A Phase Field Model for the Electromigration of Intergranular Voids

John W. Barrett[†] Harald Garcke[‡] Robert Nürnberg[†]

Abstract

We consider a fully practical finite element approximation of the nonlinear degenerate parabolic system

$$\gamma \, \tfrac{\partial u}{\partial t} - \nabla.(\, b(u)\, \nabla[w + \alpha\, \phi]\,) = 0\,, \quad \ell(\gamma)\, \tfrac{\partial v}{\partial t} + z = 0\,, \quad w = -\gamma\, \Delta u + \gamma^{-1}\, \Psi_{,u}(u,v)\,,$$
$$z = -\gamma\, \Delta v + \gamma^{-1}\, \Psi_{,v}(u,v)\,, \quad \nabla.(\, c(u)\, \nabla \phi) = 0$$

subject to initial conditions $u^0(\cdot) \in [-1,1]$ on u, $v^0(\cdot) \in [-\frac{1}{\sqrt{3}}\,(1+u^0(\cdot)), \frac{1}{\sqrt{3}}\,(1+u^0(\cdot))]$ on v and flux boundary conditions. Here $\gamma \in \mathbb{R}_{>0}$, $\alpha \in \mathbb{R}_{\geq 0}$, Ψ is a nonconvex obstacle potential, $\ell(\gamma) := \beta\,\gamma^2$ or $\beta\,\gamma$, with $\beta \in \mathbb{R}_{>0}$, and $c(u) := 1+u$, $b(u) := 1-u^2$ are degenerate coefficients. The degeneracy in b restricts $u(\cdot,\cdot) \in [-1,1]$. The above, in the limit $\gamma \to 0$, models the evolution of voids by surface diffusion and electromigration in an electrically conducting solid with a grain boundary. In addition to showing stability bounds for our approximation; we prove convergence, and hence existence of a solution to this nonlinear degenerate parabolic system in two space dimensions. Furthermore, an iterative scheme for solving the resulting nonlinear discrete system is introduced and analysed. Moreover, some numerical experiments are presented. Finally, in the Appendix we discuss the formal asymptotics leading to the sharp interface limit, as the interfacial parameter $\gamma \to 0$, of the above degenerate system.

Key words. void electromigration, phase field model, degenerate Cahn–Hilliard/ Allen–Cahn equation, fourth order degenerate parabolic equation, finite elements, convergence analysis, matched asymptotic expansions

AMS subject classifications. 65M60, 65M12, 35K55, 35K65, 35K35, 35Q60, 82C26, 65M50, 34E05

1 Introduction

Small voids that form in interconnect lines in microelectronic circuits can change their shape due to diffusion of atoms along the void surface. This surface diffusion is driven

[†]Department of Mathematics, Imperial College, London, SW7 2AZ, U.K.
[‡]NWF I – Mathematik, Universität Regensburg, 93040 Regensburg, Germany

by a diffusion potential which contains terms stemming from capillary effects, from an electrical potential and from elastic stresses. Elastic effects are neglected in this paper but can be incorporated (see Barrett, Garcke, and Nürnberg (2005)). The electric field can cause a so called "electron wind" force and this leads to the transport of atoms which results in migration of voids. In particular it can happen that voids which are initially contained in one grain (i.e. a region with a certain orientation of the crystal lattice) of the interconnect can get into contact with another grain (i.e. a region with a different lattice orientation). The modelling and computing of the interaction between voids and grain boundaries is the subject of this paper.

There are two approaches to model the evolution of coupled grain boundary/void systems. In the classical approach interfaces (i.e. the grain boundaries and the void surfaces) are modelled by a sharp interface, i.e. a hypersurface. A second more recent approach models interfaces by a diffusive interfacial layer. Let us first discuss roughly the sharp interface approach (for more details see Averbuch, Israeli, and Ravve (2003) and the references therein). Here a quite complicated system has to be studied. Along the void surface a fourth order parabolic equation has to be solved whereas at grain boundaries a second order parabolic equation holds. These equations are then coupled at triple junctions where boundary conditions such as angle conditions and flux balances have to hold. To approximate this problem numerically is quite difficult since the topology of the interfaces can change drastically (e.g. voids can attach to and detach from a grain boundary) and no satisfactory approach is known to us. For example in the paper by Averbuch, Israeli, and Ravve (2003) quite severe symmetry conditions are assumed.

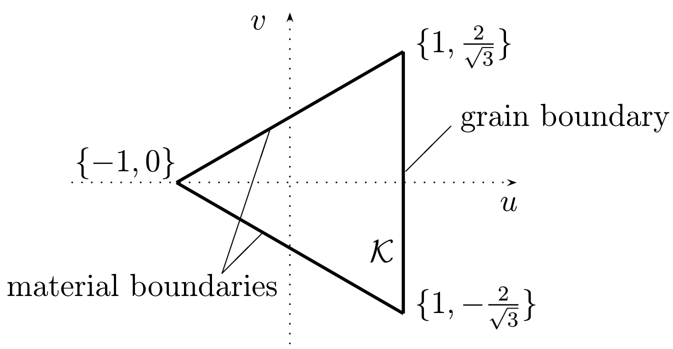

Figure 1: The $\{u, v\}$ space $\mathcal{K} = \triangle ABC$, where $A = \{-1, 0\}$, $B = \{1, -\frac{2}{\sqrt{3}}\}$, $C = \{1, \frac{2}{\sqrt{3}}\}$.

In this paper we therefore introduce a new model based on the idea of modelling the interface by a diffusive interfacial layer (our model will be a so called phase field model). We formulate a model for a system of two grains (we call them grain I and grain II) but natural generalizations are possible (see e.g. Garcke and Nestler (2000)). Each point in space either belongs to grain I, grain II or to the void. We now introduce a vector order parameter (or phase field) $\{u, v\}$ and the order parameter u describes whether we are in the void ($u = -1$) or not ($u = 1$). If $u = 1$ (i.e. in the material) then the order parameter v describes whether we are in grain I ($v = \frac{2}{\sqrt{3}}$) or in grain II ($v = -\frac{2}{\sqrt{3}}$). If $u = -1$ it makes no sense to distinguish between the grains and we set $v = 0$. This means the three values $A = \{-1, 0\}$, $B = \{1, -\frac{2}{\sqrt{3}}\}$, $C = \{1, \frac{2}{\sqrt{3}}\}$ for $\{u, v\}$ are relevant to distinguish between void, grain I and grain II (see Figure 1). We choose $\pm\frac{2}{\sqrt{3}}$ as values for v in the grains because this makes the triangle $\mathcal{K}$ with vertices A, B and C, see Figure 1, equilateral. Other values of v to distinguish the grains are possible, but these would complicate matters slightly. Our idea now is to generalize a phase field model introduced by Mahadevan and Bradley (1999) and studied later by Bhate, Kumar, and Bower (2000) and Barrett, Nürnberg, and Styles (2004) to include grain boundaries. We make use also of ideas by Cahn and Novick-Cohen (1994), Cahn and Novick-Cohen (1996), Cahn and Novick-Cohen (2000) who studied a degenerate Cahn–Hilliard/Allen–Cahn equation to study coupled surface diffusion and curvature flow.

The first step is to introduce the correct free energy: It is by now well established that a Ginzburg–Landau energy

$$\mathcal{E}(u, v) := \int_\Omega (\tfrac{\gamma}{2}\, |\nabla u|^2 + \tfrac{\gamma}{2}\, |\nabla v|^2 + \gamma^{-1}\Psi(u, v))\, \mathrm{d}x,$$

dependent on a vector-valued order parameter $\{u, v\}$, for a domain Ω, a parameter $\gamma > 0$, which is related to the interfacial thickness, and a nonconvex free energy density Ψ can model the interfacial energy of systems having different types of interfaces, see e.g. Baldo (1990), Bronsard and Reitich (1993), Garcke, Nestler, and Stoth (1998) and the references therein. To model the interfacial energy of our intergranular void system we need to assume that Ψ has three global minima at the points A, B and C. As mentioned above γ is related to the interfacial thickness. It can be shown with the help of formally matched asymptotic expansions or with Γ–convergence methods that $\mathcal{E}$ leads to a sharp interface free energy with surface energy densities (sometimes also called surface tensions)

$$\sigma^{ij} = 2 \inf_p \int_{-1}^{1} |p'(s)| \sqrt{\tfrac{1}{2}\, \Psi(p(s))}\, \mathrm{d}s, \qquad i, j \in \{A, B, C\}\,, \tag{1.1}$$

where the infimum is over all $p \in C^1([-1, 1], \mathbb{R}^2)$ with $p(-1) = i$ and $p(1) = j$. Again we refer to Baldo (1990), Bronsard and Reitich (1993), Garcke, Nestler, and Stoth (1998) for more details.

To formulate equations for the time evolution of the interfaces we introduce the potentials

$$w = \tfrac{\delta\mathcal{E}}{\delta u} = -\gamma\Delta u + \gamma^{-1}\, \Psi_{,u}(u, v) \quad \text{and} \quad z = \tfrac{\delta\mathcal{E}}{\delta v} = -\gamma\Delta v + \gamma^{-1}\, \Psi_{,v}(u, v)\,,$$

where $\frac{\delta\mathcal{E}}{\delta u}$ and $\frac{\delta\mathcal{E}}{\delta v}$ are the variational derivatives of $\mathcal{E}$ with respect to u and v respectively. The potential w is the chemical potential for the diffusion of atoms in the void-material interfacial layer and z acts at the driving force for the grain boundary motion. Taking into account that diffusion of atoms is also caused by the electrical field (see e.g. Mahadevan and Bradley (1999)) we propose the following set of evolution equations

$$\gamma\, \tfrac{\partial u}{\partial t} - \nabla\,.\, (b(u)\, \nabla[w + \alpha\, \phi]) = 0, \tag{1.2a}$$

$$\ell(\gamma)\, \tfrac{\partial v}{\partial t} + z = 0, \tag{1.2b}$$

which are coupled to the equation for the electric potential ϕ

$$\nabla . (c(u)\, \nabla\phi) = 0. \tag{1.2c}$$

Here α, $\ell(\gamma)$ are nonnegative coefficients and later we will use the scalings $\ell(\gamma) := \beta\,\gamma$ and $\ell(\gamma) := \beta\,\gamma^2$, where $\beta \in \mathbb{R}_{>0}$. The equation (1.2a) models diffusion in the void-material interfacial layer when we choose the degenerate mobility $b(u) := 1 - u^2$ and equation (1.2c) reduces to Laplace's equation in the material and is absent in the void if we take $c(u) := 1 + u$. As first shown by Cahn, Elliott, and Novick-Cohen (1996), we expect in the case that there is no coupling to a v-equation and $\alpha = 0$ that (1.2a) will model surface diffusion in the sharp interface limit. The equation (1.2b) is expected to describe the evolution of the grain boundary (see e.g. Garcke, Nestler, and Stoth (1998)).

The resulting system couples the degenerate Cahn–Hilliard equation (1.2a) to a non-degenerate Allen–Cahn equation (1.2b). We note that this is different to a similar set of equations introduced by Cahn and Novick-Cohen (1994), where the Allen–Cahn equation was also degenerate. For their system, which is a model for simultaneous order-disorder and phase separation, they showed that under an appropriate scaling and under certain assumptions on the geometry one obtains coupled mean curvature and surface diffusion in the sharp interface limit. We will show that we obtain a similar sharp interface limit also for our system with a nondegenerate Allen–Cahn equation. However, our sharp interface limit is different in some aspects and leads to some interesting new effects. We will discuss the formal asymptotics leading to the sharp interface model in the Appendix, §A. Here we will only outline the results. The domain Ω will split into regions where $\{u, v\}$ attains the values A, B and C and into interfacial layers separating these regions which have a thickness that is proportional to γ. Now depending on the scaling we derive different geometric evolution laws for the interfaces. For a detailed formulation of these laws we refer to the Appendix (see (A.41) and (A.42)). Here we only discuss the case when no coupling to the electric field is present. For the scaling $\ell(\gamma) := \beta\,\gamma^2$ we obtain that the interfaces which bound the void move by surface diffusion, i.e.

$$\mathcal{V} = -\tfrac{M\sigma}{4}\,\Delta_s\kappa\,,$$

where $\mathcal{V}$ is the normal velocity of the interface, κ is the (mean) curvature, Δ_s is the surface Laplacian, and M and σ are constants, whose precise definition can be found in §A. For a grain boundary we obtain that its mean curvature is zero. These evolution laws are coupled at triple junctions were angle conditions, flux conditions and continuity conditions have to hold.

If we scale the v-equation with $\ell(\gamma) := \beta\,\gamma$ we obtain for void boundaries an evolution law which combines surface diffusion and surface attachment limited kinetics (SALK). The evolution equation is

$$\mathcal{V} = \tfrac{M}{4}\,\Delta_s(-\sigma\,\kappa + \beta\,\omega\,\mathcal{V})\,,$$

where ω is a constant. This law has been derived by Taylor and Cahn (1994) and studied by Elliott and Garcke (1997). It links the fourth order surface diffusion flow to a second order flow, which is called motion by averaged mean curvature (see Taylor and Cahn

(1994) for details). For this second scaling one obtains the mean curvature flow

$$\beta\,\omega\,\mathcal{V} = \sigma\,\kappa$$

as the evolution law for grain boundaries. We remark that in Novick-Cohen (2000) and in Novick-Cohen and Hari (2005) a singular limit of an Allen-Cahn/Cahn-Hilliard system has also been analyzed. However, in our asymptotics we use a system different to theirs; firstly ours is not degenerate with respect to the Allen-Cahn part, and secondly we use a completely different scaling. This leads to new couplings of second and fourth order geometric evolution laws, which have not been studied before.

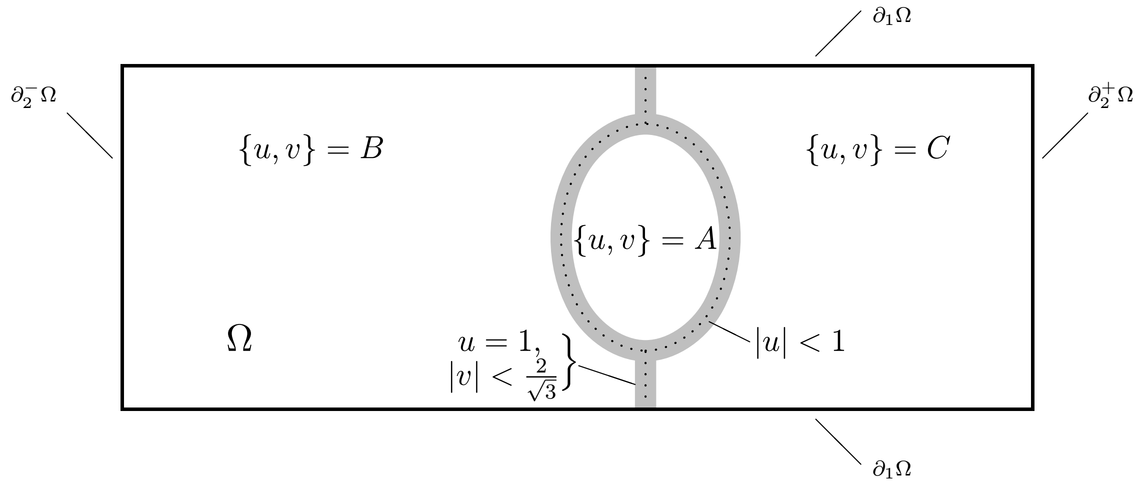

Figure 2: The order parameters for a typical intergranular void. Note that $A = \{-1, 0\}$, $B = \{1, -\frac{2}{\sqrt{3}}\}$ and $C = \{1, \frac{2}{\sqrt{3}}\}$.

In a recent paper by Barrett, Nürnberg, and Styles (2004), the following phase field model for void electromigration was considered:

$$\gamma\,\tfrac{\partial u}{\partial t} - \nabla\,.\,(b(u)\,\nabla[w + \alpha\,\phi]\,) = 0\,, \quad w = -\gamma\,\Delta u + \gamma^{-1}\,\overline{\Psi}'(u)\,, \quad \nabla\,.\,(c(u)\,\nabla\phi) = 0 \quad (1.3)$$

subject to an initial condition $u^0(\cdot) \in [-1, 1]$ on u and flux boundary conditions on all three equations. Here $u(\cdot, t) \in [-1, 1] \subset \mathbb{R}$ is the conserved order parameter, where at any time $t \in [0, T]$ $u(\cdot, t) = -1$ denotes the void and $u(\cdot, t) = 1$ denotes the conductor, while the void boundary is approximated by the $u(\cdot, t) = 0$ contour line inside the $|u(\cdot, t)| < 1$ interfacial region. In addition, $w(\cdot, t)$ is the chemical potential and $\overline{\Psi}$ is a non-smooth double obstacle potential. While, as in (1.2a–c), $\phi(\cdot, t)$ is the electric potential, $\gamma \in \mathbb{R}_{>0}$ is the interfacial parameter, $\alpha \in \mathbb{R}_{\geq 0}$ is a parameter denoting the relative strength of the electric field, and $b(u) := 1 - u^2$ and $c(u) := 1 + u$ are degenerate coefficients. The authors extended the technique of formal asymptotic expansions in Cahn, Elliott, and Novick-Cohen (1996) to show that the zero level sets of u_γ, the solution to (1.3) for a fixed $\gamma > 0$, converge as $\gamma \to 0$ to an interface, $\Gamma(t)$ with unit normal n_Γ, evolving with normal velocity

$$\mathcal{V} = -\tfrac{\pi^2}{16}\,\Delta_s\,\kappa + \alpha\,\tfrac{\pi}{4}\,\Delta_s\,\overline{\phi} \qquad \text{on } \Gamma(t)\,, \qquad (1.4a)$$

where κ is the curvature of $\Gamma(t)$ (positive if it is curved in the direction of n_Γ). The limiting electric potential, $\overline{\phi}(\cdot,t)$, satisfies

$$\Delta\overline{\phi} = 0 \quad \text{in } \Omega^+(t) := \Omega \setminus \overline{\Omega^-(t)}, \qquad \tfrac{\partial\overline{\phi}}{\partial n_\Gamma} = 0 \quad \text{on } \Gamma(t)\,, \tag{1.4b}$$

where $\Omega^-(t)$ is the void with boundary $\Gamma(t)$. For a discussion of different approaches to approximate (1.4a,b), see Barrett, Nürnberg, and Styles (2004). For further details on void electromigration see e.g. Xia, Bower, Suo, and Shih (1997), Cummings, Richardson, and Amar (2001) and the references therein.

The present paper extends the phase field model (1.3) to take into account grain boundaries. In summary the evolution of intergranular voids is described by the following nonlinear degenerate parabolic system:

(P) Find functions and u, v, w, z, $\phi : \Omega \times [0,T] \to \mathbb{R}$ such that $\{u(x,t), v(x,t)\} \in \mathcal{K}$ and for all $\{\eta_1(x,t), \eta_2(x,t)\} \in \mathcal{K}$

$$\gamma\,\tfrac{\partial u}{\partial t} - \nabla\,.\,(b(u)\,\nabla[w + \alpha\,\phi]\,) = 0 \qquad \text{in } \Omega_T, \tag{1.5a}$$

$$\ell(\gamma)\tfrac{\partial v}{\partial t} + z = 0 \qquad \text{in } \Omega_T, \tag{1.5b}$$

$$\begin{aligned}(-\gamma\,\Delta u + \gamma^{-1}\,\Psi_{,u}(u,v) - w)\,(\eta_1 - u)\\ + (-\gamma\,\Delta v + \gamma^{-1}\,\Psi_{,v}(u,v) - z)\,(\eta_2 - v) \geq 0\end{aligned} \qquad \text{in } \Omega_T, \tag{1.5c}$$

$$\{u(x,0), v(x,0)\} = \{u^0(x), v^0(x)\} \in \mathcal{K} \qquad \forall\ x \equiv (x_1,x_2)^T \in \Omega,$$

$$\tfrac{\partial u}{\partial \nu} = b(u)\,\tfrac{\partial[w+\alpha\,\phi]}{\partial \nu} = \tfrac{\partial v}{\partial \nu} = 0 \qquad \text{on } \partial\Omega \times (0,T]\,, \tag{1.5d}$$

$$\nabla\,.\,(c(u)\,\nabla\phi) = 0 \qquad \text{in } \Omega_T, \tag{1.5e}$$

$$c(u)\,\tfrac{\partial\phi}{\partial\nu} = 0 \quad \text{on } \partial_1\Omega \times (0,T], \quad c(u)\,\tfrac{\partial\phi}{\partial\nu} + \phi = g^\pm := x_1 \pm 2 \text{ on } \partial_2^\pm\Omega \times (0,T]\,; \tag{1.5f}$$

where $T > 0$ is a fixed positive time, $\Omega_T := \Omega \times (0,T]$ and $\Omega := (-L_1, L_1) \times (-L_2, L_2)$ is a rectangular domain in $\mathbb{R}^2$, representing the interconnect line, with boundary $\partial\Omega = \partial_1\Omega \cup \partial_2\Omega$, where $\partial_1\Omega \cap \partial_2\Omega = \emptyset$ and

$$\partial_2\Omega = \partial_2^-\Omega \cup \partial_2^+\Omega \quad \text{with} \quad \partial_2^\pm\Omega := \{\pm L_1\} \times [-L_2, L_2]\,,$$

and ν is the outward unit normal to $\partial\Omega$; see Figure 2. Hence $\partial_1\Omega$ is the insulated boundary of Ω, whilst the Robin boundary conditions on the ends $\partial_2^\pm\Omega$ model a uniform parallel electric field, as $L_1 \to \infty$. We note that one could alternatively model this with either (a) the Dirichlet condition $\phi = x_1$ or (b) the Neumann condition $c(u)\,\frac{\partial\phi}{\partial\nu} = \pm 2$ on $\partial_2^\pm\Omega$. However, in deriving energy bounds for (P) it is convenient to have weak boundary conditions; that is, Neumann or Robin conditions. The chosen Robin condition on $\partial_2^\pm\Omega$, (1.5f), has the added advantage that one obtains an immediate $L^2(\partial_2\Omega)$ bound on ϕ for the degenerate elliptic equation (1.5e). In (1.5a–d), γ, $\ell(\gamma) \in \mathbb{R}_{>0}$ and $\alpha \in \mathbb{R}_{\geq 0}$ are given constants and

$$\Psi(r,s) := \begin{cases} \Psi_0(r,s) & \text{if } \{r,s\} \in \mathcal{K}, \\ \infty & \text{if } \{r,s\} \notin \mathcal{K}, \end{cases} \qquad \text{with } \Psi_0 \in C^2(\overline{\mathcal{K}}), \tag{1.6}$$

is an obstacle free energy which restricts $\{u(\cdot,\cdot), v(\cdot,\cdot)\} \in \mathcal{K}$. Here we assume that $\Psi_0 \geq 0$ is a concave function with $\Psi_0(A) = \Psi_0(B) = \Psi_0(C) = 0$, e.g.

