.*

[8] *F.A. Bornemann, B. Erdmann, and R. Kornhuber. Adaptive multilevel methods in three space dimensions.* Int. J. Numer. Meth. Engrg., *36:3187–3203, 1993.*

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

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

[11] *I. Ekeland and R. Temam.* Convex Analysis and Variational Problems. *North–Holland, Amsterdam, 1976.*

[12] *D.J. Eyre. An unconditionally stable one-step scheme for gradient systems. Preprint, University of Utah, Salt Lake City, 1997.*

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

[14] *D. De Fontaine. An analysis of clustering and ordering in multicomponent solid solutions I. Stability criteria.* J. Phys. Chem. Solids, *33:287–310, 1972.*

[15] *H. Garcke, B. Nestler, and B. Stoth. On anisotropic order parameter models for multi–phase systems and their sharp interface limits.* Physica D, *115:87–108, 1998.*

[16] *H. Garcke, B. Nestler, and B. Stoth. Anisotropy in multi phase systems: A phase field approach.* Interf. Free Bound., *1:175–198, 1999.*

[17] *H. Garcke, B. Nestler, and B. Stoth. A multi phase concept: Numerical simulations of moving phase boundaries and multiple junctions.* SIAM J. Appl. Math., *60:295 – 315, 1999.*

[18] *H. Garcke and V. Styles. Bi-directional diffusion induced grain boundary motion with triple junctions.* Interf. Free Bound., *to appear.*

[19] *W. Hackbusch.* Multi-Grid Methods and Applications. *Springer, Berlin, 1985.*

[20] *T. Ilmanen. Convergence of the Allen-Cahn equation to Brakke's motion by mean curvature.* J. Differ. Geom., *38:417–461, 1993.*

[21] *D. Kessler, R. Nochetto, and A. Schmidt. A posteriori error control for thhe Allen-Cahn problem: circumventing Gronwal's inequality.* M2AN, *38:129–142, 2004.*

[22] *R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I.* Numer. Math, *69:167–184, 1994.*

[23] *R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I.* Numer. Math., *69:167 – 184, 1994.*

[24] *R. Kornhuber.* Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems. *Teubner, Stuttgart, 1997.*

[25] *R. Kornhuber. Monotone iterations for elliptic variational inequalities. In I. Athanasopoulos, G. Makrakis, and J.F. Rodrigues, editors,* Free Boundary Problems: Theory and Applications, *Research Notes in Mathematics, pages 335–343. Chapman & Hall, 1999.*

[26] *R. Kornhuber. On constrained Newton linearization and multigrid for variational inequalities.* Numer. Math., *91:699–721, 2002.*

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

[28] *R. Kornhuber, R. Krause, and V. Styles. Multigrid computations of diffusion induced grain boundary motion. In preparation.*

[29] *J. Mandel. A multilevel iterative method for symmetric, positive definite linear complementarity problems.* Appl. Math. Optimization*, 11:77–95, 1984.*

[30] *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.*

[31] *J. Xu. Iterative methods by space decomposition and subspace correction.* SIAM Review*, 34:581–613, 1992.*

[32] *G.M. Ziegler. Private communication. Technical report, 2003.*

Prof. Dr. Ralf Kornhuber, Freie Universität Berlin, Institut für Mathematik II, Arnimallee 2-6, D - 14195 Berlin, Germany

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

Prof. Dr. Rolf Krause, Universität Bonn, Institut für Numerische Simulation, Wegelerstrasse 6, D - 53115 Bonn, Germany

E-mail address: krause@iam.uni-bonn.de