$$\Psi_0(r,s) := \tfrac{8}{9} - \tfrac{1}{2}\left[(r - \tfrac{1}{3})^2 + (1-\mu)\, s^2 + \tfrac{2}{3}\,\mu\,(r+1)\right], \tag{1.7}$$

where $\mu < 1$ is a parameter. In addition, we define the degenerate diffusion coefficients

$$c(s) := 1 + s, \qquad b(s) := 1 - s^2 = c(s)\, c(-s) \qquad \forall\ s \in -[1,1]. \tag{1.8}$$

The basic ingredients of our approach are some key energy estimates. Let us now briefly in a formal way describe how we obtain these estimates. Multiplying (1.5a) by $\gamma^{-1} w = \gamma^{-1} \frac{\delta \mathcal{E}}{\delta u}$ and (1.5b) by $[\ell(\gamma)]^{-1} z = [\ell(\gamma)]^{-1} \frac{\delta \mathcal{E}}{\delta v}$ yields, after integration of the sum of the two terms, the following free energy identity

$$\tfrac{d}{dt}\mathcal{E}(u,v) + \int_\Omega \left[\gamma^{-1} b(u)\, |\nabla w|^2 + [\ell(\gamma)]^{-1} z^2\right] \mathrm{d}x = -\gamma^{-1}\alpha \int_\Omega b(u)\, \nabla w \,.\, \nabla \phi \,\mathrm{d}x\,. \tag{1.9}$$

If we multiply (1.5e) by ϕ we can estimate $\int_\Omega b(u)\,\nabla w \,.\, \nabla\phi\,\mathrm{d}x$ and this enables us to control the right hand side of (1.9); leading to H^1-estimates in space for the phase field $\{u, v\}$. Relating F to c and G to b by the identities

$$c(s)\, F''(s) = 1 \qquad \text{and} \qquad b(s)\, G''(s) = 1\,, \tag{1.10}$$

and testing (1.5a) with $G'(u)$, (1.5b) with $-\Delta v$, and adding leads to

$$\begin{aligned} \tfrac{d}{dt} \int_\Omega \left[\gamma\, G(u) + \tfrac{l(\gamma)}{2}\, |\nabla v|^2\right] \mathrm{d}x + \gamma \int_\Omega \left[\, |\Delta u|^2 + |\Delta v|^2\right] \mathrm{d}x \\ = -\int_\Omega \nabla u \,.\, \nabla(\gamma^{-1}\Psi_{,u} + \alpha\,\phi)\,\mathrm{d}x - \gamma^{-1}\int_\Omega \nabla v \,.\, \nabla \Psi_{,v}\,\mathrm{d}x\,; \end{aligned} \tag{1.11}$$

where the term $\int_\Omega \nabla u \,.\, \nabla\phi\,\mathrm{d}x$ can be controlled if we test (1.5e) with $F'(u)$. This approach will lead to H^2-estimates in space for $\{u, v\}$. Discrete analogues of the above testing procedures will lead to the main a priori estimates for our finite element discretization (see Section 2). It is the goal of this paper to derive a finite element approximation of (P) that is consistent with these energy estimates, which then enables us to establish convergence in two space dimensions. In addition, in order to derive a discrete analogue of the necessary energy estimates we adapt a technique introduced in Zhornitskaya and Bertozzi (2000), and Grün and Rumpf (2000) for deriving a discrete entropy bound for the thin film equation. Finally, we note that a finite element approximation of the degenerate Allen–Cahn/Cahn–Hilliard system introduced by Cahn and Novick-Cohen (1994) can be found in Barrett and Blowey (2001). However due to the lack of a corresponding entropy bound, convergence of that approximation was only established for one space dimension.

This paper is organised as follows. In Section 2 we formulate a fully practical finite element approximation of the degenerate system (P) and derive important discrete energy estimates. In Section 3 we prove convergence, and hence existence of a solution to the system (P) in two space dimensions. In Section 4 we introduce and prove convergence of a "Gauss–Seidel type" iterative scheme for solving the nonlinear discrete system for the approximations of $\{u, v, w, z\}$ at each time level. In Section 5 we present some numerical experiments. Finally, in the Appendix we discuss the formal asymptotics leading to the sharp interface limit, as the interfacial parameter $\gamma \to 0$, of (P).

Notation and auxiliary results

For $D \subset \mathbb{R}$ or $D \subset \mathbb{R}^2$, we adopt the standard notation for Sobolev spaces, denoting the norm of $W^{m,q}(D)$ ($m \in \mathbb{N}$, $q \in [1,\infty]$) by $\|\cdot\|_{m,q,D}$ and the semi-norm by $|\cdot|_{m,q,D}$. We extend these norms and semi-norms in the natural way to the corresponding spaces of vector and matrix valued functions. For $q = 2$, $W^{m,2}(D)$ will be denoted by $H^m(D)$ with the associated norm and semi-norm written as, respectively, $\|\cdot\|_{m,D}$ and $|\cdot|_{m,D}$. For notational convenience, we drop the domain subscript on the above norms and semi-norms in the case $D \equiv \Omega$. Throughout $(\cdot,\cdot)$ denotes the standard L^2 inner product over Ω. In addition we define $\underline{m}(\Omega)$ as the measure of Ω and $\fint \eta := \frac{1}{\underline{m}(\Omega)}(\eta, 1)$ for all $\eta \in L^1(\Omega)$.

For later purposes, we recall the following compactness results. Let X_1, X_2 and X_3 be Banach spaces with a compact embedding $X_1 \hookrightarrow X_2$ and a continuous embedding $X_2 \hookrightarrow X_3$. Then we have the compact embeddings

$$\{\, \eta \in L^2(0,T;X_1) : \tfrac{\partial \eta}{\partial t} \in L^2(0,T;X_3) \,\} \hookrightarrow L^2(0,T;X_2) \tag{1.12a}$$

$$\text{and} \qquad \{\, \eta \in L^\infty(0,T;X_1) : \tfrac{\partial \eta}{\partial t} \in L^2(0,T;X_3) \,\} \hookrightarrow C([0,T];X_2). \tag{1.12b}$$

It is convenient to introduce the "inverse Laplacian" operator $\mathcal{G} : Y_1 \to Y_2$ such that

$$(\nabla[\mathcal{G}\eta_1], \nabla\eta_2) = \langle \eta_1, \eta_2 \rangle \quad \forall\, \eta_2 \in H^1(\Omega), \tag{1.13}$$

where $Y_1 := \{\eta \in (H^1(\Omega))' : \langle \eta, 1\rangle = 0\}$ and $Y_2 := \{\eta \in H^1(\Omega) : (\eta, 1) = 0\}$. Here and throughout $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $(H^1(\Omega))'$ and $H^1(\Omega)$. The well-posedness of $\mathcal{G}$ follows from the Lax–Milgram theorem and the Poincaré inequality

$$|\eta|_0 \leq C\,(\,|\eta|_1 + |(\eta,1)|\,) \quad \forall\, \eta \in H^1(\Omega). \tag{1.14}$$

We note also for future reference Young's inequality

$$r\,s \leq \tfrac{\theta}{2}\,r^2 + \tfrac{1}{2\theta}\,s^2 \quad \forall\, r,s \in \mathbb{R},\ \theta \in \mathbb{R}_{>0}\,. \tag{1.15}$$

Throughout C denotes a generic constant independent of h, τ and ε; the mesh and temporal discretization parameters and the regularization parameter. In addition $C(a_1, \cdots, a_I)$ denotes a constant depending on the arguments $\{a_i\}_{i=1}^{I}$.

2 Finite element approximation

We consider the finite element approximation of (P) under the following assumptions on the mesh:

(A) Let Ω be the rectangular domain $(-L_1, L_1) \times (-L_2, L_2)$. Let $\{\mathcal{T}^h\}_{h>0}$ be a quasi-uniform family of partitionings of Ω into disjoint open simplices σ with $h_\sigma := \text{diam}(\sigma)$ and $h := \max_{\sigma\in\mathcal{T}^h} h_\sigma$, so that $\overline{\Omega} = \cup_{\sigma\in\mathcal{T}^h}\overline{\sigma}$. In addition, it is assumed that all simplices $\sigma \in \mathcal{T}^h$ are right-angled.

We note that the right-angled simplices assumption is not a severe constraint, as there exist adaptive finite element codes that satisfy this requirement, see e.g. Schmidt and Siebert (2004).

Associated with $\mathcal{T}^h$ is the finite element space

$$S^h := \{\chi \in C(\overline{\Omega}) : \chi\,|_\sigma \text{ is linear } \forall\, \sigma \in \mathcal{T}^h\} \subset H^1(\Omega).$$

We introduce also

$$K := \{\, \{\eta_1, \eta_2\} \in [H^1(\Omega)]^2 : \{\eta_1(x), \eta_2(x)\} \in \mathcal{K} \ a.e. \text{ in } \Omega\,\} \quad \text{and} \quad K^h := K \cap [S^h]^2\,.$$

Let J be the set of nodes of $\mathcal{T}^h$ and $\{p_j\}_{j\in J}$ the coordinates of these nodes. Let $\{\chi_j\}_{j\in J}$ be the standard basis functions for S^h; that is $\chi_j \in S^h$ and $\chi_j(p_i) = \delta_{ij}$ for all $i, j \in J$. The right angle constraint on the partitioning is required for our approximations of $b(\cdot)$ and $c(\cdot)$, see (2.12a,b) and (2.8a,b) below, but one consequence is that

$$\int_\sigma \nabla\chi_i \,.\, \nabla\chi_j \,\mathrm{d}x \le 0 \quad i \neq j, \quad \forall\, \sigma \in \mathcal{T}^h. \tag{2.1}$$

We introduce $\pi^h : C(\overline{\Omega}) \to S^h$, the interpolation operator, such that $(\pi^h\eta)(p_j) = \eta(p_j)$ for all $j \in J$. A discrete semi-inner product on $C(\overline{\Omega})$ is then defined by

$$(\eta_1, \eta_2)^h := \int_\Omega \pi^h(\eta_1(x)\,\eta_2(x))\,\mathrm{d}x = \sum_{j\in J} m_j\,\eta_1(p_j)\,\eta_2(p_j), \tag{2.2}$$

where $m_j := (1, \chi_j) > 0$. The induced discrete semi-norm is then $|\eta|_h := [\,(\eta, \eta)^h\,]^{\frac{1}{2}}$, where $\eta \in C(\overline{\Omega})$. We introduce also the projection $Q^h : L^2(\Omega) \to S^h$ defined by

$$(Q^h\eta, \chi)^h = (\eta, \chi) \quad \forall\, \chi \in S^h. \tag{2.3}$$

On recalling (1.8) and (1.10), we then define functions F and G such that $c(\eta)\,\nabla[F'(\eta)] = \nabla\eta$ and $b(\eta)\,\nabla[G'(\eta)] = \nabla\eta$; that is,

$$F''(s) = \tfrac{1}{c(s)} = \tfrac{1}{1+s} \quad \text{and} \quad G''(s) = \tfrac{1}{b(s)} = \tfrac{1}{c(s)\,c(-s)} = \tfrac{1}{1-s^2}\,. \tag{2.4}$$

We take F, $G \in C^\infty(-1, 1)$, such that

$$F(s) = (1+s)\,\log(\tfrac{1+s}{2}) + (1-s) \qquad \text{and} \qquad G(s) = \tfrac{1}{2}\,[F(s) + F(-s)]\,; \tag{2.5}$$

and, for computational purposes, we replace F, G for any $\varepsilon \in (0, 1)$ by the regularized functions F_ε, $G_\varepsilon \in C^{2,1}(\mathbb{R})$ such that

$$F_\varepsilon(s) := \begin{cases} F(\varepsilon - 1) + (s - \varepsilon + 1)\,F'(\varepsilon - 1) + \frac{(s-\varepsilon+1)^2}{2}\,F''(\varepsilon - 1) & s \le \varepsilon - 1 \\ F(s) & s \ge \varepsilon - 1 \end{cases},$$
$$G_\varepsilon(s) := \tfrac{1}{2}\,[F_\varepsilon(s) + F_\varepsilon(-s)]\,. \tag{2.6}$$

We note for later purposes that for all $s \in [-1, 1]$

$$\tfrac{1}{2} \leq F_\varepsilon''(s) \leq \varepsilon^{-1}, \qquad \tfrac{1}{2}\, F_\varepsilon''(s) \leq G_\varepsilon''(s) \leq [\varepsilon\,(2-\varepsilon)]^{-1} \leq \varepsilon^{-1}\,. \tag{2.7}$$

Similarly to the approach in Zhornitskaya and Bertozzi (2000), Grün and Rumpf (2000), we introduce $\Lambda_\varepsilon : S^h \to [L^\infty(\Omega)]^{2\times 2}$ such that for all $\eta^h \in S^h$ and *a.e.* in Ω

$$\Lambda_\varepsilon(\eta^h) \text{ is symmetric and positive semi-definite,} \tag{2.8a}$$

$$\Lambda_\varepsilon(\eta^h)\, \nabla \pi^h[F_\varepsilon'(\eta^h)] = \nabla \eta^h. \tag{2.8b}$$

We now give the construction of Λ_ε. Let $\{e_i\}_{i=1}^2$ be the orthonormal vectors in $\mathbb{R}^2$, such that the $j^{\rm th}$ component of e_i is δ_{ij}, i, $j = 1 \to 2$. Given non-zero constants ζ_i, $i = 1 \to 2$; let $\widehat{\sigma}(\{\zeta_i\}_{i=1}^2)$ be the reference open simplex in $\mathbb{R}^2$ with vertices $\{\widehat{p}_i\}_{i=0}^2$, where $\widehat{p}_0$ is the origin and $\widehat{p}_i = \zeta_i\, e_i$, $i = 1 \to 2$. Given a $\sigma \in \mathcal{T}^h$ with vertices $\{p_{j_i}\}_{i=0}^2$, such that p_{j_0} is the right-angled vertex, then there exists a rotation matrix R_σ and non-zero constants $\{\zeta_i\}_{i=1}^2$ such that the mapping $\mathcal{R}_\sigma : \widehat{x} \in \mathbb{R}^2 \to p_{j_0} + R_\sigma \widehat{x} \in \mathbb{R}^2$ maps the vertex $\widehat{p}_i$ to p_{j_i}, $i = 0 \to 2$, and hence $\widehat{\sigma} \equiv \widehat{\sigma}(\{\zeta_i\}_{i=1}^2)$ to σ. For any $\eta^h \in S^h$, we then set

$$\Lambda_\varepsilon(\eta^h)\,|_\sigma := R_\sigma\, \widehat{\Lambda}_\varepsilon(\widehat{\eta}^h)\,|_{\widehat{\sigma}}\; R_\sigma^T, \tag{2.9}$$

where $\widehat{\eta}^h(\widehat{x}) \equiv \eta^h(\mathcal{R}_\sigma \widehat{x})$ for all $\widehat{x} \in \overline{\widehat{\sigma}}$ and $\widehat{\Lambda}_\varepsilon(\widehat{\eta}^h)\,|_{\widehat{\sigma}}$ is the 2×2 diagonal matrix with diagonal entries, $k = 1 \to 2$,

$$[\widehat{\Lambda}_\varepsilon(\widehat{\eta}^h)\,|_{\widehat{\sigma}}]_{kk} := \begin{cases} \frac{\widehat{\eta}^h(\widehat{p}_k) - \widehat{\eta}^h(\widehat{p}_0)}{F_\varepsilon'(\widehat{\eta}^h(\widehat{p}_k)) - F_\varepsilon'(\widehat{\eta}^h(\widehat{p}_0))} \equiv \frac{\eta^h(p_{j_k}) - \eta^h(p_{j_0})}{F_\varepsilon'(\eta^h(p_{j_k})) - F_\varepsilon'(\eta^h(p_{j_0}))} & \text{if } \eta^h(p_{j_k}) \neq \eta^h(p_{j_0}), \\ \frac{1}{F_\varepsilon''(\widehat{\eta}^h(\widehat{p}_0))} \equiv \frac{1}{F_\varepsilon''(\eta^h(p_{j_0}))} & \text{if } \eta^h(p_{j_k}) = \eta^h(p_{j_0}). \end{cases} \tag{2.10}$$

As $R_\sigma^T \equiv R_\sigma^{-1}$, $\nabla \eta^h \equiv R_\sigma\, \widehat{\nabla} \widehat{\eta}^h$, where $x \equiv (x_1, x_2)^T$, $\nabla \equiv (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2})^T$, $\widehat{x} \equiv (\widehat{x}_1, \widehat{x}_2)^T$ and $\widehat{\nabla} \equiv (\frac{\partial}{\partial \widehat{x}_1}, \frac{\partial}{\partial \widehat{x}_2})^T$, it easily follows that $\Lambda_\varepsilon(\eta^h)$ constructed in (2.9) and (2.10) satisfies (2.8a,b). It is this construction that requires the right angle constraint on the partitioning $\mathcal{T}^h$. Another consequence of this constraint is that

$$\int_{\widehat{\sigma}} \frac{\partial \widehat{\chi}_i}{\partial \widehat{x}_k} \frac{\partial \widehat{\chi}_j}{\partial \widehat{x}_k}\, {\rm d}\widehat{x} \leq 0 \quad i \neq j, \quad k = 1 \to 2, \quad \forall\, \sigma \in \mathcal{T}^h\,. \tag{2.11}$$

In a similar fashion we introduce $\Xi_\varepsilon : S^h \to [L^\infty(\Omega)]^{2\times 2}$ such that for all $\eta^h \in S^h$ and *a.e.* in Ω

$$\Xi_\varepsilon(\eta^h) \text{ is symmetric and positive semi-definite,} \tag{2.12a}$$

$$\Xi_\varepsilon(\eta^h)\, \nabla \pi^h[G_\varepsilon'(\eta^h)] = \nabla \eta^h; \tag{2.12b}$$

by extending the construction (2.9)–(2.10) for Λ_ε to Ξ_ε. Similarly to (2.1), it follows from (2.11), the above construction and (2.7) that for all $\eta^h \in S^h$

$$\int_\sigma \Xi_\varepsilon(\eta^h)\, \nabla \chi_i \,.\, \nabla \chi_j\, {\rm d}x \equiv \int_{\widehat{\sigma}} \widehat{\Xi}_\varepsilon(\widehat{\eta}^h)\, \widehat{\nabla} \widehat{\chi}_i \,.\, \widehat{\nabla} \widehat{\chi}_j\, {\rm d}\widehat{x} \leq 0 \quad i \neq j, \quad \forall\, \sigma \in \mathcal{T}^h\,. \tag{2.13}$$

Obviously, the above result also holds with Ξ_ε replaced by Λ_ε.

In addition to $\mathcal{T}^h$, let $0 = t_0 < t_1 < \ldots < t_{N-1} < t_N = T$ be a partitioning of $[0,T]$ into possibly variable time steps $\tau_n := t_n - t_{n-1}$, $n = 1 \to N$. We set $\tau := \max_{n=1\to N} \tau_n$. For any given $\varepsilon \in (0,1)$, we then consider the following fully practical finite element approximation of (P):

($\mathrm{P}_\varepsilon^{h,\tau}$) For $n \geq 1$ find $\{\Phi_\varepsilon^n, U_\varepsilon^n, V_\varepsilon^n, W_\varepsilon^n, Z_\varepsilon^n\} \in [S^h]^5$ such that $\{U_\varepsilon^n, V_\varepsilon^n\} \in K^h$ and

$$(\Lambda_\varepsilon(U_\varepsilon^{n-1})\,\nabla\Phi_\varepsilon^n, \nabla\chi) + \int_{\partial_2\Omega} \Phi_\varepsilon^n\,\chi\,\mathrm{d}s = \int_{\partial_2\Omega} g\,\chi\,\mathrm{d}s \quad \forall\,\chi \in S^h, \tag{2.14a}$$

$$\gamma\left(\tfrac{U_\varepsilon^n - U_\varepsilon^{n-1}}{\tau_n}, \chi\right)^h + (\Xi_\varepsilon(U_\varepsilon^{n-1})\,\nabla[W_\varepsilon^n + \alpha\,\Phi_\varepsilon^n], \nabla\chi) = 0 \quad \forall\,\chi \in S^h, \tag{2.14b}$$

$$\ell(\gamma)\left(\tfrac{V_\varepsilon^n - V_\varepsilon^{n-1}}{\tau_n}, \chi\right)^h + (Z_\varepsilon^n, \chi)^h = 0 \quad \forall\,\chi \in S^h, \tag{2.14c}$$

$$\begin{aligned}\gamma\,(\nabla U_\varepsilon^n, \nabla[\chi_1 - U_\varepsilon^n]) + \gamma\,(\nabla V_\varepsilon^n, \nabla[\chi_2 - V_\varepsilon^n]) &\geq (W_\varepsilon^n - \gamma^{-1}\,\Psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \chi_1 - U_\varepsilon^n)^h \\ &+ (Z_\varepsilon^n - \gamma^{-1}\,\Psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \chi_2 - V_\varepsilon^n)^h \quad \forall\,\{\chi_1,\chi_2\} \in K^h,\end{aligned} \tag{2.14d}$$

where $g := g^\pm \equiv \pm(2 + L_1)$ on $\partial_2^\pm\Omega$ and $\{U_\varepsilon^0, V_\varepsilon^0\} \in K^h$ is an approximation of $\{u^0, v^0\} \in K$, e.g. $U_\varepsilon^0 \equiv \pi^h u^0$, if $u^0 \in C(\overline{\Omega})$; and similarly V_ε^0.

Below we recall some well-known results concerning S^h for any $\sigma \in \mathcal{T}^h$, $\chi, \eta^h \in S^h$, $m \in \{0,1\}$, $p \in [1,\infty]$ and $q \in (2,\infty]$:

$$|(I - \pi^h)\eta|_m \leq C\,h^{2-m}\,|\eta|_2 \qquad \forall\,\eta \in H^2(\Omega)\,; \tag{2.15}$$

$$|(I - \pi^h)\eta|_{m,q} \leq C\,h^{1-m}\,|\eta|_{1,q} \qquad \forall\,\eta \in W^{1,q}(\Omega)\,; \tag{2.16}$$

$$\int_\sigma \chi^2\,\mathrm{d}x \leq \int_\sigma \pi^h[\chi^2]\,\mathrm{d}x \leq 4\int_\sigma \chi^2\,\mathrm{d}x\,; \tag{2.17}$$

$$|\int_\sigma (I - \pi^h)(\chi\,\eta^h)\,\mathrm{d}x| \leq |(I - \pi^h)(\chi\,\eta^h)|_{0,1,\sigma} \;\leq\; C\,h^{1+m}\,|\chi|_{m,\sigma}\,|\eta^h|_{1,\sigma}\,. \tag{2.18}$$

Finally, as we have a quasi-uniform family of partitionings, it holds that

$$|(I - Q^h)\eta|_m \leq C\,h^{1-m}\,|\eta|_1 \quad \forall\,\eta \in H^1(\Omega). \tag{2.19}$$

We define $Y_2^h := \{\eta^h \in S^h : (\eta^h, 1) = 0\}$ and introduce the "discrete Laplacian" operator $\Delta^h : S^h \to Y_2^h$ such that

$$(\Delta^h\eta^h, \chi)^h = -(\nabla\eta^h, \nabla\chi) \qquad \forall\,\chi \in S^h. \tag{2.20}$$

Next we introduce for all $\varepsilon \in (0,1)$, $c_\varepsilon : [-1,1] \to [\varepsilon, 2]$ and $b_\varepsilon : [-1,1] \to [\varepsilon\,(2-\varepsilon), 1]$ defined, on recalling (2.4), (2.6) and (2.7), by

$$c_\varepsilon(s) := \tfrac{1}{F_\varepsilon''(s)} \geq \tfrac{1}{F''(s)} = c(s)\,, \qquad b_\varepsilon(s) := \tfrac{1}{G_\varepsilon''(s)} \geq \tfrac{1}{G''(s)} = b(s)\,. \tag{2.21}$$

Then the following two lemmas follow immediately from the construction of Λ_ε and Ξ_ε, see Barrett, Nürnberg, and Styles (2004, Lemmas 2.2 and 2.3) for details.

LEMMA. 2.1 *Let the assumptions* (A) *hold. Then for any given* $\varepsilon \in (0,1)$ *the functions* Λ_ε, $\Xi_\varepsilon : S^h \to [L^\infty(\Omega)]^{2\times 2}$ *satisfy for all* $\eta^h \in K^h$, $\xi \in \mathbb{R}^2$ *and for all* $\sigma \in \mathcal{T}^h$

$$\varepsilon\, \xi^T \xi \leq \min_{x\in\overline{\sigma}} c_\varepsilon(\eta^h(x))\, \xi^T\xi \leq \xi^T\, \Lambda_\varepsilon(\eta^h)\,|_\sigma\, \xi \leq \max_{x\in\overline{\sigma}} c_\varepsilon(\eta^h(x))\, \xi^T\xi \leq 2\, \xi^T\xi\,, \tag{2.22a}$$

$$\varepsilon\,(2-\varepsilon)\, \xi^T \xi \leq \min_{x\in\overline{\sigma}} b_\varepsilon(\eta^h(x))\, \xi^T\xi \leq \xi^T\, \Xi_\varepsilon(\eta^h)\,|_\sigma\, \xi \leq \max_{x\in\overline{\sigma}} b_\varepsilon(\eta^h(x))\, \xi^T\xi \leq \xi^T\xi\,, \tag{2.22b}$$

$$\xi^T\, \Xi_\varepsilon(\eta^h)\,|_\sigma\, \xi \leq 2\, \xi^T\, \Lambda_\varepsilon(\eta^h)\,|_\sigma\, \xi\,. \tag{2.22c}$$

LEMMA. 2.2 *Let the assumptions* (A) *hold and let* $\|\cdot\|$ *denote the spectral norm on* $\mathbb{R}^{2\times 2}$. *Then for any given* $\varepsilon \in (0,1)$ *the functions* $\Lambda_\varepsilon : S^h \to [L^\infty(\Omega)]^{2\times 2}$ *and* $\Xi_\varepsilon : S^h \to [L^\infty(\Omega)]^{2\times 2}$ *are such that for all* $\eta^h \in K^h$ *and for all* $\sigma \in \mathcal{T}^h$

$$\max_{x\in\sigma} \|\{\Lambda_\varepsilon(\eta^h) - c_\varepsilon(\eta^h)\,\mathcal{I}\}(x)\| \leq h_\sigma\, |\nabla[c_\varepsilon(\eta^h)]\,|_{0,\infty,\sigma} \leq h_\sigma\, |\nabla \eta^h\,|_\sigma\,|\,, \tag{2.23a}$$

$$\max_{x\in\sigma} \|\{\Xi_\varepsilon(\eta^h) - b_\varepsilon(\eta^h)\,\mathcal{I}\}(x)\| \leq h_\sigma\, |\nabla[b_\varepsilon(\eta^h)]\,|_{0,\infty,\sigma} \leq 2\, h_\sigma\, |\nabla \eta^h\,|_\sigma\,|\,, \tag{2.23b}$$

where $\mathcal{I}$ *is the* 2×2 *identity matrix.*

We now derive discrete analogues of the energy estimates (1.9) and (1.11).

LEMMA. 2.3 *Let the assumptions* (A) *hold and* $\{U^{n-1}_\varepsilon, V^{n-1}_\varepsilon\} \in K^h$. *Then for all* $\varepsilon \in (0,1)$ *and for all* h, $\tau_n > 0$ *there exists a solution* $\{\Phi^n_\varepsilon, U^n_\varepsilon, V^n_\varepsilon, W^n_\varepsilon, Z^n_\varepsilon\}$ *to the* n*-th step of* $(\mathrm{P}^{h,\tau}_\varepsilon)$ *with* $⨍ U^n_\varepsilon = ⨍ U^{n-1}_\varepsilon$. $\{\Phi^n_\varepsilon, U^n_\varepsilon, V^n_\varepsilon, Z^n_\varepsilon\}$ *is unique. In addition,* W^n_ε *is unique if there exists* $j \in J$ *such that* $U^n_\varepsilon(p_j) \in (-1,1)$. *Moreover, it holds that*

$$(\Lambda_\varepsilon(U^{n-1}_\varepsilon)\, \nabla\Phi^n_\varepsilon, \nabla\Phi^n_\varepsilon) + \tfrac{1}{2}\, |\Phi^n_\varepsilon|^2_{0,\partial_2\Omega} \leq \tfrac{1}{2}\, |g|^2_{0,\partial_2\Omega}, \tag{2.24}$$

$$|(\nabla\Phi^n_\varepsilon, \nabla U^{n-1}_\varepsilon)| \leq 2\, |g|_{0,\partial_2\Omega}\, |\pi^h[F'_\varepsilon(U^{n-1}_\varepsilon)]|_{0,\partial_2\Omega} \tag{2.25}$$

and

$$\begin{aligned}\mathcal{E}(U^n_\varepsilon, V^n_\varepsilon) &+ \tfrac{1}{2}\left[\gamma\, |U^n_\varepsilon - U^{n-1}_\varepsilon|^2_1 + \gamma\, |V^n_\varepsilon - V^{n-1}_\varepsilon|^2_1\right] + [\ell(\gamma)]^{-1}\, \tau_n\, |Z^n_\varepsilon|^2_h \\ &+ \tfrac{1}{2}\, \gamma^{-1}\, \tau_n\, |\, [\Xi_\varepsilon(U^{n-1}_\varepsilon)]^{\frac{1}{2}}\, \nabla W^n_\varepsilon|^2_0 \leq \mathcal{E}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon) + \tfrac{1}{2}\, \alpha^2\, \gamma^{-1}\, \tau_n\, |g|^2_{0,\partial_2\Omega}\,, \end{aligned} \tag{2.26a}$$

where

$$\mathcal{E}(U^n_\varepsilon, V^n_\varepsilon) := \tfrac{1}{2}\,[\,\gamma\, |U^n_\varepsilon|^2_1 + \gamma\, |V^n_\varepsilon|^2_1\,] + \gamma^{-1}\,(\Psi(U^n_\varepsilon, V^n_\varepsilon), 1)^h\,. \tag{2.26b}$$

Furthermore, it holds that

$$\begin{aligned}\gamma\,(G_\varepsilon(U^n_\varepsilon) - G_\varepsilon(U^{n-1}_\varepsilon), 1)^h &+ \gamma\, \tau_n\, |\Delta^h U^n_\varepsilon|^2_h + \tfrac{\ell(\gamma)}{2}\,\left[\,|V^n_\varepsilon|^2_1 + |V^n_\varepsilon - V^{n-1}_\varepsilon|^2_1 - |V^{n-1}_\varepsilon|^2_1\,\right] \\ &+ \gamma\, \tau_n\, |\Delta^h V^n_\varepsilon|^2_h \\ \leq \varepsilon^{-1}\, \gamma\, |U^n_\varepsilon - U^{n-1}_\varepsilon|^2_h &+ \tau_n\,[\,(\nabla W^n_\varepsilon, \nabla[U^n_\varepsilon - U^{n-1}_\varepsilon]\,) - \alpha\,(\nabla\Phi^n_\varepsilon, \nabla U^{n-1}_\varepsilon) \\ &- \gamma^{-1}\,(\nabla\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \nabla U^n_\varepsilon) - \gamma^{-1}\,(\nabla\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \nabla V^n_\varepsilon)\,]\,. \end{aligned} \tag{2.27}$$

Proof. Deriving the existence of a unique solution $\Phi^n_\varepsilon \in S^h$, (2.24) and (2.25) is straightforward and can be found in Barrett, Nürnberg, and Styles (2004, Lemma 2.4). In order to prove existence of a solution $\{\{U^n_\varepsilon, V^n_\varepsilon\}, W^n_\varepsilon, Z^n_\varepsilon\} \in K^h \times [S^h]^2$ to (2.14b–d), we introduce, similarly to (1.13), for $q^h \in K^h$ the discrete anisotropic Green's operator $\mathcal{G}^h_{q^h} : Y^h_2 \to Y^h_2$ such that

$$(\Xi_\varepsilon(q^h)\,\nabla[\mathcal{G}^h_{q^h}\eta^h], \nabla\chi) = (\eta^h, \chi)^h \qquad \forall\, \chi \in S^h. \tag{2.28}$$

It follows immediately from (2.22b) and (1.14) that $\mathcal{G}^h_{q^h}$ is well-posed. It follows from (2.14b) and (2.28) that

$$W^n_\varepsilon \equiv -\alpha\,\Phi^n_\varepsilon - \gamma\,\mathcal{G}^h_{U^{n-1}_\varepsilon}[\tfrac{U^n_\varepsilon - U^{n-1}_\varepsilon}{\tau_n}] + \lambda^n, \tag{2.29}$$

where $\lambda^n \in \mathbb{R}$. Hence (2.14b–d) can be restated as: Find $\{U^n_\varepsilon, V_\varepsilon\} \in K^h(U^{n-1}_\varepsilon) := \{\{\chi_1, \chi_2\} \in K^h : \chi_1 - U^{n-1}_\varepsilon \in Y^h_2\}$ and a Lagrange multiplier $\lambda^n \in \mathbb{R}$ such that for all $\{\chi_1, \chi_2\} \in K^h$

$$\begin{aligned}
\gamma\,(\nabla U^n_\varepsilon, \nabla(\chi_1 - U^n_\varepsilon)) + \gamma\,(\nabla V^n_\varepsilon, \nabla(\chi_2 - V^n_\varepsilon)) + \gamma\,(\mathcal{G}^h_{U^{n-1}_\varepsilon}[\tfrac{U^n_\varepsilon - U^{n-1}_\varepsilon}{\tau_n}], \chi_1 - U^n_\varepsilon)^h \\
+ \ell(\gamma)\,(\tfrac{V^n_\varepsilon - V^{n-1}_\varepsilon}{\tau_n}, \chi_2 - V^n_\varepsilon)^h \\
\geq (-\gamma^{-1}\,\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon) - \alpha\,\Phi^n_\varepsilon + \lambda^n, \chi_1 - U^n_\varepsilon)^h - \gamma^{-1}\,(\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \chi_2 - V^n_\varepsilon)^h.
\end{aligned} \tag{2.30}$$

It follows from (2.30) that $\{U^n_\varepsilon, V^n_\varepsilon\} \in K^h(U^{n-1}_\varepsilon)$ is such that for all $\{\chi_1, \chi_2\} \in K^h(U^{n-1}_\varepsilon)$

$$\begin{aligned}
\gamma\,(\nabla U^n_\varepsilon, \nabla(\chi_1 - U^n_\varepsilon)) + \gamma\,(\nabla V^n_\varepsilon, \nabla(\chi_2 - V^n_\varepsilon)) + \gamma\,(\mathcal{G}^h_{U^{n-1}_\varepsilon}[\tfrac{U^n_\varepsilon - U^{n-1}_\varepsilon}{\tau_n}], \chi_1 - U^n_\varepsilon)^h \\
+ \ell(\gamma)\,(\tfrac{V^n_\varepsilon - V^{n-1}_\varepsilon}{\tau_n}, \chi_2 - V^n_\varepsilon)^h \\
\geq (-\gamma^{-1}\,\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon) - \alpha\,\Phi^n_\varepsilon, \chi_1 - U^n_\varepsilon)^h - \gamma^{-1}\,(\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \chi_2 - V^n_\varepsilon)^h.
\end{aligned} \tag{2.31}$$

There exists a unique $\{U^n_\varepsilon, V^n_\varepsilon\} \in K^h(U^{n-1}_\varepsilon)$ solving (2.31) since, on noting (2.28), this is the Euler–Lagrange variational inequality of the strictly convex minimization problem

$$\begin{aligned}
\min_{\{\eta^h_1, \eta^h_2\} \in K^h(U^{n-1}_\varepsilon)} \Big\{ \tfrac{\gamma}{2}\,|\eta^h_1|^2_1 + \tfrac{\gamma}{2}\,|\eta^h_2|^2_1 + \tfrac{\gamma}{2\,\tau_n}\,|[\Xi_\varepsilon(U^{n-1}_\varepsilon)]^{\frac{1}{2}}\,\nabla\mathcal{G}^h_{U^{n-1}_\varepsilon}(\eta^h_1 - U^{n-1}_\varepsilon)|^2_0 \\
+ \tfrac{\ell(\gamma)}{2\,\tau_n}\,|\eta^h_2 - V^{n-1}_\varepsilon|^2_0 + (\gamma^{-1}\,\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon) + \alpha\,\Phi^n_\varepsilon, \eta^h_1)^h + (\gamma^{-1}\,\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \eta^h_2)^h \Big\}.
\end{aligned}$$

Existence of the Lagrange multiplier λ^n in (2.30) then follows from standard optimisation theory, see e.g. Ciarlet (1988). Hence we have existence of a solution $\{\{U^n_\varepsilon, V^n_\varepsilon\}, W^n_\varepsilon, Z^n_\varepsilon\} \in K^h \times [S^h]^2$ to (2.14b–d). If $|U^n_\varepsilon(p_j)| < 1$ for some $j \in J$ then $\pi^h[1 - (U^n_\varepsilon)^2] \not\equiv 0$ and choosing $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon \pm \delta\pi^h[1 - (U^n_\varepsilon)^2], V^n_\varepsilon \pm \frac{\delta}{\sqrt{3}}\pi^h[1 - (U^n_\varepsilon)^2]\}$ in (2.30) for $\delta > 0$ sufficiently small yields uniqueness of λ^n and, on noting (2.29), uniqueness of W_ε. Furthermore, choosing $\chi \equiv 1$ in (2.14b) yields $⨍ U^n_\varepsilon = ⨍ U^{n-1}_\varepsilon$.

Choosing $\chi \equiv W_\varepsilon^n$ in (2.14b), $\chi \equiv Z_\varepsilon^n$ in (2.14c) and $\{\chi_1, \chi_2\} \equiv \{U_\varepsilon^{n-1}, V_\varepsilon^{n-1}\}$ in (2.14d) yields that

$$\gamma\,(U_\varepsilon^n - U_\varepsilon^{n-1}, W_\varepsilon^n)^h + \tau_n\,(\Xi_\varepsilon(U_\varepsilon^{n-1})\,\nabla[W_\varepsilon^n + \alpha\,\Phi_\varepsilon^n], \nabla W_\varepsilon^n) = 0, \tag{2.32a}$$

$$\ell(\gamma)\,(V_\varepsilon^n - V_\varepsilon^{n-1}, Z_\varepsilon^n)^h + \tau_n\,(Z_\varepsilon^n, Z_\varepsilon^n)^h = 0, \tag{2.32b}$$

$$\begin{aligned}\gamma\,(\nabla U_\varepsilon^n, \nabla[U_\varepsilon^{n-1} - U_\varepsilon^n]) + \gamma\,(\nabla V_\varepsilon^n, \nabla[V_\varepsilon^{n-1} - V_\varepsilon^n])&\\ \geq (W_\varepsilon^n - \gamma^{-1}\,\Psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), U_\varepsilon^{n-1} - U_\varepsilon^n)^h&\\ + (Z_\varepsilon^n - \gamma^{-1}\,\Psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), V_\varepsilon^{n-1} - V_\varepsilon^n)^h.&\end{aligned} \tag{2.32c}$$

On noting the elementary identity

$$2\,r\,(r-s) = (r^2 - s^2) + (r-s)^2 \quad \forall\, r, s \in \mathbb{R},$$

it follows from (2.32a–c), (1.15), (2.22c) and the convexity of $-\Psi_0$, recall (1.6), that

$$\begin{aligned}&\tfrac{\gamma}{2}\,[\,|U_\varepsilon^n|_1^2 + |U_\varepsilon^n - U_\varepsilon^{n-1}|_1^2 - |U_\varepsilon^{n-1}|_1^2 + |V_\varepsilon^n|_1^2 + |V_\varepsilon^n - V_\varepsilon^{n-1}|_1^2 - |V_\varepsilon^{n-1}|_1^2\,]\\ &\quad + \gamma^{-1}\,\tau_n\,|\,[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{\frac12}\,\nabla W_\varepsilon^n|_0^2 + [\ell(\gamma)]^{-1}\,\tau_n\,|Z_\varepsilon^n|_h^2\\ &\quad \leq -\gamma^{-1}\,(\Psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), U_\varepsilon^n - U_\varepsilon^{n-1})^h - \gamma^{-1}\,(\Psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), V_\varepsilon^n - V_\varepsilon^{n-1})^h\\ &\quad\quad - \alpha\,\gamma^{-1}\,\tau_n\,(\Xi_\varepsilon(U_\varepsilon^{n-1})\,\nabla\Phi_\varepsilon^n, \nabla W_\varepsilon^n)\\ &\quad \leq \gamma^{-1}\,(\Psi(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}) - \Psi(U_\varepsilon^n, V_\varepsilon^n), 1)^h\\ &\quad\quad + \gamma^{-1}\,\tfrac{\tau_n}{2}\,\left[\,|\,[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{\frac12}\,|\nabla W_\varepsilon^n|_0^2 + 2\,\alpha^2\,|\,[\Lambda_\varepsilon(U_\varepsilon^{n-1})]^{\frac12}\,\nabla\Phi_\varepsilon^n|_0^2\right].\end{aligned} \tag{2.33}$$

Hence the desired result (2.26a) follows from (2.33), (2.26b) and (2.24).

Choosing $\chi \equiv \pi^h[G'_\varepsilon(U_\varepsilon^{n-1})]$ in (2.14b), and noting (2.12b) yields that

$$\gamma\,(U_\varepsilon^n - U_\varepsilon^{n-1}, G'_\varepsilon(U_\varepsilon^{n-1}))^h + \tau_n\,(\nabla[W_\varepsilon^n + \alpha\,\Phi_\varepsilon^n], \nabla U_\varepsilon^{n-1}) = 0\,; \tag{2.34}$$

while choosing $\chi \equiv -\Delta^h V_\varepsilon^n$ in (2.14c), and noting (2.20) yields that

$$\tfrac{\ell(\gamma)}{2}\,\left[\,|V_\varepsilon^n|_1^2 + |V_\varepsilon^n - V_\varepsilon^{n-1}|_1^2 - |V_\varepsilon^{n-1}|_1^2\right] = \ell(\gamma)\,(V_\varepsilon^n - V_\varepsilon^{n-1}, -\Delta^h V_\varepsilon^n)^h = -\tau_n\,(\nabla Z_\varepsilon^n, \nabla V_\varepsilon^n)\,. \tag{2.35}$$

We now extend an argument in Barrett, Blowey, and Garcke (2001, Theorem 2.3), where the authors treated the one dimensional case of $\mathcal{K} = [-1, 1]$. The case $\mathcal{K} = \triangle ABC \subset \mathbb{R}^2$ studied here, requires some special considerations. Let $j \in J$, then for $\{U_\varepsilon^n(p_j), V_\varepsilon^n(p_j)\} \in \mathcal{K}$ we distinguish the following cases. For ease of notation, let $v_b := \frac{2}{\sqrt{3}}$.

(*i*) $\{U_\varepsilon^n(p_j), V_\varepsilon^n(p_j)\} \in \mathcal{K} \setminus \partial\mathcal{K}$, (*ii*) $U_\varepsilon^n(p_j) = 1,\ V_\varepsilon^n(p_j) \in (-v_b, v_b)$,
(*iii*) $U_\varepsilon^n(p_j) \in (-1, 1),\ V_\varepsilon^n(p_j) = \frac{v_b}{2}\,(U_\varepsilon^n(p_j) + 1)$,
(*iv*) $U_\varepsilon^n(p_j) \in (-1, 1),\ V_\varepsilon^n(p_j) = -\frac{v_b}{2}\,(U_\varepsilon^n(p_j) + 1)$,
(*v*) $\{U_\varepsilon^n(p_j), V_\varepsilon^n(p_j)\} = \{1, v_b\}$, (*vi*) $\{U_\varepsilon^n(p_j), V_\varepsilon^n(p_j)\} = \{1, -v_b\}$,
(*vii*) $\{U_\varepsilon^n(p_j), V_\varepsilon^n(p_j)\} = \{-1, 0\}$.

In what follows, we choose $\delta > 0$ sufficiently small so that the specified $\{\chi_1, \chi_2\} \in \mathcal{K}$ can be chosen in (2.14d). In case (i) we have on choosing $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon \pm \delta\,\chi_j, V^n_\varepsilon\}$ and $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon, V^n_\varepsilon \pm \delta\,\chi_j\}$, respectively, that

$$A^U_j := \gamma\,(\nabla U^n_\varepsilon, \nabla\chi_j) - (W^n_\varepsilon - \gamma^{-1}\,\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \chi_j)^h = 0 \tag{2.36a}$$

$$\text{and} \quad A^V_j := \gamma\,(\nabla V^n_\varepsilon, \nabla\chi_j) - (Z^n_\varepsilon - \gamma^{-1}\,\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \chi_j)^h = 0\,. \tag{2.36b}$$

In case (ii) we have on choosing $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon, V^n_\varepsilon \pm \delta\,\chi_j\}$ and $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon - \delta\,\chi_j, V^n_\varepsilon\}$, respectively, that $A^U_j \leq 0$ and $A^V_j = 0$. For case (iii) we choose $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon \pm \delta\,\chi_j, V^n_\varepsilon \pm \frac{v_b}{2}\,\delta\,\chi_j\}$ and $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon, V^n_\varepsilon - \delta\,\chi_j\}$, respectively, so that $A^U_j + \frac{v_b}{2}\,A^V_j = 0$ and $A^V_j \leq 0$. Similarly, we obtain for case (iv), on choosing $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon \pm \delta\,\chi_j, V^n_\varepsilon \mp \frac{v_b}{2}\delta\,\chi_j\}$ and $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon, V^n_\varepsilon + \delta\,\chi_j\}$, respectively, that $A^U_j - \frac{v_b}{2}\,A^V_j = 0$ and $A^V_j \geq 0$. In case (v) we obtain that $A^V_j \leq 0$ and $A^U_j + \frac{v_b}{2}\,A^V_j \leq 0$; on choosing $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon, V^n_\varepsilon - \delta\,\chi_j\}$ and $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon - \delta\,\chi_j, V^n_\varepsilon - \frac{v_b}{2}\delta\,\chi_j\}$, respectively. Similarly, for case (vi) we have that $A^V_j \geq 0$ and $A^U_j - \frac{v_b}{2}\,A^V_j \leq 0$ hold; on choosing $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon, V^n_\varepsilon + \delta\,\chi_j\}$ and $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon - \delta\,\chi_j, V^n_\varepsilon + \frac{v_b}{2}\delta\,\chi_j\}$, respectively. Finally, in case (vii) we have that $A^U_j \pm \frac{v_b}{2}\,A^V_j \geq 0$ hold; on choosing $\{\chi_1, \chi_2\} \equiv \{U^n_\varepsilon + \delta\,\chi_j, V^n_\varepsilon \pm \frac{v_b}{2}\delta\,\chi_j\}$.

From (2.20), (2.2) and (2.1) it follows for the cases (ii), (v) and (vi) that

$$U^n_\varepsilon(p_j) = 1 \implies U^n_\varepsilon(p_j) \geq U^n_\varepsilon(p_i) \quad \forall\, i \in J \implies \Delta^h U^n_\varepsilon(p_j) \leq 0. \tag{2.37a}$$

Similarly, in the cases (iii), (v) and (vii) it holds that

$$\Delta^h(U^n_\varepsilon - \sqrt{3}\,V_\varepsilon)(p_j) \geq 0, \tag{2.37b}$$

while in the cases (iv), (vi) and (vii) we have that

$$\Delta^h(U^n_\varepsilon + \sqrt{3}\,V_\varepsilon)(p_j) \geq 0. \tag{2.37c}$$

Combining (2.36a,b) and (2.37a–c) yields for all cases (i) – (vii) that

$$-[\,A^U_j\,\Delta^h U^n_\varepsilon(p_j) + A^V_j\,\Delta^h V^n_\varepsilon(p_j)\,] \leq 0\,. \tag{2.38}$$

Summing (2.38) for all $j \in J$ yields, on noting (2.36a,b), (2.20) and (2.2), that

$$\begin{aligned}
&\gamma\,|\Delta^h U^n_\varepsilon|^2_h + \gamma\,|\Delta^h V^n_\varepsilon|^2_h \\
&\quad \leq -(W^n_\varepsilon - \gamma^{-1}\,\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \Delta^h U^n_\varepsilon)^h - (Z^n_\varepsilon - \gamma^{-1}\,\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \Delta^h V^n_\varepsilon)^h \\
&\quad = (\nabla[W^n_\varepsilon - \gamma^{-1}\,\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon)], \nabla U^n_\varepsilon) + (\nabla[Z^n_\varepsilon - \gamma^{-1}\,\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon)], \nabla V^n_\varepsilon).
\end{aligned} \tag{2.39}$$

It follows from (2.34), (2.35), (2.7) and (2.39) that

$$
\begin{aligned}
&\gamma\,(G_\varepsilon(U^n_\varepsilon) - G_\varepsilon(U^{n-1}_\varepsilon), 1)^h + \gamma\,\tau_n\,|\Delta^h U^n_\varepsilon|^2_h + \tfrac{\ell(\gamma)}{2}\,\left[\,|V^n_\varepsilon|^2_1 + |V^n_\varepsilon - V^{n-1}_\varepsilon|^2_1 - |V^{n-1}_\varepsilon|^2_1\,\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \gamma\,\tau_n\,|\Delta^h V^n_\varepsilon|^2_h \\
&\quad \le \gamma\,(U^n_\varepsilon - U^{n-1}_\varepsilon, G'_\varepsilon(U^n_\varepsilon))^h + \tau_n\,[\,(\nabla[W^n_\varepsilon - \gamma^{-1}\,\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon)], \nabla U^n_\varepsilon) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad - \gamma^{-1}\,(\nabla\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \nabla V^n_\varepsilon)\,] \\
&\quad \le \gamma\,(U^n_\varepsilon - U^{n-1}_\varepsilon, G'_\varepsilon(U^n_\varepsilon) - G'_\varepsilon(U^{n-1}_\varepsilon))^h \\
&\qquad + \tau_n\,[\,(\nabla W^n_\varepsilon, \nabla[U^n_\varepsilon - U^{n-1}_\varepsilon]\,) - \alpha\,(\nabla \Phi^n_\varepsilon, \nabla U^{n-1}_\varepsilon) \\
&\qquad\quad - \gamma^{-1}\,(\nabla\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \nabla U^n_\varepsilon) - \gamma^{-1}\,(\nabla\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \nabla V^n_\varepsilon)\,] \\
&\quad \le \varepsilon^{-1}\,\gamma\,|U^n_\varepsilon - U^{n-1}_\varepsilon|^2_h + \tau_n\,[\,(\nabla W^n_\varepsilon, \nabla[U^n_\varepsilon - U^{n-1}_\varepsilon]\,) - \alpha\,(\nabla \Phi^n_\varepsilon, \nabla U^{n-1}_\varepsilon) \\
&\qquad\qquad\quad - \gamma^{-1}\,(\nabla\Psi_{,u}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \nabla U^n_\varepsilon) - \gamma^{-1}\,(\nabla\Psi_{,v}(U^{n-1}_\varepsilon, V^{n-1}_\varepsilon), \nabla V^n_\varepsilon)\,]
\end{aligned}
$$

and hence the desired result (2.27). □

The results of the preceding lemma will now be used to derive fundamental a priori estimates.

THEOREM. 2.1 *Let the assumptions* (A) *hold and* $\{U^0_\varepsilon, V^0_\varepsilon\} \in K^h$. *Then for all* $\varepsilon \in (0,1)$, $h > 0$ *and for all time partitions* $\{\tau_n\}_{n=1}^N$, *the solution* $\{\Phi^n_\varepsilon, U^n_\varepsilon, V^n_\varepsilon, W^n_\varepsilon, Z^n_\varepsilon\}_{n=1}^N$ *to* $(\mathrm{P}^{h,\tau}_\varepsilon)$ *is such that* $⨍ U^n_\varepsilon = ⨍ U^0_\varepsilon$, $n = 1 \to N$, *and*

$$
\begin{aligned}
&\gamma \max_{n=1\to N} \|U^n_\varepsilon\|^2_1 + \gamma \max_{n=1\to N} \|V^n_\varepsilon\|^2_1 + \gamma \sum_{n=1}^N \left[\,|U^n_\varepsilon - U^{n-1}_\varepsilon|^2_1 + |V^n_\varepsilon - V^{n-1}_\varepsilon|^2_1\,\right] \\
&\quad + \sum_{n=1}^N \tau_n \left[\gamma^{-1}\,|\,[\Xi_\varepsilon(U^{n-1}_\varepsilon)]^{\frac12}\,\nabla W^n_\varepsilon|^2_0 + [\ell(\gamma)]^{-1}\,|Z^n_\varepsilon|^2_h + \ell(\gamma)\,|\tfrac{V^n_\varepsilon - V^{n-1}_\varepsilon}{\tau_n}|^2_h\right] \\
&\qquad\qquad \le C\left[\,\gamma\,\|U^0_\varepsilon\|^2_1 + \gamma\,\|V^0_\varepsilon\|^2_1 + \gamma^{-1}\,(1 + T\,|g|^2_{0,\partial_2\Omega})\,\right]. \qquad (2.40)
\end{aligned}
$$

In addition

$$
\begin{aligned}
&\gamma \sum_{n=1}^N \tau_n \left|\mathcal{G}[\tfrac{U^n_\varepsilon - U^{n-1}_\varepsilon}{\tau_n}]\right|^2_1 + \gamma\,\tau^{-\frac12} \sum_{n=1}^N |U^n_\varepsilon - U^{n-1}_\varepsilon|^2_0 \\
&\qquad\qquad \le C\left[\,\gamma\,\|U^0_\varepsilon\|^2_1 + \gamma\,\|V^0_\varepsilon\|^2_1 + \gamma^{-1}\,(1 + T\,|g|^2_{0,\partial_2\Omega})\,\right] \qquad (2.41)
\end{aligned}
$$

and

$$
\begin{aligned}
&\gamma \max_{n=1\to N} (G_\varepsilon(U^n_\varepsilon), 1)^h + \gamma \sum_{n=1}^N \tau_n\,|\Delta^h U^n_\varepsilon|^2_h + \gamma \sum_{n=1}^N \tau_n\,|\Delta^h V^n_\varepsilon|^2_h \\
&\quad \le \gamma\,(G_\varepsilon(U^0_\varepsilon), 1)^h + \alpha^2 \sum_{n=1}^N \tau_n\,|\pi^h[F'_\varepsilon(U^{n-1}_\varepsilon)]|^2_{0,\partial_2\Omega} \\
&\qquad + C(T)\,[\,1 + \gamma^{-2} + \varepsilon^{-1}\tau^{\frac12}\,]\,\left[\,\gamma\,\|U^0_\varepsilon\|^2_1 + \gamma\,\|V^0_\varepsilon\|^2_1 + \gamma^{-1}\,(1 + T\,|g|^2_{0,\partial_2\Omega})\,\right]. \qquad (2.42)
\end{aligned}
$$

Proof. Summing (2.26a) from $n = 1 \to k$ yields for any $k \le N$ that

$$\mathcal{E}(U_\varepsilon^k, V_\varepsilon^k) + \tfrac{1}{2}\,\gamma \sum_{n=1}^{k} \left[|U_\varepsilon^n - U_\varepsilon^{n-1}|_1^2 + |V_\varepsilon^n - V_\varepsilon^{n-1}|_1^2\right] + [\ell(\gamma)]^{-1} \sum_{n=1}^{k} \tau_n\, |Z_\varepsilon^n|_h^2$$
$$+ \tfrac{1}{2}\,\gamma^{-1} \sum_{n=1}^{k} \tau_n\, |\,[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{\frac{1}{2}}\, \nabla W_\varepsilon^n|_0^2 \le \mathcal{E}(U_\varepsilon^0, V_\varepsilon^0) + \tfrac{1}{2}\,\alpha^2\,\gamma^{-1}\, t_k\, |g|_{0,\partial_2\Omega}^2\,. \qquad (2.43)$$

The desired result (2.40) then follows from (2.43), (2.26b), (2.2), (2.17), (2.14c) and the fact that $\{U_\varepsilon^n(p_j), V_\varepsilon^n(p_j)\} \in \mathcal{K}$, $\forall\, j \in J$, $n = 0 \to N$. Then (2.41) follows from (1.13), (2.3), (2.14b), (2.22b,c), (2.19), (2.24) and (2.40); see Barrett, Nürnberg, and Styles (2004, Theorem 2.6) for details.

Finally, summing (2.27) from $n = 1 \to k$ and noting (1.6), (2.2), (2.17) and (2.22b) yields for any $k \le N$ that

$$\gamma\,(G_\varepsilon(U_\varepsilon^k), 1)^h + \gamma \sum_{n=1}^{k} \tau_n\, |\Delta^h U_\varepsilon^n|_h^2 + \gamma \sum_{n=1}^{k} \tau_n\, |\Delta^h V_\varepsilon^n|_h^2 \le \gamma\,(G_\varepsilon(U_\varepsilon^0), 1)^h$$
$$+ \sum_{n=1}^{k} \left[\,\varepsilon^{-1}\,\gamma\, |U_\varepsilon^n - U_\varepsilon^{n-1}|_0^2 + \tau_n\, |\alpha\,(\nabla\Phi_\varepsilon^n, \nabla U_\varepsilon^{n-1})|\,\right] + \gamma^{-1}\, t_k\, |\Psi_0|_{2,\infty,\mathcal{K}} \left[\max_{n=0\to k} \|U_\varepsilon^n\|_1^2\right.$$
$$\left. + \max_{n=0\to k} \|V_\varepsilon^n\|_1^2\right] + \left[\varepsilon^{-1} \sum_{n=1}^{k} \tau_n\, |[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{\frac{1}{2}}\, \nabla W_\varepsilon^n|_0^2\right]^{\frac{1}{2}} \left[\sum_{n=1}^{k} \tau_n\, |U_\varepsilon^n - U_\varepsilon^{n-1}|_1^2\right]^{\frac{1}{2}}. \qquad (2.44)$$

The desired result (2.42) then follows from (2.44), (2.25), (1.15), (2.40) and (2.41). □

LEMMA. 2.4 *Let* $\{u^0, v^0\} \in K \cap [W^{1,p}(\Omega)]^2$ *with* $p > 2$, *and the assumptions* (A) *hold. On choosing* $\{U_\varepsilon^0, V_\varepsilon^0\} \equiv \{\pi^h u^0, \pi^h v^0\}$ *it follows that* $\{U_\varepsilon^0, V_\varepsilon^0\} \in K^h$ *is such that for all* $h > 0$

$$\|U_\varepsilon^0\|_1^2 + \|V_\varepsilon^0\|_1^2 + (G_\varepsilon(U_\varepsilon^0), 1)^h \le C(T). \qquad (2.45)$$

Proof. The desired result (2.45) follows from (2.16), (2.6) and (2.5). □

REMARK. 2.1 The approximation $(\mathrm{P}_\varepsilon^{h,\tau})$ of (P) requires solving for $\{\Phi_\varepsilon^n, U_\varepsilon^n, W_\varepsilon^n\}$ over the whole domain Ω, due to the non-degeneracy of $\Lambda_\varepsilon(\cdot)$ and $\Xi_\varepsilon(\cdot)$, see (2.22a,b). For computational speed it would be more convenient to solve for Φ_ε^n just in the conductor and interfacial regions, $U_\varepsilon^{n-1} > -1$, and for $\{U_\varepsilon^n, W_\varepsilon^n\}$ just in the interfacial region, $|U_\varepsilon^{n-1}| < 1$. With this in mind, we recall Remark 2.10 in Barrett, Nürnberg, and Styles (2004) and introduce the following approximation of (P). Adopting the notation (2.9) and (2.10), let $\widetilde{\Lambda}_\varepsilon, \widetilde{\Xi}_\varepsilon : S^h \to [L^\infty(\Omega)]^{2\times 2}$ such that $\widetilde{\Lambda}_\varepsilon(\eta^h)\,|_\sigma := R_\sigma\, \widehat{\Lambda}_\varepsilon^\star(\widehat{\eta}^h)\,|_{\widehat{\sigma}}\, R_\sigma^T$ and $\widetilde{\Xi}_\varepsilon(\eta^h)\,|_\sigma := R_\sigma\, \widehat{\Xi}_\varepsilon^\star(\widehat{\eta}^h)\,|_{\widehat{\sigma}}\, R_\sigma^T$, where

$$[\widehat{\Lambda}_\varepsilon^\star(\widehat{\eta}^h)\,|_{\widehat{\sigma}}]_{kk} := \begin{cases} 0 & \text{if } \widehat{\eta}^h(p_{j_k}) = \widehat{\eta}^h(p_{j_0}) = -1, \\ [\widehat{\Lambda}_\varepsilon(\widehat{\eta}^h)\,|_{\widehat{\sigma}}]_{kk} & \text{otherwise;} \end{cases}$$

$$\text{and} \quad [\widehat{\Xi}_\varepsilon^\star(\widehat{\eta}^h)\,|_{\widehat{\sigma}}]_{kk} := \begin{cases} 0 & \text{if } \widehat{\eta}^h(p_{j_k}) = \widehat{\eta}^h(p_{j_0}) = \pm 1, \\ [\widehat{\Xi}_\varepsilon(\widehat{\eta}^h)\,|_{\widehat{\sigma}}]_{kk} & \text{otherwise.} \end{cases}$$

We note that the key identities, $\Lambda_\varepsilon(\eta^h)$ in (2.8a,b) replaced by $\widetilde{\Lambda}_\varepsilon(\eta^h)$ and $\Xi_\varepsilon(\eta^h)$ in (2.12a,b) replaced by $\widetilde{\Xi}_\varepsilon(\eta^h)$, still hold. We then introduce the approximation $(\widetilde{\mathrm{P}}^{h,\tau}_\varepsilon)$ of (P), which is the same as $(\mathrm{P}^{h,\tau}_\varepsilon)$ but with $\Lambda_\varepsilon(U^{n-1}_\varepsilon)$ in (2.14a) replaced by $\widetilde{\Lambda}_\varepsilon(U^{n-1}_\varepsilon)$ and $\Xi_\varepsilon(U^{n-1}_\varepsilon)$ in (2.14b) replaced by $\widetilde{\Xi}_\varepsilon(U^{n-1}_\varepsilon)$. As $\widetilde{\Lambda}_\varepsilon(\cdot)$ and $\widetilde{\Xi}_\varepsilon(\cdot)$ are now degenerate, existence of a solution $\{\Phi^n_\varepsilon, U^n_\varepsilon, V^n_\varepsilon, W^n_\varepsilon, Z^n_\varepsilon\}$ to $(\widetilde{\mathrm{P}}^{h,\tau}_\varepsilon)$ does not appear to be trivial. However, this can easily be established by splitting the nodes into passive and active sets, see e.g. (Barrett, Blowey, and Garcke 1999). Moreover, one can show that $\{U^n_\varepsilon, V^n_\varepsilon, Z^n_\varepsilon\}$ is unique. Furthermore, one can establish analogues of the energy estimates (2.40) and (2.41). Unfortunately, it does not appear possible to establish an analogue of the key energy estimate (2.42) for $(\widetilde{\mathrm{P}}^{h,\tau}_\varepsilon)$.

3 Convergence

Let

$$U_\varepsilon(t) := \tfrac{t-t_{n-1}}{\tau_n}\, U^n_\varepsilon + \tfrac{t_n - t}{\tau_n}\, U^{n-1}_\varepsilon \qquad t \in [t_{n-1}, t_n] \quad n \geq 1, \tag{3.1a}$$

$$U^+_\varepsilon(t) := U^n_\varepsilon, \qquad U^-_\varepsilon(t) := U^{n-1}_\varepsilon \qquad t \in (t_{n-1}, t_n] \quad n \geq 1. \tag{3.1b}$$

We note for future reference that

$$U_\varepsilon - U^\pm_\varepsilon = (t - t^\pm_n)\, \tfrac{\partial U_\varepsilon}{\partial t} \qquad t \in (t_{n-1}, t_n) \quad n \geq 1, \tag{3.2}$$

where $t^+_n := t_n$ and $t^-_n := t_{n-1}$. We introduce also

$$\bar{\tau}(t) := \tau_n \qquad t \in (t_{n-1}, t_n] \quad n \geq 1. \tag{3.3}$$

Using the above notation, and introducing analogous notation for $V_\varepsilon, W^+_\varepsilon, Z^+_\varepsilon$ and Φ^+_ε, $(\mathrm{P}^{h,\tau}_\varepsilon)$ can be restated as: Find $\{\Phi^+_\varepsilon, \{U_\varepsilon, V_\varepsilon\}, W^+_\varepsilon, Z^+_\varepsilon\} \in L^\infty(0,T;S^h) \times C([0,T];K^h) \times [L^\infty(0,T;S^h)]^2$ such that for all $\chi \in L^\infty(0,T;S^h)$

$$\int_0^T (\Lambda_\varepsilon(U^-_\varepsilon)\, \nabla \Phi^+_\varepsilon, \nabla \chi)\, \mathrm{d}t + \int_0^T \int_{\partial_2\Omega} \Phi^+_\varepsilon\, \chi \,\mathrm{d}s\, \mathrm{d}t = \int_0^T \int_{\partial_2\Omega} g\, \chi \,\mathrm{d}s\, \mathrm{d}t, \tag{3.4a}$$

$$\int_0^T \left[\gamma \left(\tfrac{\partial U_\varepsilon}{\partial t}, \chi\right)^h + (\Xi_\varepsilon(U^-_\varepsilon)\, \nabla[W^+_\varepsilon + \alpha\, \Phi^+_\varepsilon], \nabla \chi)\right] \mathrm{d}t = 0, \tag{3.4b}$$

$$\int_0^T \left[\ell(\gamma)\left(\tfrac{\partial V_\varepsilon}{\partial t}, \chi\right)^h + (Z^+_\varepsilon, \chi)^h\right] \mathrm{d}t = 0; \tag{3.4c}$$

where for *a.a.* $t \in (0,T)$

$$\begin{aligned} &\gamma\left[(\nabla U^+_\varepsilon, \nabla[\chi_1 - U^+_\varepsilon]) + (\nabla V^+_\varepsilon, \nabla[\chi_2 - V^+_\varepsilon])\right] \\ &\quad \geq \left[(W^+_\varepsilon - \gamma^{-1}\, \Psi_{,u}(U^-_\varepsilon, V^-_\varepsilon), \chi_1 - U^+_\varepsilon)^h + (Z^+_\varepsilon - \gamma^{-1}\, \Psi_{,v}(U^-_\varepsilon, V^-_\varepsilon), \chi_2 - V^+_\varepsilon)^h\right] \\ &\qquad\qquad \forall\ \{\chi_1, \chi_2\} \in K^h. \end{aligned} \tag{3.4d}$$

LEMMA. 3.1 *Let* $\{u^0, v^0\} \in K \cap [W^{1,p}(\Omega)]^2$ *with* $p > 2$, *and* $\fint u^0 \in (-1,1)$. *Let* $\{\mathcal{T}^h, U_\varepsilon^0, V_\varepsilon^0, \{\tau_n\}_{n=1}^N, \varepsilon\}_{h>0}$ *be such that*

(i) $\{U_\varepsilon^0, V_\varepsilon^0\} \equiv \{\pi^h u^0, \pi^h v^0\}$;

(ii) Ω *and* $\{\mathcal{T}^h\}_{h>0}$ *fulfil assumption* (A), $\varepsilon \in (0,1)$ *with* $\varepsilon \to 0$ *as* $h \to 0$ *and* $\tau_n \leq C\,\tau_{n-1} \leq C\,\varepsilon^2$, $n = 2 \to N$;

Then there exists a subsequence of $\{\Phi_\varepsilon^+, U_\varepsilon, V_\varepsilon, W_\varepsilon^+, Z_\varepsilon^+\}_h$, *where* $\{\Phi_\varepsilon^+, U_\varepsilon, V_\varepsilon, W_\varepsilon^+, Z_\varepsilon^+\}$ *solve* $(\mathrm{P}_\varepsilon^{h,\tau})$, *and functions*

$$u \in L^\infty(0,T;H^1(\Omega)) \cap H^1(0,T;(H^1(\Omega))'), \tag{3.5a}$$
$$v \in L^\infty(0,T;H^1(\Omega)) \cap H^1(0,T;L^2(\Omega)) \qquad \textit{and} \qquad z \in L^2(\Omega_T); \tag{3.5b}$$

such that $\{u(x,t), v(x,t)\} \in \mathcal{K}$ *for a.e.* $(x,t) \in \Omega_T$, *with* $u(\cdot,0) = u^0(\cdot)$, $v(\cdot,0) = v^0(\cdot)$ *in* $L^2(\Omega)$ *and* $\fint u(\cdot,t) = \fint u^0$ *for a.a.* $t \in (0,T)$, *such that as* $h \to 0$

$$\begin{aligned}
&U_\varepsilon,\ U_\varepsilon^\pm \to u \quad \textit{and} \quad V_\varepsilon,\ V_\varepsilon^\pm \to v && \textit{weak-}*\ \textit{in } L^\infty(0,T;H^1(\Omega)), && (3.6a)\\
&\mathcal{G}\tfrac{\partial U_\varepsilon}{\partial t} \to \mathcal{G}\tfrac{\partial u}{\partial t} && \textit{weakly in } L^2(0,T;H^1(\Omega)), && (3.6b)\\
&\tfrac{\partial V_\varepsilon}{\partial t} \to \tfrac{\partial v}{\partial t} \quad \textit{and} \quad Z_\varepsilon^+ \to z && \textit{weakly in } L^2(\Omega_T), && (3.6c)\\
&U_\varepsilon,\ U_\varepsilon^\pm \to u \quad \textit{and} \quad V_\varepsilon,\ V_\varepsilon^\pm \to v && \textit{strongly in } L^2(0,T;L^s(\Omega)), && (3.7a)\\
&\Xi_\varepsilon(U_\varepsilon^-) \to b(u)\,\mathcal{I} \quad \textit{and} \quad \Lambda_\varepsilon(U_\varepsilon^-) \to c(u)\,\mathcal{I} && \textit{strongly in } L^2(0,T;L^s(\Omega)); && (3.7b)
\end{aligned}$$

for all $s \in [2,\infty)$.

If in addition $u^0 \in H^2(\Omega)$ *with* $\frac{\partial u^0}{\partial \nu} = 0$ *on* $\partial\Omega$ *and*

$$\alpha^2 \int_0^T |\pi^h[F_\varepsilon'(U_\varepsilon^-)]|_{0,\partial_2\Omega}^2 \,\mathrm{d}t \leq C, \tag{3.8}$$

then u *in addition to* (3.5a) *satisfies*

$$u \in L^2(0,T;H^2(\Omega)) \tag{3.9}$$

and there exists a subsequence of $\{U_\varepsilon\}_h$ *satisfying* (3.6a,b), (3.7a,b) *and as* $h \to 0$

$$\begin{aligned}
&\Delta^h U_\varepsilon,\ \Delta^h U_\varepsilon^\pm \to \Delta u && \textit{weakly in } L^2(\Omega_T), && (3.10a)\\
&U_\varepsilon,\ U_\varepsilon^\pm \to u && \textit{weakly in } L^2(0,T;W^{1,s}(\Omega)), \quad \textit{for any } s \in [2,\infty), && (3.10b)\\
&U_\varepsilon,\ U_\varepsilon^\pm \to u && \textit{strongly in } L^2(0,T;C^{0,\zeta}(\overline{\Omega})), \quad \textit{for any } \zeta \in (0,1), && (3.10c)\\
&U_\varepsilon,\ U_\varepsilon^\pm \to u && \textit{strongly in } L^2(0,T;H^1(\Omega))\,. && (3.10d)
\end{aligned}$$

Similarly, if in addition $v^0 \in H^2(\Omega)$ *with* $\frac{\partial v^0}{\partial \nu} = 0$ *on* $\partial\Omega$, *then* v *in addition to* (3.5b) *satisfies*

$$v \in L^2(0,T;H^2(\Omega)) \tag{3.11}$$

and there exists a subsequence of $\{V_\varepsilon\}_h$ *satisfying* (3.6a,c), (3.7a) *and as* $h \to 0$

$$\Delta^h V_\varepsilon,\ \Delta^h V_\varepsilon^\pm \to \Delta v \qquad \textit{weakly in } L^2(\Omega_T), \tag{3.12a}$$
$$V_\varepsilon,\ V_\varepsilon^\pm \to v \qquad \textit{weakly in } L^2(0,T;W^{1,s}(\Omega)), \quad \textit{for any } s \in [2,\infty), \tag{3.12b}$$
$$V_\varepsilon,\ V_\varepsilon^\pm \to v \qquad \textit{strongly in } L^2(0,T;C^{0,\zeta}(\overline{\Omega})), \quad \textit{for any } \zeta \in (0,1), \tag{3.12c}$$
$$V_\varepsilon,\ V_\varepsilon^\pm \to v \qquad \textit{strongly in } L^2(0,T;H^1(\Omega))\,. \tag{3.12d}$$

Proof. Noting the definitions (3.1a,b), (3.3), the bounds in (2.24), (2.40), (2.41) and (2.42) together with (1.14), (2.45) and our assumption (i) imply that

$$\begin{aligned}
&\|\,[\Lambda_\varepsilon(U_\varepsilon^-)]^{\frac{1}{2}}\,\nabla\Phi_\varepsilon^+\|_{L^2(\Omega_T)}^2 + \|\Phi_\varepsilon^+\|_{L^2(0,T;L^2(\partial_2\Omega))}^2 + \|U_\varepsilon^{(\pm)}\|_{L^\infty(0,T;H^1(\Omega))}^2 + \|V_\varepsilon^{(\pm)}\|_{L^\infty(0,T;H^1(\Omega))}^2 \\
&\quad + \|\bar\tau^{\frac{1}{2}}\tfrac{\partial U_\varepsilon}{\partial t}\|_{L^2(0,T;H^1(\Omega))}^2 + \|\bar\tau^{\frac{1}{2}}\tfrac{\partial V_\varepsilon}{\partial t}\|_{L^2(0,T;H^1(\Omega))}^2 + \|\,[\Xi_\varepsilon(U_\varepsilon^-)]^{\frac{1}{2}}\,\nabla W_\varepsilon^+\|_{L^2(\Omega_T)}^2 \\
&\quad + \|\mathcal{G}\tfrac{\partial U_\varepsilon}{\partial t}\|_{L^2(0,T;H^1(\Omega))}^2 + \tau^{-\frac{1}{2}}\,\|\bar\tau^{\frac{1}{2}}\tfrac{\partial U_\varepsilon}{\partial t}\|_{L^2(\Omega_T)}^2 + \|\tfrac{\partial V_\varepsilon}{\partial t}\|_{L^2(\Omega_T)}^2 + \|Z_\varepsilon^+\|_{L^2(\Omega_T)}^2 \le C,
\end{aligned} \tag{3.13a}$$

and

$$\|\Delta^h U_\varepsilon^+\|_{L^2(\Omega_T)}^2 + \|\Delta^h V_\varepsilon^+\|_{L^2(\Omega_T)}^2 \le C. \tag{3.13b}$$

Furthermore, we deduce from (3.2) and (3.13a) that

$$\begin{aligned}
&\|U_\varepsilon - U_\varepsilon^\pm\|_{L^2(0,T;H^1(\Omega))}^2 + \|V_\varepsilon - V_\varepsilon^\pm\|_{L^2(0,T;H^1(\Omega))}^2 \\
&\qquad \le \|\bar\tau\,\tfrac{\partial U_\varepsilon}{\partial t}\|_{L^2(0,T;H^1(\Omega))}^2 + \|\bar\tau\,\tfrac{\partial V_\varepsilon}{\partial t}\|_{L^2(0,T;H^1(\Omega))}^2 \le C\,\tau\,.
\end{aligned} \tag{3.14}$$

Hence on noting (3.13a), (3.14), $\{U_\varepsilon(\cdot,t), V_\varepsilon(\cdot,t)\} \in K^h$, and (1.12a) we can choose a subsequence $\{\Phi_\varepsilon^+, U_\varepsilon, V_\varepsilon, W_\varepsilon^+, Z_\varepsilon^+\}_h$ such that the convergence results (3.5a,b), (3.6a–c) and (3.7a) hold. Then (3.5a,b) and Theorem 2.1 yield, on noting (1.12b), assumption (i) and (2.16) that the subsequence satisfies the additional initial and integral conditions.

The proof of (3.7b) can be found in the proof of Lemma 3.1 in Barrett, Nürnberg, and Styles (2004). Moreover, the proof of the results (3.9)–(3.10b) and the result on U_ε in (3.10c), are also in Barrett, Nürnberg, and Styles (2004, Lemma 3.1), where they are derived from the key entropy bound (3.13b). We now establish (3.10c) for $U_\varepsilon^\pm$. For any $\zeta \in (0,1)$, $s \in (\frac{2}{1-\zeta},\infty)$ and any $\bar s \in (\frac{2}{1-\zeta}, s)$ it holds on noting the compact embedding $W^{1,\bar s}(\Omega) \hookrightarrow C^{0,\zeta}(\overline{\Omega})$, (3.14) and (3.10b) that

$$\begin{aligned}
\|U_\varepsilon - U_\varepsilon^\pm\|_{L^2(0,T;C^{0,\zeta}(\overline{\Omega}))} &\le \|U_\varepsilon - U_\varepsilon^\pm\|_{L^2(0,T;W^{1,\bar s}(\Omega))} \\
&\le \|U_\varepsilon - U_\varepsilon^\pm\|_{L^2(0,T;H^1(\Omega))}^q\,\|U_\varepsilon - U_\varepsilon^\pm\|_{L^2(0,T;W^{1,s}(\Omega))}^{1-q} \le C\,\tau^{\frac{q}{2}}\,,
\end{aligned} \tag{3.15}$$

where $q = \frac{2\,(s-\bar s)}{(s-2)\,\bar s} \in (0,1)$. Combining (3.15), assumption (ii) and the established result on U_ε in (3.10c) yields the desired result on $U_\varepsilon^\pm$ in (3.10c).

We now prove (3.10d). We have that

$$\begin{aligned}
\|\nabla(U_\varepsilon^+ - u)\|_{L^2(\Omega_T)}^2 &\le \left|\int_{\Omega_T} \nabla(U_\varepsilon^+ - u)\,.\,\nabla u \,\mathrm{d}x\,\mathrm{d}t\right| \\
&\quad + \left|\int_{\Omega_T} \nabla(U_\varepsilon^+ - \pi^h u)\,.\,\nabla U_\varepsilon^+ \,\mathrm{d}x\,\mathrm{d}t\right| + \left|\int_{\Omega_T} \nabla(\pi^h u - u)\,.\,\nabla U_\varepsilon^+ \,\mathrm{d}x\,\mathrm{d}t\right|,
\end{aligned} \tag{3.16a}$$

where, on noting (2.20) and (2.17),

$$\left|\int_{\Omega_T} \nabla(U_\varepsilon^+ - \pi^h u) \,.\, \nabla U_\varepsilon^+ \,\mathrm{d}x\,\mathrm{d}t\right| = \left|-\int_0^T (\Delta^h U_\varepsilon^+, U_\varepsilon^+ - \pi^h u)^h \,\mathrm{d}t\right|$$
$$\leq C\,\|\Delta^h U_\varepsilon^+\|_{L^2(\Omega_T)}\,\|U_\varepsilon^+ - \pi^h u\|_{L^2(\Omega_T)}\,. \qquad (3.16b)$$

Combining (3.16a,b), (3.6a), (3.13b), (2.15), (3.9), (3.7a) and (3.14) yields (3.10d).

Finally, the proof of the results (3.11)–(3.12d) for V_ε is exactly the same as the proof of (3.9)–(3.10d) for U_ε. □

REMARK. 3.1 The conditions $u^0 \in H^2(\Omega)$ with $\frac{\partial u^0}{\partial \nu} = 0$ on $\partial\Omega$ for the results (3.10a–d), and similarly for v^0, can be replaced by a restriction on τ_1 in terms of h, see Barrett and Nürnberg (2004, Lemma 3.1), but they are not particularly restrictive. The assumption (3.8) holds if $U_\varepsilon(x,t) = 1$ for all $x \in \partial_2\Omega$ and $t \in [0,T]$, and this condition held in all our numerical experiments provided $u^0 = 1$ on $\partial_2\Omega$ and either L_1 is chosen sufficiently large or T is chosen sufficiently small. This can be made rigorous for the approximation $(\widetilde{\mathrm{P}}_\varepsilon^{h,\tau})$, see Remark 2.1, as the degeneracy of $\widetilde{\Xi}_\varepsilon$ leads to finite speed of propagation of the numerical material interfacial region, $|U_\varepsilon| < 1$; at each time level it can move locally at most one mesh point, see Barrett, Blowey, and Garcke (1999).

From (3.13a), (2.22a,b), (2.21), (1.8) and (3.10c) we see that we can only control $\nabla\Phi_\varepsilon^+$ and ∇W_ε^+ on the sets where $\Lambda_\varepsilon(U_\varepsilon^-)$ and $\Xi_\varepsilon(U_\varepsilon^-)$ are bounded below independently of ε, and hence h on noting (ii), i.e. on the sets where $u > -1$ and $|u| < 1$, respectively. Therefore in order to construct the appropriate limits as $h \to 0$, we introduce the following open subsets of $\overline{\Omega}$. For any $\delta \in (0,1)$, we define for *a.a.* $t \in (0,T)$

$$B_\delta(t) := \{\, x \in \overline{\Omega} : |u(x,t)| < 1-\delta \,\} \subset D_\delta(t) := \{\, x \in \overline{\Omega} : -1+\delta < u(x,t) \,\}\,, \quad (3.17a)$$
$$B_{\delta,I}(t) := \{\, x \in B_\delta(t) : |v(x,t)| < \tfrac{1}{\sqrt{3}}(1 + u(x,t) - \delta) \,\}\,, \quad (3.17b)$$
$$B_{\delta,+}(t) := \{\, x \in B_\delta(t) : v(x,t) - \tfrac{1}{\sqrt{3}}(1 + u(x,t)) \in [-\tfrac{\delta}{\sqrt{3}}, 0] \,\}\,, \quad (3.17c)$$
$$B_{\delta,-}(t) := \{\, x \in B_\delta(t) : v(x,t) + \tfrac{1}{\sqrt{3}}(1 + u(x,t)) \in [0, \tfrac{\delta}{\sqrt{3}}] \,\}\,. \quad (3.17d)$$

From (3.10c) and (3.12c) we have that there exist positive constants $C_x(t)$ such that

$$|u(y_1,t) - u(y_2,t)| + |v(y_1,t) - v(y_2,t)| \leq C_x(t)\,|y_1 - y_2|^\varsigma$$
$$\forall\; y_1,\, y_2 \in \overline{\Omega} \quad \text{for } a.a.\ t \in (0,T). \quad (3.18)$$

As $⨍ u(\cdot,t) = ⨍ u^0 \in (-1,1)$ for *a.a.* $t \in (0,T)$, it follows that there exists a $\delta_0 \in (0, 1 - |⨍ u^0|)$ such that $D_{\delta_0}(t) \supset B_{\delta_0}(t) \not\equiv \emptyset$ for *a.a.* $t \in (0,T)$. It immediately follows from (3.17a–d) and (3.18) for *a.a.* $t \in (0,T)$ and for any δ_1, $\delta_2 \in (0,\delta_0)$ with $\delta_1 > \delta_2$ that

$$\text{either}\quad y_1 \in B_{\delta_1}(t) \text{ and } y_2 \in \partial B_{\delta_2}(t) \quad\text{or}\quad y_1 \in D_{\delta_1}(t) \text{ and } y_2 \in \partial D_{\delta_2}(t) \;\text{ with } y_2 \notin \partial\Omega$$
$$\Longrightarrow\; C_x(t)\,|y_1 - y_2|^\varsigma \geq |u(y_1,t) - u(y_2,t)| > (\delta_1 - \delta_2), \quad (3.19a)$$
$$\text{and}\quad y_1 \in B_{\delta_1,I}(t) \text{ and } y_2 \in \partial B_{\delta_2,I}(t) \;\text{ with } y_2 \notin \partial\Omega$$
$$\Longrightarrow\; C_x(t)\,|y_1 - y_2|^\varsigma \geq |u(y_1,t) - u(y_2,t)| + |v(y_1,t) - v(y_2,t)| > \tfrac{1}{\sqrt{3}}\,(\delta_1 - \delta_2); \quad (3.19b)$$

where $\partial B_\delta(t)$, $\partial D_\delta(t)$ and $\partial B_{\delta,I}(t)$ are the boundaries of $B_\delta(t)$, $D_\delta(t)$ and $B_{\delta,I}(t)$, respectively. This implies that for $a.a.$ $t \in (0,T)$ and any $\delta \in (0,\delta_0)$, there exists an $h_0(\delta,t)$ such that for all $h \leq h_0(\delta,t)$ there exist collections of simplices $\mathcal{T}^h_{B,\delta,I}(t) \subset \mathcal{T}^h_{B,\delta}(t) \subset \mathcal{T}^h_{D,\delta}(t) \subset \mathcal{T}^h$ such that

$$B_\delta(t) \subset B^h_\delta(t) := \cup_{\sigma\in\mathcal{T}^h_{B,\delta}(t)}\,\overline{\sigma} \subset B_{\frac{\delta}{2}}(t)\,, \quad D_\delta(t) \subset D^h_\delta(t) := \cup_{\sigma\in\mathcal{T}^h_{D,\delta}(t)}\,\overline{\sigma} \subset D_{\frac{\delta}{2}}(t)\,, \tag{3.20a}$$

$$B_{\delta,I}(t) \subset B^h_{\delta,I}(t) := \cup_{\sigma\in\mathcal{T}^h_{B,\delta,I}(t)}\,\overline{\sigma} \subset B_{\frac{\delta}{2},I}(t)\,. \tag{3.20b}$$

Clearly, we have from (3.17a,b) that

$$\delta_2 < \delta_1 < \delta_0 \qquad \Longrightarrow \qquad h_0(\delta_2,t) \leq h_0(\delta_1,t)\,.$$

For $a.a.$ $t \in (0,T)$ and any fixed $\delta \in (0,\widehat{\delta}_0)$, where $\widehat{\delta}_0 := \min\{\delta_0,\frac{1}{2}\}$, it follows from (3.17a–d), (3.10c), (3.12c) and our assumption (ii) of Lemma 3.1 that there exists an $\widehat{h}_0(\delta,t) \leq h_0(\delta,t)$ such that for $h \leq \widehat{h}_0(\delta,t)$

$$1-2\,\delta \leq |U^\pm_\varepsilon(x,t)| \quad \forall\, x \notin B_\delta(t), \quad |U^\pm_\varepsilon(x,t)| < 1-\tfrac{\delta}{2} \qquad \forall\, x \in B_\delta(t)\,, \tag{3.21a}$$

$$U^\pm_\varepsilon(x,t) \leq -1+2\,\delta \quad \forall\, x \notin D_\delta(t), \qquad -1+\tfrac{\delta}{2} < U^\pm_\varepsilon(x,t) \quad \forall\, x \in D_\delta(t)\,; \tag{3.21b}$$

$$|V^\pm_\varepsilon(x,t)| < \tfrac{1}{\sqrt{3}}\,(1+U^\pm_\varepsilon(x,t)\,) \qquad\qquad \forall\, x \in B_{\delta,I}(t)\,, \tag{3.22a}$$

$$V^\pm_\varepsilon(x,t) - \tfrac{1}{\sqrt{3}}(1+U^\pm_\varepsilon(x,t)\,) \in [-\tfrac{2\,\delta}{\sqrt{3}},0] \quad \forall\, x \in B_{\delta,+}(t)\,, \tag{3.22b}$$

$$V^\pm_\varepsilon(x,t) + \tfrac{1}{\sqrt{3}}(1+U^\pm_\varepsilon(x,t)\,) \in [0,\tfrac{2\,\delta}{\sqrt{3}}] \qquad \forall\, x \in B_{\delta,-}(t)\,; \tag{3.22c}$$

and

$$\varepsilon \leq \delta\,. \tag{3.23}$$

Lemma. 3.2 *Let all the assumptions of Lemma* 3.1 *hold. Then for a.a.* $t \in (0,T)$ *there exist functions*

$$\phi(\cdot,t) \in H^1_{loc}(\{u(\cdot,t) > -1\}), \;\; w(\cdot,t) \in H^1_{loc}(\{|u(\cdot,t)| < 1\}); \tag{3.24}$$

where $\{u(\cdot,t) > -1\} := \{x \in \Omega : u(x,t) > -1\}$ *and* $\{|u(\cdot,t)| < 1\} := \{x \in \Omega : |u(x,t)| < 1\}$*; such that on extracting a further subsequence from the subsequence* $\{\Phi^+_\varepsilon, U_\varepsilon, V_\varepsilon, W^+_\varepsilon, Z^+_\varepsilon\}_h$ *in Lemma* 3.1*, it holds as* $h \to 0$ *that*

$$\Lambda_\varepsilon(U^-_\varepsilon)\,\nabla\Phi^+_\varepsilon \to \mathcal{H}_{\{u>-1\}}\,c(u)\,\nabla\phi \qquad \textit{weakly in } L^2(\Omega_T), \tag{3.25a}$$

$$\Xi_\varepsilon(U^-_\varepsilon)\,\nabla\Phi^+_\varepsilon \to \mathcal{H}_{\{|u|<1\}}\,b(u)\,\nabla\phi \qquad \textit{weakly in } L^2(\Omega_T), \tag{3.25b}$$

$$\Xi_\varepsilon(U^-_\varepsilon)\,\nabla W^+_\varepsilon \to \mathcal{H}_{\{|u|<1\}}\,b(u)\,\nabla w \qquad \textit{weakly in } L^2(\Omega_T); \tag{3.25c}$$

where $\mathcal{H}_{\{u>-1\}}$ *and* $\mathcal{H}_{\{|u|<1\}}$ *are the characteristic functions of the sets* $\{u > -1\} := \{(x,t) \in \Omega_T : u(x,t) > -1\}$ *and* $\{|u| < 1\} := \{(x,t) \in \Omega_T : |u(x,t)| < 1\}$*, respectively.*

Moreover for a.a. $t \in (0,T)$, $\{u(\cdot,t), v(\cdot,t)\} \in K$ *and* $w(\cdot,t)$, $z(\cdot,t)$ *satisfy*

$$\int_{\{|u(\cdot,t)|<1\}} [\gamma \nabla u \,.\, \nabla(\eta_1 - u) + (\gamma^{-1} \Psi_{,u}(u,v) - w)\,(\eta_1 - u)]\,\mathrm{d}x$$
$$+ \int_\Omega [\gamma \nabla v \,.\, \nabla(\eta_2 - v) + (\gamma^{-1} \Psi_{,v}(u,v) - z)\,(\eta_2 - v)]\,\mathrm{d}x \geq 0$$
$$\forall\ \{\eta_1, \eta_2\} \in K \text{ with } \operatorname{supp}(\eta_1 - u) \subset \{|u(\cdot,t)| < 1\}. \tag{3.26}$$

Finally if $\alpha \neq 0$, *on assuming that*

$$u(x,t) = 1 \qquad \forall\ x \in \partial_2\Omega, \quad \textit{for a.a. } t \in (0,T)\,; \tag{3.27}$$

it follows as $h \to 0$ *that*

$$\Phi_\varepsilon^+ \to \phi \qquad \textit{weakly in } L^2(0,T; L^2(\partial_2\Omega))\,. \tag{3.28}$$

Proof. This lemma is a generalisation of Lemma 3.4 in Barrett, Nürnberg, and Styles (2004). The proof of the results (3.24) for ϕ, (3.25a,b) and (3.28) can be found there, on using the results on $D_\delta(t)$ in (3.17a), (3.20a) and (3.21b). The key difference here is the identification of w on $\{|u| < 1\}$ via the variational inequality (3.26), which is now more delicate to establish. On recalling (3.9), (3.11) and (1.6), let

$$a^u := -\gamma\,\Delta u + \gamma^{-1}\,\Psi_{,u}(u,v), \quad a^v := -\gamma\,\Delta v + \gamma^{-1}\,\Psi_{,v}(u,v) \in L^2(\Omega_T).$$

For *a.a.* $t \in (0,T)$, we define $w(\cdot,t)$ on $\{|u(\cdot,t)| < 1\}$ such that

$$w(\cdot,t) \equiv \begin{cases} a^u(\cdot,t) - \frac{1}{\sqrt{3}}(a^v(\cdot,t) - z(\cdot,t)) & \text{if} \quad v(\cdot,t) \in [-\frac{1}{\sqrt{3}}(1 + u(\cdot,t)), 0)\,, \\ a^u(\cdot,t) & \text{if} \quad v(\cdot,t) = 0\,, \\ a^u(\cdot,t) + \frac{1}{\sqrt{3}}(a^v(\cdot,t) - z(\cdot,t)) & \text{if} \quad v(\cdot,t) \in (0, \frac{1}{\sqrt{3}}(1 + u(\cdot,t))]\,. \end{cases} \tag{3.29}$$

We will deduce below that for *a.a.* $t \in (0,T)$

$$a^v(\cdot,t) \equiv z(\cdot,t) \quad \text{if} \quad |v(\cdot,t)| < \tfrac{1}{\sqrt{3}}\,(1 + u(\cdot,t))\,. \tag{3.30}$$

It follows from (3.13a) and (2.22b) that

$$\|\Xi_\varepsilon(U_\varepsilon^-)\nabla W_\varepsilon^+\|^2_{L^2(\Omega_T)} \leq C. \tag{3.31}$$

Hence (3.31) implies that there exists a vector function $f \in L^2(\Omega_T)$, and on extracting a further subsequence from the subsequence $\{\Phi_\varepsilon^+, U_\varepsilon, V_\varepsilon, W_\varepsilon^+, Z_\varepsilon^+\}_h$ in Lemma 3.1, it holds as $h \to 0$ that

$$\Xi_\varepsilon(U_\varepsilon^-)\nabla W_\varepsilon^+ \to f \quad \text{weakly in } L^2(\Omega_T)\,. \tag{3.32}$$

We now identify the function f.

First, we consider a fixed $\delta \in (0, \widehat{\delta}_0)$. It follows from (1.8), (2.21), (2.22b), (3.21a) and (3.13a) that for *a.a.* $t \in (0,T)$ and for all $h \leq \widehat{h}_0(\delta, t)$

$$\begin{aligned} \delta\,(1-\tfrac{\delta}{4})\,|\nabla W_\varepsilon^+(\cdot,t)|^2_{0,B_\delta(t)} &= b(1-\tfrac{\delta}{2})\,|\nabla W_\varepsilon^+(\cdot,t)|^2_{0,B_\delta(t)} \leq b_\varepsilon(1-\tfrac{\delta}{2})\,|\nabla W_\varepsilon^+(\cdot,t)|^2_{0,B_\delta(t)} \\ &\leq |\,(\,[\Xi_\varepsilon(U_\varepsilon^-)]^{\frac{1}{2}}\,\nabla W_\varepsilon^+)(\cdot,t)|^2_0 \leq C(t)\,. \end{aligned} \tag{3.33}$$

From (3.33), (3.20a), (2.22b), (3.21b) and (3.23) we have for *a.a.* $t \in (0,T)$ and for all $h \leq \widehat{h}_0(\delta, t)$

$$\begin{aligned} |\,(\Xi_\varepsilon(U_\varepsilon^-)\,\nabla W_\varepsilon^+)(\cdot,t)|^2_{0,\Omega\setminus B_\delta(t)} &\leq \max_{x\in\Omega\setminus B_{2\delta}(t)} b_\varepsilon(U_\varepsilon^-(x))\,|\,(\,[\Xi_\varepsilon(U_\varepsilon^-)]^{\frac{1}{2}}\,\nabla W_\varepsilon^+)(\cdot,t)|^2_{0,\Omega\setminus B_\delta(t)} \\ &\leq C(t)\,b_\varepsilon(1-4\,\delta) \leq C(t)\,\max\{4\delta,\varepsilon\} \leq C(t)\,\delta\,. \end{aligned} \tag{3.34}$$

On noting (3.13b) we have for *a.a.* $t \in (0,T)$ that

$$|\Delta^h U_\varepsilon^+(\cdot,t)|_0 + |\Delta^h V_\varepsilon^+(\cdot,t)|_0 \leq C(t)\,. \tag{3.35}$$

This yields for *a.a.* $t \in (0,T)$ that as $h \to 0$

$$\Delta^h U_\varepsilon^+(\cdot,t) \to \Delta u(\cdot,t), \qquad \Delta^h V_\varepsilon^+(\cdot,t) \to \Delta v(\cdot,t) \qquad \text{weakly in } L^2(\Omega); \tag{3.36}$$

see Barrett, Nürnberg, and Styles (2004, (3.18)) for details. Recalling the notation (2.36a,b), we have from cases (i) $A_j^U = A_j^V = 0$, (iii) $A_j^U + \frac{1}{\sqrt{3}}\,A_j^V = 0$ and (iv) $A_j^U - \frac{1}{\sqrt{3}}\,A_j^V = 0$ in the proof of Lemma 2.3, on noting (2.20), (3.1b), (3.22a–c) and (3.20b), that for *a.a.* $t \in (0,T)$ and for all $h \leq \widehat{h}_0(\frac{\delta}{2}, t)$

$$\begin{aligned} W_\varepsilon^+(\cdot,t) &\equiv -\gamma\,\Delta^h U_\varepsilon^+(\cdot,t) + \gamma^{-1}\,\Psi_{,u}(U_\varepsilon^-(\cdot,t), V_\varepsilon^-(\cdot,t))\,, \\ \text{and} \qquad Z_\varepsilon^+(\cdot,t) &\equiv -\gamma\,\Delta^h V_\varepsilon^+(\cdot,t) + \gamma^{-1}\,\Psi_{,v}(U_\varepsilon^-(\cdot,t), V_\varepsilon^-(\cdot,t)) \quad \text{on } B_{\delta,I}(t)\,; \end{aligned} \tag{3.37a}$$

$$\begin{aligned} W_\varepsilon^+(\cdot,t) \pm \tfrac{1}{\sqrt{3}}\,Z_\varepsilon^+(\cdot,t) &\equiv [-\gamma\,\Delta^h U_\varepsilon^+(\cdot,t) + \gamma^{-1}\,\Psi_{,u}(U_\varepsilon^-(\cdot,t), V_\varepsilon^-(\cdot,t))] \\ &\quad \pm \tfrac{1}{\sqrt{3}}\,[-\gamma\,\Delta^h V_\varepsilon^+(\cdot,t) + \gamma^{-1}\,\Psi_{,v}(U_\varepsilon^-(\cdot,t), V_\varepsilon^-(\cdot,t))] \\ &\qquad \text{on } B_{\delta,\pm}(t)\,. \end{aligned} \tag{3.37b}$$

It follows from (3.37a,b), (3.36), (3.10c), (3.12c) and (3.6c) for *a.a.* $t \in (0,T)$ that as $h \to 0$

$$\begin{aligned} &W_\varepsilon^+(\cdot,t) \to a^u(\cdot,t), \quad Z_\varepsilon^+(\cdot,t) \to a^v(\cdot,t) \equiv z(\cdot,t) &&\text{weakly in } L^2(B_{\delta,I}(t))\,, \\ &W_\varepsilon^+(\cdot,t) \to a^u(\cdot,t) \pm \tfrac{1}{\sqrt{3}}\,(a^v(\cdot,t) - z(\cdot,t)\,) &&\text{weakly in } L^2(B_{\delta,\pm}(t))\,. \end{aligned}$$

This together with (3.29) and (3.33) yields that

$$W_\varepsilon^+(\cdot,t) \to w(\cdot,t) \quad \text{weakly in } H^1(B_\delta(t)). \tag{3.38}$$

Combining (3.32), (3.38) and (3.7b) yields for *a.a.* $t \in (0,T)$ that as $h \to 0$

$$(\Xi_\varepsilon(U_\varepsilon^-)\nabla W_\varepsilon^+)(\cdot,t) \to b(u(\cdot,t))\,\nabla w(\cdot,t) \qquad \text{weakly in } L^2(B_\delta(t)).$$

We now work on establishing the variational inequality (3.26). For $a.a.\ t \in (0,T)$, let $\{\eta_1, \eta_2\} \in K$ with $\eta_1(\cdot) \equiv u(\cdot,t) + \xi(\cdot)$ and $\operatorname{supp} \xi \subset B_{3\delta}(t)$. For the ensuing analysis it is necessary to prescribe the following extensions in order to control the support of a mollified version of ξ, see (3.40) below. Let $\widetilde{\Omega} := (-\widetilde{L}_1, \widetilde{L}_1) \times (-\widetilde{L}_2, \widetilde{L}_2)$, where $\widetilde{L}_i := \frac{3}{2} L_i$. By reflection about $x_i = \pm L_i$, $i = 1 \to 2$, there exist extensions $\widetilde{u}(\cdot,t)$, $\widetilde{\xi} \in H^1(\widetilde{\Omega})$ and $\{\widetilde{\eta}_1, \widetilde{\eta}_2\} \in [H^1(\widetilde{\Omega})]^2$ such that $\{\widetilde{\eta}_1(x), \widetilde{\eta}_2(x)\} \in \mathcal{K}$ for $a.e.\ x \in \widetilde{\Omega}$, $\widetilde{\eta}_1(\cdot) \equiv \widetilde{u}(\cdot,t) + \widetilde{\xi}(\cdot)$ with $\operatorname{supp} \widetilde{\xi} \subset \{x \in \overline{\widetilde{\Omega}} : |\widetilde{u}(x,t)| < 1 - 3\delta\}$, and $\widetilde{\eta}_i \,|_\Omega \equiv \eta_i$, $\widetilde{u}(\cdot,t)\,|_\Omega \equiv u(\cdot,t)$, $\widetilde{\xi}\,|_\Omega \equiv \xi$. Applying the standard Friedrichs mollifier to $\widetilde{\eta}_i$, $\widetilde{u}(\cdot,t)$ and $\widetilde{\xi}$, there exist $C_0^\infty(\mathbb{R}^2)$ functions with their restrictions to $\overline{\Omega}$ satisfying

$$\begin{aligned} &\{\eta_1^{(\ell)}, \eta_2^{(\ell)}\} \in K \cap [C^\infty(\overline{\Omega})]^2 \quad \text{such that} \quad \eta_1^{(\ell)}(\cdot) \equiv u^{(\ell)}(\cdot,t) + \xi^{(\ell)}(\cdot) \text{ in } C^\infty(\overline{\Omega}) \\ \text{with} \quad &\eta_i^{(\ell)} \to \eta_i, \quad u^{(\ell)}(\cdot,t) \to u(\cdot,t), \quad \xi^{(\ell)} \to \xi \quad \text{strongly in } H^1(\Omega) \text{ as } \ell \to \infty. \end{aligned} \tag{3.39}$$

Moreover, there exists an $\ell_0(\delta) \in \mathbb{N}$ such that

$$\operatorname{supp} \xi^{(\ell)} \subset B_{2\delta}(t) \qquad \forall\ \ell \geq \ell_0(\delta). \tag{3.40}$$

It follows that $\{\chi_1^{(\ell)}, \chi_2^{(\ell)}\} \in K^h$, where

$$\chi_1^{(\ell)}(\cdot) \equiv U_\varepsilon^+(\cdot,t) + \mathcal{R}^1_{U_\varepsilon^+}(\pi^h \xi^{(\ell)}(\cdot)), \quad \chi_2^{(\ell)} \equiv \mathcal{R}^2_{U_\varepsilon^+}(\pi^h \eta_2^{(\ell)}),$$

and $\mathcal{R}^i_{U_\varepsilon^+} : S^h \to S^h$, $i = 1 \to 2$, are such that for all $\chi \in S^h$ and for all $j \in J$

$$[\mathcal{R}^1_{U_\varepsilon^+}(\chi)](p_j) := \begin{cases} \chi(p_j) & \text{if } |U_\varepsilon^+(p_j,t) + \chi(p_j)| \leq 1\,, \\ 1 - U_\varepsilon^+(p_j,t) & \text{if } U_\varepsilon^+(p_j,t) + \chi(p_j) > 1\,, \\ -1 - U_\varepsilon^+(p_j,t) & \text{if } U_\varepsilon^+(p_j,t) + \chi(p_j) < -1\,; \end{cases} \tag{3.41}$$

and

$$[\mathcal{R}^2_{U_\varepsilon^+}(\chi)](p_j) := \begin{cases} \chi(p_j) & \text{if } |\chi(p_j)| \leq \frac{1}{\sqrt{3}}(1 + \chi_1^{(\ell)}(p_j))\,, \\ \frac{1}{\sqrt{3}}(1 + \chi_1^{(\ell)}(p_j)) & \text{if } \chi(p_j) > \frac{1}{\sqrt{3}}(1 + \chi_1^{(\ell)}(p_j))\,, \\ -\frac{1}{\sqrt{3}}(1 + \chi_1^{(\ell)}(p_j)) & \text{if } \chi(p_j) < -\frac{1}{\sqrt{3}}(1 + \chi_1^{(\ell)}(p_j))\,. \end{cases} \tag{3.42}$$

We note from (3.41), (3.20a) and (3.40) that for all $\ell \geq \ell_0(\delta)$ and for all $h \leq h_0(2\delta, t)$

$$\operatorname{supp} \mathcal{R}^1_{U_\varepsilon^+}(\pi^h \xi^{(\ell)}) \subset \operatorname{supp} \pi^h \xi^{(\ell)} \subset B_\delta(t)\,. \tag{3.43}$$

Moreover, it follows from (3.41), (3.42) and (3.39) that

$$|\pi^h \xi^{(\ell)} - \mathcal{R}^1_{U_\varepsilon^+}(\pi^h \xi^{(\ell)})|_h \leq |\pi^h u^{(\ell)}(\cdot,t) - U_\varepsilon^+(\cdot,t)|_h\,, \tag{3.44a}$$

$$\begin{aligned} |\pi^h \eta_2^{(\ell)} - \mathcal{R}^2_{U_\varepsilon^+}(\pi^h \eta_2^{(\ell)})|_h &\leq \tfrac{1}{\sqrt{3}} \left[|\pi^h u^{(\ell)}(\cdot,t) - U_\varepsilon^+(\cdot,t)|_h + |\pi^h \xi^{(\ell)} - \mathcal{R}^1_{U_\varepsilon^+}(\pi^h \xi^{(\ell)})|_h \right] \\ &\leq \tfrac{2}{\sqrt{3}}\, |\pi^h u^{(\ell)}(\cdot,t) - U_\varepsilon^+(\cdot,t)|_h\,. \end{aligned} \tag{3.44b}$$

We now choose $\{\chi_1, \chi_2\} \equiv \{\chi_1^{(\ell)}, \chi_2^{(\ell)}\}$ in (3.4d) and analyse the subsequent terms. First, we have from (2.20), (3.35), (3.44a) and (2.17) that $U_\varepsilon^+(\cdot,t)$, $\chi_1^{(\ell)}(\cdot)$ and $\xi^{(\ell)}(\cdot)$ satisfy

$$\begin{aligned}|(\nabla U_\varepsilon^+, \nabla(\chi_1^{(\ell)} - U_\varepsilon^+)) - (\nabla U_\varepsilon^+, \nabla(\pi^h \xi^{(\ell)}))| &= |(\Delta^h U_\varepsilon^+, (I - \mathcal{R}^1_{U_\varepsilon^+})(\pi^h \xi^{(\ell)}))^h| \\ &\leq C(t)\, |\pi^h u^{(\ell)}(\cdot,t) - U_\varepsilon^+(\cdot,t)|_0 . \qquad (3.45)\end{aligned}$$

Similarly to (3.45), we deduce from (2.20), (3.35), (3.44b) and (2.17) that $V_\varepsilon^+(\cdot,t)$, $\chi_2^{(\ell)}(\cdot)$ and $\eta_2^{(\ell)}(\cdot)$ satisfy

$$|(\nabla V_\varepsilon^+, \nabla(\chi_2^{(\ell)} - V_\varepsilon^+)) - (\nabla V_\varepsilon^+, \nabla(\pi^h \eta_2^{(\ell)} - V_\varepsilon^+))| \leq C(t)\, |\pi^h u^{(\ell)}(\cdot,t) - U_\varepsilon^+(\cdot,t)|_0 . \quad (3.46)$$

Next, it follows from (3.43), (2.17), (3.44a), (1.6), (2.18), (3.38) and (2.15) that $U_\varepsilon^\pm(\cdot,t)$, $V_\varepsilon^\pm(\cdot,t)$, $W_\varepsilon^+(\cdot,t)$, $\chi_1^{(\ell)}(\cdot)$ and $\xi^{(\ell)}(\cdot)$ satisfy

$$\begin{aligned}&|(W_\varepsilon^+ - \gamma^{-1}\Psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \chi_1^{(\ell)} - U_\varepsilon^+)^h - (W_\varepsilon^+ - \gamma^{-1}\Psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \pi^h \xi^{(\ell)})| \\ &\qquad \leq |(W_\varepsilon^+ - \gamma^{-1}\Psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), (I - \mathcal{R}^1_{U_\varepsilon^+})(\pi^h \xi^{(\ell)}))^h| \\ &\qquad\qquad + |(W_\varepsilon^+ - \gamma^{-1}\Psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \pi^h \xi^{(\ell)}) - (W_\varepsilon^+ - \gamma^{-1}\Psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \pi^h \xi^{(\ell)})^h| \\ &\qquad \leq C\,[1 + |W_\varepsilon|_{0,B_\delta(t)}]\,\left[|\pi^h u^{(\ell)}(\cdot,t) - U_\varepsilon^+(\cdot,t)|_0 + h\, |\pi^h \xi^{(\ell)}|_1\right] \\ &\qquad \leq C(t)\,\left[|\pi^h u^{(\ell)}(\cdot,t) - U_\varepsilon^+(\cdot,t)|_0 + h\, |\xi^{(\ell)}|_2\right] . \qquad (3.47)\end{aligned}$$

Similarly, it follows from (3.44b), (2.17), (2.18), (3.13a) and (2.15) that $U_\varepsilon^\pm(\cdot,t)$, $V_\varepsilon^\pm(\cdot,t)$, $Z_\varepsilon^+(\cdot,t)$, $\chi_2^{(\ell)}(\cdot)$ and $\eta_2^{(\ell)}(\cdot)$ satisfy

$$\begin{aligned}&|(Z_\varepsilon^+ - \gamma^{-1}\Psi_{,v}(U_\varepsilon^-, V_\varepsilon^-), \chi_2^{(\ell)} - V_\varepsilon^+)^h - (Z_\varepsilon^+ - \gamma^{-1}\Psi_{,v}(U_\varepsilon^-, V_\varepsilon^-), \pi^h \eta_2^{(\ell)} - V_\varepsilon^+)| \\ &\qquad \leq C(t)\,\left[|\pi^h u^{(\ell)}(\cdot,t) - U_\varepsilon^+(\cdot,t)|_0 + h\,(1 + |\eta_2^{(\ell)}|_2)\right] . \qquad (3.48)\end{aligned}$$

Combining (3.45)–(3.48), noting (3.4d), (3.10d), (3.12d), (3.38), (3.6c), (2.15) and letting $h \to 0$, we obtain, on possibly extracting another subsequence from $\{\Phi_\varepsilon^+, U_\varepsilon, V_\varepsilon, W_\varepsilon^+, Z_\varepsilon^+\}_h$, that $u(\cdot,t)$, $v(\cdot,t)$, $w(\cdot,t)$, $z(\cdot,t)$ and $u^{(\ell)}(\cdot,t)$, $\eta_i^{(\ell)}(\cdot)$ satisfy

$$\begin{aligned}&\int_{B_\delta(t)} [\,\gamma\, \nabla u \,.\, \nabla(\eta_1^{(\ell)} - u^{(\ell)}) + (\gamma^{-1}\, \Psi_{,u}(u,v) - w)\,(\eta_1^{(\ell)} - u^{(\ell)})]\,\mathrm{d}x \\ &\qquad + \int_\Omega [\gamma\, \nabla v \,.\, \nabla(\eta_2^{(\ell)} - v) + (\gamma^{-1}\, \Psi_{,v}(u,v) - z)\,(\eta_2^{(\ell)} - v)]\,\mathrm{d}x \geq r^{(\ell)}(t), \qquad (3.49)\end{aligned}$$

where $|r^{(\ell)}(t)| \leq C\, |(u - u^{(\ell)})(\cdot,t)|_0$. Letting $\ell \to \infty$ in (3.49), it follows from (3.39) that $u(\cdot,t)$, $v(\cdot,t)$, $w(\cdot,t)$, $z(\cdot,t)$ and $\eta_i(\cdot)$ satisfy

$$\begin{aligned}&\int_{B_\delta(t)} [\,\gamma\, \nabla u \,.\, \nabla(\eta_1 - u) + (\gamma^{-1}\, \Psi_{,u}(u,v) - w)\,(\eta_1 - u)]\,\mathrm{d}x \\ &\qquad + \int_\Omega [\gamma\, \nabla v \,.\, \nabla(\eta_2 - v) + (\gamma^{-1}\, \Psi_{,v}(u,v) - z)\,(\eta_2 - v)]\,\mathrm{d}x \geq 0 . \qquad (3.50)\end{aligned}$$

Repeating (3.33), (3.34) and (3.37a)–(3.50) for all $\delta \in (0, \widehat{\delta}_0)$ yields, on recalling (3.10c), that (3.24) for w, and (3.26) hold; and, on noting (3.34) and (3.32), the desired result (3.25c). In addition, we deduce the identity (3.30). Of course, the identities (3.29) and (3.30) can be deduced from the derived variational inequality (3.26); and hence, their omission in the statement of the Lemma. □

REMARK. 3.2 The assumption (3.27) is similar to the assumption (3.8), see Remark 3.1.

THEOREM. 3.1 *Let the assumptions of Lemma* 3.2 *hold. Then there exists a subsequence of* $\{\Phi_\varepsilon^+, U_\varepsilon, V_\varepsilon, W_\varepsilon^+, Z_\varepsilon^+\}_h$, *where* $\{\Phi_\varepsilon^+, U_\varepsilon, V_\varepsilon, W_\varepsilon^+, Z_\varepsilon^+\}$ *solve* $(\mathrm{P}_\varepsilon^{h,\tau})$, *and functions* $\{\phi, u, v, w, z\}$ *satisfying* (3.5a,b), (3.9), (3.11) *and* (3.24). *In addition, as* $h \to 0$ *the following hold:* (3.6a–c), (3.7a,b), (3.10a–d), (3.12a–d) *and* (3.28–d). *Furthermore, we have that* $\{\phi, u, v, w, z\}$ *fulfil* $u(\cdot, 0) = u^0(\cdot)$, $v(\cdot, 0) = v^0(\cdot)$ *in* $L^2(\Omega)$ *and* $⨍ u(\cdot, t) = ⨍ u^0$ *for a.a.* $t \in (0, T)$. *Moreover, they satisfy for all* $\eta \in L^2(0, T; H^1(\Omega))$

$$\int_{\{u>-1\}} c(u)\, \nabla\phi \,.\, \nabla\eta \, \mathrm{d}x\, \mathrm{d}t + \int_0^T \int_{\partial_2\Omega} \phi\, \eta \, \mathrm{d}s\, \mathrm{d}t = \int_0^T \int_{\partial_2\Omega} g\, \eta \, \mathrm{d}s\, \mathrm{d}t\,, \tag{3.51a}$$

$$\gamma \int_0^T \langle \tfrac{\partial u}{\partial t}, \eta \rangle \, \mathrm{d}t + \int_{\{|u|<1\}} b(u)\, \nabla\left[w + \alpha\, \phi\right] .\, \nabla\eta \, \mathrm{d}x\, \mathrm{d}t = 0\,, \tag{3.51b}$$

$$\ell(\gamma) \int_0^T (\tfrac{\partial v}{\partial t}, \eta) \, \mathrm{d}t + \int_0^T (z, \eta) \, \mathrm{d}t = 0\,; \tag{3.51c}$$

where for a.a. $t \in (0, T)$, $\{u(\cdot, t), v(\cdot, t)\} \in K$ *and* $w(\cdot, t)$, $z(\cdot, t)$ *satisfy*

$$\begin{aligned}\int_{\{|u(\cdot,t)|<1\}} &[\gamma\, \nabla u \,.\, \nabla(\eta_1 - u) + (\gamma^{-1}\, \Psi_{,u}(u, v) - w)\, (\eta_1 - u)] \, \mathrm{d}x \\ &+ \int_\Omega [\gamma\, \nabla v \,.\, \nabla(\eta_2 - v) + (\gamma^{-1}\, \Psi_{,v}(u, v) - z)\, (\eta_2 - v)] \, \mathrm{d}x \geq 0 \\ &\qquad \forall\ \{\eta_1, \eta_2\} \in K \text{ with } \operatorname{supp}(\eta_1 - u) \subset \{|u(\cdot, t)| < 1\}.\end{aligned} \tag{3.52}$$

Proof. Only (3.51a–c) need to be established, as (3.52) was established in Lemma 3.2 above. The proof of (3.51a,b) can be found in Barrett, Nürnberg, and Styles (2004, Theorem 3.6), and (3.51c) is similarly established. □

4 Solution of the discrete system

We now discuss algorithms for solving the resulting system of algebraic equations for $\{\Phi_\varepsilon^n, U_\varepsilon^n, V_\varepsilon^n, W_\varepsilon^n, Z_\varepsilon^n\}$ arising at each time level from the approximation $(\mathrm{P}_\varepsilon^{h,\tau})$. As (2.14a) in $(\mathrm{P}_\varepsilon^{h,\tau})$ is independent of $\{U_\varepsilon^n, V_\varepsilon^n, W_\varepsilon^n, Z_\varepsilon^n\}$, we solve it first to obtain Φ_ε^n; then solve (2.14b–d) for $\{U_\varepsilon^n, V_\varepsilon^n, W_\varepsilon^n, Z_\varepsilon^n\}$. Solving (2.14a) is straightforward, as it is linear. Adopting the obvious notation, the system (2.14b–d) can be rewritten as: Find $\{\{\underline{U}_\varepsilon^n, \underline{V}_\varepsilon^n\}, \underline{W}_\varepsilon^n,$

$\underline{Z}_\varepsilon^n\} \in \mathcal{K}^\mathcal{J} \times [\mathbb{R}^\mathcal{J}]^2$ such that

$$\gamma\,\mathcal{M}\,\underline{U}_\varepsilon^n + \tau_n\,\mathcal{A}^{n-1}\,\underline{W}_\varepsilon^n = \underline{r}_1\,, \tag{4.1a}$$

$$\ell(\gamma)\,\mathcal{M}\,\underline{V}_\varepsilon^n + \tau_n\,\mathcal{M}\,\underline{Z}_\varepsilon^n = \underline{r}_2\,, \tag{4.1b}$$

$$\begin{aligned}(\underline{\chi}_1 - \underline{U}_\varepsilon^n)^T\,(\gamma\,\mathcal{B}\,\underline{U}_\varepsilon^n - \mathcal{M}\,\underline{W}_\varepsilon^n) + (\underline{\chi}_2 - \underline{V}_\varepsilon^n)^T\,(\gamma\,\mathcal{B}\,\underline{V}_\varepsilon^n - \mathcal{M}\,\underline{Z}_\varepsilon^n)\\ \geq (\underline{\chi}_1 - \underline{U}_\varepsilon^n)^T\,\underline{s}_1 + (\underline{\chi}_2 - \underline{V}_\varepsilon^n)^T\,\underline{s}_2 \qquad \forall\,\{\underline{\chi}_1, \underline{\chi}_2\} \in \mathcal{K}^\mathcal{J};\end{aligned} \tag{4.1c}$$

where $\mathcal{M}$, $\mathcal{B}$ and $\mathcal{A}^{n-1}$ are symmetric $\mathcal{J}\times\mathcal{J}$ matrices, $\mathcal{J} := \#J$, with entries

$$\mathcal{M}_{ij} := (\chi_i, \chi_j)^h, \qquad \mathcal{B}_{ij} := (\nabla\chi_i, \nabla\chi_j), \qquad \mathcal{A}_{ij}^{n-1} := (\Xi_\varepsilon(U_\varepsilon^{n-1})\,\nabla\chi_i, \nabla\chi_j)\,.$$

In addition,

$$\begin{aligned}\underline{r}_1 &:= \gamma\,\mathcal{M}\,\underline{U}_\varepsilon^{n-1} - \alpha\,\tau_n\,\mathcal{A}^{n-1}\,\underline{\Phi}_\varepsilon^n \in \mathbb{R}^\mathcal{J}\,, \qquad &\underline{r}_2 &:= \ell(\gamma)\,\mathcal{M}\,\underline{V}_\varepsilon^{n-1} \in \mathbb{R}^\mathcal{J},\\ \underline{s}_1 &:= -\gamma^{-1}\,\mathcal{M}\,\underline{\Psi_{,u}}(\underline{U}_\varepsilon^{n-1}, \underline{V}_\varepsilon^{n-1}) \in \mathbb{R}^\mathcal{J}, \qquad &\underline{s}_2 &:= -\gamma^{-1}\,\mathcal{M}\,\underline{\Psi_{,v}}(\underline{U}_\varepsilon^{n-1}, \underline{V}_\varepsilon^{n-1}) \in \mathbb{R}^\mathcal{J}\,;\end{aligned}$$

where $[\underline{\Psi_{,\bullet}}(\underline{U}_\varepsilon^{n-1}, \underline{V}_\varepsilon^{n-1})]_j := \Psi_{,\bullet}([\underline{U}_\varepsilon^{n-1}]_j, [\underline{V}_\varepsilon^{n-1}]_j)$. Let $\mathcal{A}^{n-1} \equiv \mathcal{A}_D - \mathcal{A}_L - \mathcal{A}_L^T$, with $\mathcal{A}_L$ and $\mathcal{A}_D$ being the lower triangular and diagonal parts of the matrix $\mathcal{A}^{n-1}$, similarly for $\mathcal{B}$. We use this formulation in constructing our "Gauss–Seidel type" iterative method to solve (4.1a–c).

Given $\{\{U_\varepsilon^{n,0}, V_\varepsilon^{n,0}\}, W_\varepsilon^{n,0}, Z_\varepsilon^{n,0}\} \in K^h \times [S^h]^2$, for $k \geq 1$ find $\{\{U_\varepsilon^{n,k}, V_\varepsilon^{n,k}\}, W_\varepsilon^{n,k}, Z_\varepsilon^{n,k}\} \in K^h \times [S^h]^2$ such that

$$\gamma\,\mathcal{M}\,\underline{U}_\varepsilon^{n,k} + \tau_n\,(\mathcal{A}_D - \mathcal{A}_L)\,\underline{W}_\varepsilon^{n,k} = \underline{r}_1 + \tau_n\,\mathcal{A}_L^T\,\underline{W}_\varepsilon^{n,k-1}\,, \tag{4.2a}$$

$$\ell(\gamma)\,\mathcal{M}\,\underline{V}_\varepsilon^{n,k} + \tau_n\,\mathcal{M}\,\underline{Z}_\varepsilon^{n,k} = \underline{r}_2\,, \tag{4.2b}$$

$$\begin{aligned}(\underline{\chi}_1 - \underline{U}_\varepsilon^{n,k})^T\,(\gamma\,(\mathcal{B}_D - \mathcal{B}_L)\,\underline{U}_\varepsilon^{n,k} - \mathcal{M}\,\underline{W}_\varepsilon^{n,k}) + (\underline{\chi}_2 - \underline{V}_\varepsilon^{n,k})^T\,(\gamma\,(\mathcal{B}_D - \mathcal{B}_L)\,\underline{V}_\varepsilon^{n,k} - \mathcal{M}\,\underline{Z}_\varepsilon^{n,k})\\ \geq (\underline{\chi}_1 - \underline{U}_\varepsilon^{n,k})^T\,(\underline{s}_1 + \gamma\,\mathcal{B}_L^T\,\underline{U}_\varepsilon^{n,k-1}) + (\underline{\chi}_2 - \underline{V}_\varepsilon^{n,k})^T\,(\underline{s}_2 + \gamma\,\mathcal{B}_L^T\,\underline{V}_\varepsilon^{n,k-1}) \quad \forall\,\{\underline{\chi}_1, \underline{\chi}_2\} \in \mathcal{K}^\mathcal{J}.\end{aligned} \tag{4.2c}$$

The above is the natural extension of the iterative method in Barrett, Nürnberg, and Styles (2004) for solving the corresponding nonlinear algebraic system arising from the corresponding finite element approximation of (1.3). Below, we prove convergence of (4.2a–c) for our nonlinear system (2.14b–d) using an energy method.

THEOREM. 4.1 *Let the assumptions* (A) *hold. Then for* $\{\{U_\varepsilon^{n,0}, V_\varepsilon^{n,0}\}, W_\varepsilon^{n,0}, Z_\varepsilon^{n,0}\} \in K^h \times [S^h]^2$ *the sequence* $\{\{U_\varepsilon^{n,k}, V_\varepsilon^{n,k}\}, W_\varepsilon^{n,k}, Z_\varepsilon^{n,k}\}_{k\geq 0}$ *generated by the algorithm* (4.2a–c) *satisfies*

$$\|U_\varepsilon^n - U_\varepsilon^{n,k}\|_1 \to 0, \quad |[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{\frac{1}{2}}\,\nabla(W_\varepsilon^n - W_\varepsilon^{n,k})|_0 \to 0, \tag{4.3a}$$

$$\|V_\varepsilon^n - V_\varepsilon^{n,k}\|_1 \to 0, \quad \textit{and} \quad |Z_\varepsilon^n - Z_\varepsilon^{n,k}|_h \to 0 \quad \textit{as } k \to \infty. \tag{4.3b}$$

Proof. The proof is similar to the proof of Theorem 4.1 in Barrett, Nürnberg, and Styles (2004). Let $\underline{E}^{n,k} := \underline{U}_\varepsilon^n - \underline{U}_\varepsilon^{n,k}$, $\underline{F}^{n,k} := \underline{V}_\varepsilon^n - \underline{V}_\varepsilon^{n,k}$, $\underline{P}^{n,k} := \underline{W}_\varepsilon^n - \underline{W}_\varepsilon^{n,k}$ and

$\underline{Q}^{n,k} := \underline{Z}_\varepsilon^n - \underline{Z}_\varepsilon^{n,k}$. Now subtracting (4.2a) from (4.1a) and testing the resulting equation with $\underline{P}^{n,k}$ yields

$$\gamma\,[\underline{P}^{n,k}]^T\,\mathcal{M}\,\underline{E}^{n,k} + \tau_n\,[\underline{P}^{n,k}]^T\,(\mathcal{A}_D - \mathcal{A}_L)\,\underline{P}^{n,k} = \tau_n\,[\underline{P}^{n,k}]^T\,\mathcal{A}_L^T\,\underline{P}^{n,k-1}; \tag{4.4a}$$

and similarly it follows from subtracting (4.2b) from (4.1b) that

$$\ell(\gamma)\,[\underline{Q}^{n,k}]^T\,\mathcal{M}\,\underline{F}^{n,k} + \tau_n\,[\underline{Q}^{n,k}]^T\,\mathcal{M}\,\underline{Q}^{n,k} = 0. \tag{4.4b}$$

Choosing $\{\underline{\chi_1}, \underline{\chi_2}\} \equiv \{\underline{U}_\varepsilon^{n,k}, \underline{V}_\varepsilon^{n,k}\}$ in (4.1c) and $\{\underline{\chi_1}, \underline{\chi_2}\} \equiv \{\underline{U}_\varepsilon^{n}, \underline{V}_\varepsilon^{n}\}$ in (4.2c) yields

$$\begin{aligned} -\gamma\,\Big[\,[\underline{E}^{n,k}]^T\,(\mathcal{B}_D - \mathcal{B}_L)\,\underline{E}^{n,k} + [\underline{F}^{n,k}]^T\,(\mathcal{B}_D - \mathcal{B}_L)\,\underline{F}^{n,k}\,\Big] + [\underline{E}^{n,k}]^T\,\mathcal{M}\,\underline{P}^{n,k} \\ + [\underline{F}^{n,k}]^T\,\mathcal{M}\,\underline{Q}^{n,k} \geq -\gamma\,\Big[\,[\underline{E}^{n,k}]^T\,\mathcal{B}_L^T\,\underline{E}^{n,k-1} + [\underline{F}^{n,k}]^T\,\mathcal{B}_L^T\,\underline{F}^{n,k-1}\,\Big]. \end{aligned} \tag{4.5}$$

Combining (4.4a,b) and (4.5) yields that

$$\begin{aligned} \gamma^2\,\Big[\,[\underline{E}^{n,k}]^T\,(\mathcal{B}_D - \mathcal{B}_L)\,\underline{E}^{n,k} + [\underline{F}^{n,k}]^T\,(\mathcal{B}_D - \mathcal{B}_L)\,\underline{F}^{n,k}\,\Big] \\ + \tau_n\,\Big[\,[\underline{P}^{n,k}]^T\,(\mathcal{A}_D - \mathcal{A}_L)\,\underline{P}^{n,k} + [\ell(\gamma)]^{-1}\,\gamma\,[\underline{Q}^{n,k}]^T\,\mathcal{M}\,\underline{Q}^{n,k}\,\Big] \\ \leq \gamma^2\,\Big[\,[\underline{E}^{n,k}]^T\,\mathcal{B}_L^T\,\underline{E}^{n,k-1} + [\underline{F}^{n,k}]^T\,\mathcal{B}_L^T\,\underline{F}^{n,k-1}\,\Big] + \tau_n\,[\underline{P}^{n,k}]^T\,\mathcal{A}_L^T\,\underline{P}^{n,k-1}. \end{aligned} \tag{4.6}$$

We now split the diagonal matrix $\mathcal{A}_D := \mathcal{A}_{D_1} + \mathcal{A}_{D_2}$, where $(\mathcal{A}_{D_1})_{ii} := -\sum_{j=1}^{i-1} \mathcal{A}_{ij}$ and $(\mathcal{A}_{D_2})_{ii} := -\sum_{j=i+1}^{\mathcal{J}} \mathcal{A}_{ij} = \mathcal{A}_{ii} - (\mathcal{A}_{D_1})_{ii}$. Then, on noting from (2.13) that $(\mathcal{A}_L)_{ij} \geq 0$, we have that

$$\begin{aligned} [\underline{P}^{n,k}]^T\,\mathcal{A}_L^T\,\underline{P}^{n,k-1} &= \sum_{i=1}^{\mathcal{J}} P_i^{n,k} \sum_{j=1}^{\mathcal{J}} (\mathcal{A}_L^T)_{ij}\,P_j^{n,k-1} \leq \tfrac{1}{2} \sum_{i=1}^{\mathcal{J}} \sum_{j=1}^{\mathcal{J}} (\mathcal{A}_L)_{ji}\,[(P_i^{n,k})^2 + (P_j^{n,k-1})^2] \\ &= \tfrac{1}{2} \sum_{i=1}^{\mathcal{J}} (\mathcal{A}_{D_2})_{ii}\,(P_i^{n,k})^2 + \tfrac{1}{2} \sum_{j=1}^{\mathcal{J}} (\mathcal{A}_{D_1})_{jj}\,(P_j^{n,k-1})^2. \end{aligned} \tag{4.7}$$

Combining (4.6), (4.7) and a similar argument for $\mathcal{B}$, on noting (2.1), yields that

$$\begin{aligned} \gamma^2\,\Big[\,[\underline{E}^{n,k}]^T \mathcal{B}\,\underline{E}^{n,k} + [\underline{E}^{n,k}]^T \mathcal{B}_{D_1}\,\underline{E}^{n,k} + [\underline{F}^{n,k}]^T \mathcal{B}\,\underline{F}^{n,k} + [\underline{F}^{n,k}]^T \mathcal{B}_{D_1}\,\underline{F}^{n,k}\,\Big] \\ + \tau_n\,\Big[\,[\ell(\gamma)]^{-1}\,\gamma\,[\underline{Q}^{n,k}]^T\,\mathcal{M}\,\underline{Q}^{n,k} + [\underline{P}^{n,k}]^T \mathcal{A}^{n-1}\,\underline{P}^{n,k} + [\underline{P}^{n,k}]^T \mathcal{A}_{D_1}\,\underline{P}^{n,k}\,\Big] \\ \leq \gamma^2\,\Big[\,[\underline{E}^{n,k-1}]^T \mathcal{B}_{D_1}\,\underline{E}^{n,k-1} + [\underline{F}^{n,k-1}]^T \mathcal{B}_{D_1}\,\underline{F}^{n,k-1}\,\Big] + \tau_n\,[\underline{P}^{n,k-1}]^T \mathcal{A}_{D_1}\,\underline{P}^{n,k-1}. \end{aligned} \tag{4.8}$$

Therefore, we have that $\{\gamma^2\,(\,[\underline{E}^{n,k}]^T\,\mathcal{B}_{D_1}\,\underline{E}^{n,k} + [\underline{F}^{n,k}]^T \mathcal{B}_{D_1}\,\underline{F}^{n,k}\,) + \tau_n\,[\underline{P}^{n,k}]^T\,\mathcal{A}_{D_1}\,\underline{P}^{n,k}\,\}_{k\geq 0}$ is a decreasing sequence. Since it is bounded below the sequence has a limit. Combining this and (4.8) yields that

$$\begin{aligned} &|U_\varepsilon^n - U_\varepsilon^{n,k}|_1 \to 0, \quad |V_\varepsilon^n - V_\varepsilon^{n,k}|_1 \to 0, \quad |Z_\varepsilon^n - Z_\varepsilon^{n,k}|_h \to 0, \\ \text{and} \quad &|[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{\frac{1}{2}}\,\nabla(W_\varepsilon^n - W_\varepsilon^{n,k})|_0 \to 0 \quad \text{as } k \to \infty. \end{aligned} \tag{4.9}$$

Furthermore, multiplying (4.2a) with $\underline{1}^T := (1, \ldots, 1)$, noting that $\mathcal{A}^{n-1}\,\underline{1} = \underline{0}$ and recalling the splitting of $\mathcal{A}^{n-1}$ yields that

$$\begin{aligned} \gamma\,(U_\varepsilon^{n,k} - U_\varepsilon^{n-1}, 1)^h &= \tau_n\,\underline{1}^T \mathcal{A}_L^T\,(\underline{W}_\varepsilon^{n,k-1} - \underline{W}_\varepsilon^{n,k}) = \tau_n\,\underline{1}^T \mathcal{A}_{D_1}\,(\underline{W}_\varepsilon^{n,k-1} - \underline{W}_\varepsilon^{n,k}) \\ &= \tau_n\,\underline{1}^T \mathcal{A}_{D_1}\,\underline{P}^{n,k} - \tau_n\,\underline{1}^T \mathcal{A}_{D_1}\,\underline{P}^{n,k-1} \quad \to\ 0\,; \end{aligned} \tag{4.10}$$

where we have again used the fact that $\{\, \tau_n\, [\underline{P}^{n,k}]^T\, \mathcal{A}_{D_1}\, \underline{P}^{n,k}\, \}_{k\geq 0}$ has a limit. Combining (4.9), (4.10), (2.2) and (1.14) yields the desired result (4.3a). Similarly, multiplying (4.2b) with $\underline{1}^T := (1,\ldots,1)$, yields, on noting (4.9), that

$$\ell(\gamma)\,(V_\varepsilon^{n,k} - V_\varepsilon^{n-1}, 1)^h = -\tau_n\, \underline{1}^T\, \mathcal{M}\, \underline{Z}_\varepsilon^{n,k} \to -\tau_n\, \underline{1}^T\, \mathcal{M}\, \underline{Z}_\varepsilon^{n} \quad \text{as } k\to\infty\,. \tag{4.11}$$

The desired result (4.3b) then follows from (4.11), (4.1b), (4.9), (2.2) and (1.14). □

We note that (4.2a–c) can be solved explicitly for $j = 1 \to \mathcal{J}$. In particular, let $\widehat{\underline{r}}^1 := \underline{r}_1 + \tau_n\,(\mathcal{A}_L\,\underline{W}_\varepsilon^{n,k} + \mathcal{A}_L^T\,\underline{W}_\varepsilon^{n,k-1})$, $\widehat{\underline{r}}^2 := \underline{r}_2$, $\widehat{\underline{s}}^1 := \underline{s}_1 + \gamma\,(\mathcal{B}_L\,\underline{U}_\varepsilon^{n,k} + \mathcal{B}_L^T\,\underline{U}_\varepsilon^{n,k-1})$ and $\widehat{\underline{s}}^2 := \underline{s}_2 + \gamma\,(\mathcal{B}_L\,\underline{V}_\varepsilon^{n,k} + \mathcal{B}_L^T\,\underline{V}_\varepsilon^{n,k-1})$. Then $\{[\underline{U}_\varepsilon^{n,k}]_j, [\underline{V}_\varepsilon^{n,k}]_j\}$ is the solution of: Find $\{U_j, V_j\} \in \mathcal{K}$ such that

$$(\chi_1 - U_j)\,(C_1\,U_j - b_1) + (\chi_2 - V_j)\,(C_2\,V_j - b_2) \geq 0 \qquad \forall\ \{\chi_1,\chi_2\} \in \mathcal{K}, \tag{4.12}$$

where $C_1 := \gamma\,(\,\mathcal{B}_{jj} + \frac{[\mathcal{M}_{jj}]^2}{\tau_n\,\mathcal{A}_{jj}^{n-1}}\,)$, $C_2 := \gamma\,\mathcal{B}_{jj} + \frac{\ell(\gamma)\,\mathcal{M}_{jj}}{\tau_n}$ and $b_1 := \widehat{\underline{s}}_j^1 + \frac{\mathcal{M}_{jj}\,\widehat{\underline{r}}_j^1}{\tau_n\,\mathcal{A}_{jj}^{n-1}}$, $b_2 := \widehat{\underline{s}}_j^2 + \frac{\widehat{\underline{r}}_j^2}{\tau_n}$. Clearly, the unique solution to (4.12) is

$$\{U_j, V_j\} \equiv \mathrm{P}_{\mathcal{K}}^C(\tfrac{b_1}{C_1}, \tfrac{b_2}{C_2}),$$

where $\mathrm{P}_{\mathcal{K}}^C(x_1,x_2)$ is the orthogonal projection of the point $\underline{x} = \{x_1,x_2\} \in \mathbb{R}^2$ onto $\mathcal{K}$ with respect to the $\mathbb{R}^2$ inner product $\langle \underline{p}, \underline{q}\rangle_C := \underline{p}^T C\,\underline{q}$, with $C = \left(\begin{smallmatrix} C_1 & 0 \\ 0 & C_2 \end{smallmatrix}\right)$. The projection $\underline{y} = \mathrm{P}_{\mathcal{K}}^C(\underline{x})$ can be computed as follows.

1. If $\underline{x} \in \mathcal{K}$, then $\underline{y} = \underline{x}$, else
2. If $x_1 \geq 1$ then $\underline{y} := (1, \max\{-\frac{2}{\sqrt{3}}, \min\{x_2, \frac{2}{\sqrt{3}}\}\})^T$, else
3. If $x_2 \geq 0$ then $\underline{v} := (2, \frac{2}{\sqrt{3}})^T$, else $\underline{v} := (2, -\frac{2}{\sqrt{3}})^T$.
4. $\alpha := \dfrac{\langle \underline{x} + (1,0)^T, \underline{v}\rangle_C}{\|\underline{v}\|_C^2}$.
5. $\underline{y} := (-1,0)^T + \min\{\max\{\alpha,0\},1\}\,\underline{v}$.

Hence the solution of (4.2a–c) is for $j = 1 \to \mathcal{J}$

$$\{[\underline{U}_\varepsilon^{n,k}]_j, [\underline{V}_\varepsilon^{n,k}]_j\} \equiv \mathrm{P}_{\mathcal{K}}^C\left(\frac{\mathcal{M}_{jj}\,\widehat{\underline{r}}_j^1 + \tau_n\,\mathcal{A}_{jj}^{n-1}\,\widehat{\underline{s}}_j^1}{\gamma\,[\mathcal{M}_{jj}]^2 + \tau_n\,\gamma\,\mathcal{A}_{jj}^{n-1}\,\mathcal{B}_{jj}}, \frac{\widehat{\underline{r}}_j^2 + \tau_n\,\widehat{\underline{s}}_j^2}{\ell(\gamma)\,\mathcal{M}_{jj} + \gamma\,\tau_n\,\mathcal{B}_{jj}}\right) \tag{4.13a}$$

$$\text{and}\quad [\underline{W}_\varepsilon^{n,k}]_j = \frac{\widehat{\underline{r}}_j^1 - \gamma\,\mathcal{M}_{jj}\,[\underline{U}_\varepsilon^{n,k}]_j}{\tau_n\,\mathcal{A}_{jj}^{n-1}}, \qquad [\underline{Z}_\varepsilon^{n,k}]_j = \frac{\widehat{\underline{r}}_j^2 - \ell(\gamma)\,\mathcal{M}_{jj}\,[\underline{V}_\varepsilon^{n,k}]_j}{\tau_n\,\mathcal{M}_{jj}}. \tag{4.13b}$$

We note that when using the approximation $(\widetilde{\mathrm{P}}_\varepsilon^{h,\tau})$, see Remark 2.1, there exist j with $\mathcal{A}_{jj}^{n-1} = 0$. For those j, (4.13a,b) is modified as follows:

$$U_\varepsilon^{n-1}(p_j) = -1 \Rightarrow \{[\underline{U}_\varepsilon^{n,k}]_j, [\underline{V}_\varepsilon^{n,k}]_j\} \equiv \{-1, 0\} \tag{4.14a}$$

$$U_\varepsilon^{n-1}(p_j) = 1 \quad \Rightarrow \{[\underline{U}_\varepsilon^{n,k}]_j, [\underline{V}_\varepsilon^{n,k}]_j\} \equiv \{1, \max(-\tfrac{2}{\sqrt{3}}, \min\{\tfrac{2}{\sqrt{3}}, \tfrac{\widehat{\underline{r}}_j^2 + \tau_n\,\widehat{\underline{s}}_j^2}{\ell(\gamma)\,\mathcal{M}_{jj} + \gamma\,\tau_n\,\mathcal{B}_{jj}}\}\,)\,\}\,, \tag{4.14b}$$

where in both cases $[Z_\varepsilon^{n,k}]_j$ is then defined as in (4.13b). We note that as $\mathcal{A}_{jj}^{n-1} = 0$, $[W_\varepsilon^{n,k}]_j$ is not defined and not required.

5 Numerical results

Throughout this section, we use (1.7) for Ψ_0 in (1.6), and for the initial data u^0 to (P) choose a circular void with radius $R \in \mathbb{R}_{>0}$ and centre $y \in \mathbb{R}^2$; that is,

$$u^0(x) = \rho_c(y, R;\, x) := \begin{cases} -1 & r(x) \leq R - \frac{\delta_u}{2} \\ \sin(\frac{r(x)-R}{\delta_u}\,\pi) & |r(x) - R| < \frac{\delta_u}{2} \\ 1 & r(x) \geq R + \frac{\delta_u}{2} \end{cases}, \quad \text{where } r(x) := |x - y|, \quad (5.1)$$

where $\delta_u := (1 - \frac{\mu}{4})^{-\frac{1}{2}}\,\gamma\,\pi$ is the interfacial thickness of u^0. For the initial profile v^0, on letting $\delta_v := (1 - \mu)^{-\frac{1}{2}}\,\gamma\,\pi$, we choose

$$v^0(x) = \tfrac{1}{\sqrt{3}}\,[u^0(x) + 1]\,\rho_{l_i}(y; x), \quad \text{where } \rho_{l_i}(y; x) := \begin{cases} -1 & y_i - x_i \leq -\frac{\delta_v}{2} \\ \sin(\frac{y_i - x_i}{\delta_v}\,\pi) & |y_i - x_i| < \frac{\delta_v}{2} \\ 1 & y_i - x_i \geq \frac{\delta_v}{2} \end{cases}, \quad (5.2)$$

for a vertical ($i = 1$) and a horizontal ($i = 2$) grain boundary, respectively. Note that the interfacial thickness of u^0 and v^0 is in line with the asymptotics of the phase field approach, see (A.15) and (A.14). Unless stated otherwise, we will always use the scaling $\ell(\gamma) := \gamma^2$ and set $\varepsilon = 10^{-5}$.

For the iterative algorithm (4.2a–c) we set, for $n \geq 1$, $\{U_\varepsilon^{n,0}, V_\varepsilon^{n,0}, W_\varepsilon^{n,0}, Z_\varepsilon^{n,0}\} \equiv \{U_\varepsilon^{n-1}, V_\varepsilon^{n-1}, W_\varepsilon^{n-1}, Z_\varepsilon^{n-1}\}$, where $\{U_\varepsilon^0, V_\varepsilon^0\} \equiv \{\pi^h u^0, \pi^h v^0\}$ and $W_\varepsilon^0 \equiv -\gamma\,\Delta^h U_\varepsilon^0 - \gamma^{-1}\,\pi^h[\Psi_{,u}(U_\varepsilon^0, V_\varepsilon^0)]$ $Z_\varepsilon^0 \equiv -\gamma\,\Delta^h V_\varepsilon^0 - \gamma^{-1}\,\pi^h[\Psi_{,v}(U_\varepsilon^0, V_\varepsilon^0)]$; and adopted the stopping criterion

$$\max\left\{|U_\varepsilon^{n,k} - U_\varepsilon^{n,k-1}|_{0,\infty}, |V_\varepsilon^{n,k} - V_\varepsilon^{n,k-1}|_{0,\infty}\right\} < tol,$$

with $tol = 10^{-7}$, and then setting $\{U_\varepsilon^n, V_\varepsilon^n, W_\varepsilon^n, Z_\varepsilon^n\} \equiv \{U_\varepsilon^{n,k}, V_\varepsilon^{n,k}, W_\varepsilon^{n,k}, Z_\varepsilon^{n,k}\}$.

Throughout the given domain $\Omega = (-L_1, L_1) \times (-L_2, L_2)$ is partitioned into right-angled isosceles triangles. Here we assume that L_1 and L_2 are integer multiples of L, where $L := \min\{L_1, L_2\}$. On using the adaptive finite element code Alberta 1.2, see Schmidt and Siebert (2004), we implemented the same mesh refinement strategy as in Barrett, Nürnberg, and Styles (2004). In particular, to improve efficiency we use the approximation $(\widetilde{\mathrm{P}}_\varepsilon^{h,\tau})$, see Remark 2.1 and (4.14a,b). Now we have to solve for $\{U_\varepsilon^n, W_\varepsilon^n\}$ only in the interfacial region, $|U_\varepsilon^{n-1}| < 1$, while the solution $\{V_\varepsilon^n, Z_\varepsilon^n\}$ has to be found where $U_\varepsilon^{n-1} > -1$. However, the evolution will concentrate inside the two interfacial regions $|U_\varepsilon^{n-1}| < 1$ and $|U_\varepsilon^{n-1}| = 1, |V_\varepsilon^{n-1}| < \frac{2}{\sqrt{3}}$. Hence we use a refined mesh with mesh size $h_f = \frac{2^{\frac{3}{2}}\,L}{N_f}$ in these interfacial regions, and a coarser mesh of mesh size $h_c = \frac{2^{\frac{3}{2}}\,L}{N_c}$ away from the interfaces. Here N_f and N_c are parameters, see Barrett, Nürnberg, and Styles (2004, §5). Furthermore, we choose N_f such that there are always approximately 8 mesh points across the interface in each direction. In particular, for $\mu \geq 0$ it will always hold that $h_f \leq \frac{3\sqrt{2}}{32}\,\gamma\,\pi$, whereas for $\mu < 0$ we ensure that $h_f \leq \frac{3\sqrt{2}}{32}\,(1 - \mu)^{-\frac{1}{2}}\,\gamma\,\pi$.

For our first experiments we choose $\mu = 0$ in (1.7). That means, that the function Ψ is symmetric with respect to the three vertices A, B and C of $\mathcal{K}$. In particular, the

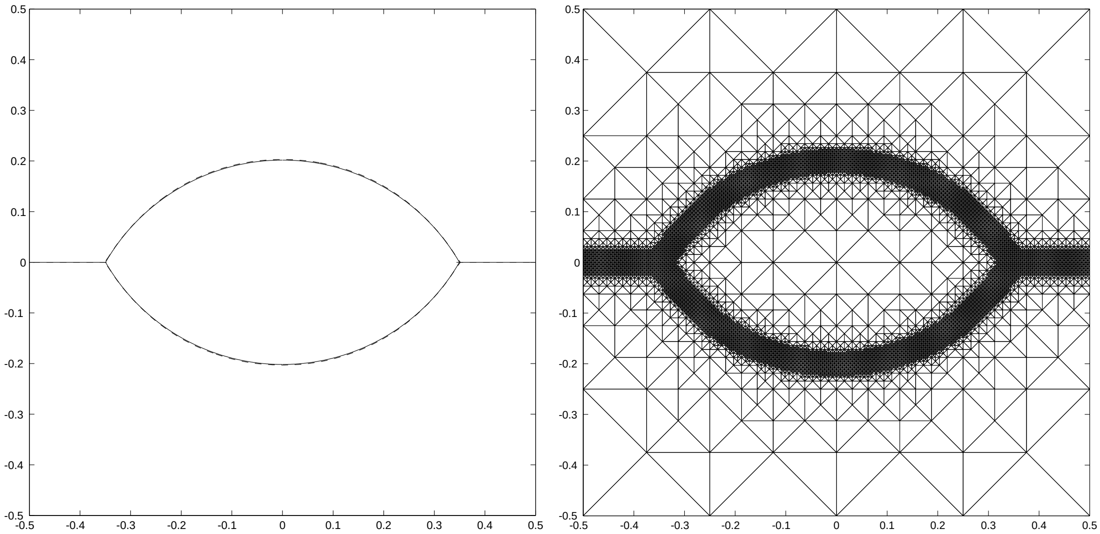

Figure 3: ($\gamma = \frac{1}{24\pi}$, $\alpha = 0$) Comparison between computed solution (dashed) and true solution (solid). The final triangulation is shown on the right.

surface energies associated with the three different interfaces will be the same, and hence we should observe a 120° degree contact angle at triple junctions between the void and the two grains. In order to check the accuracy of our approximation, we compare the evolution of an initially circular void between two horizontally aligned grains with the true steady state solution. It was shown by Ito and Kohsaka (2001) that the true solution for the void boundary consists of four symmetric branches, where one branch is given by

$$x_2 = f(x_1) := -a\,\cos\theta + (a^2 - x_1^2)^{\frac{1}{2}} \qquad \text{for } x_1 \in [-a\,\sin\theta, 0]. \tag{5.3}$$

Here $a = (\frac{A}{2\,\theta - \sin(2\,\theta)})^{\frac{1}{2}}$ with $A = \pi\,R^2$ being the total area of the void and $2\,\theta = \frac{2\pi}{3}$ being the contact angle between grains and void. We chose the following parameters for $(\widetilde{\mathrm{P}}_{\varepsilon}^{h,\tau})$ $L_1 = L_2 = 0.5$, $\gamma = \frac{1}{24\pi}$, $\alpha = 0$, $T = 10^{-2}$, $\tau_n = \tau = 5 \times 10^{-8}$. For the initial profile we chose (5.1) and (5.2) with $i = 2$, $y = \{0, 0\}$, $R = 0.25$. The refinement parameters were $N_f = 256$ and $N_c = 2$. The comparison between true solution and the numerically steady state can be seen in Figure 3, where we also include a plot of the mesh at time $t = T$. One can see that the true solution and our computation are almost graphically indistinguishable.

A short remark on the way we plot the solution $\{U_\varepsilon, V_\varepsilon\}$ is due. In our figures we show the zero level sets of the function $p(U_\varepsilon, V_\varepsilon)$ to visualize the void boundary, where

$$p(y) := \max\{|y - A|^2 - |y - B|^2, |y - A|^2 - |y - C|^2\}\,.$$

In addition, we give the zero contour line of V_ε where $U_\varepsilon > 0$, in order to show the grain boundaries. Next, we conducted the following convergence experiments for the evolution of a circular void in a vertical grain boundary under the influence of electromigration. We repeated the same experiment with decreasing values of γ, i.e. $\gamma = \frac{1}{12\pi}, \frac{1}{24\pi}, \frac{1}{48\pi}$. In particular, we set $L_1 = L_2 = 0.5$, $T = 4 \times 10^{-3}$, $\tau_n = \tau = 288\,(\gamma\,\pi)^2 \times 10^{-7}$, $\varepsilon =$

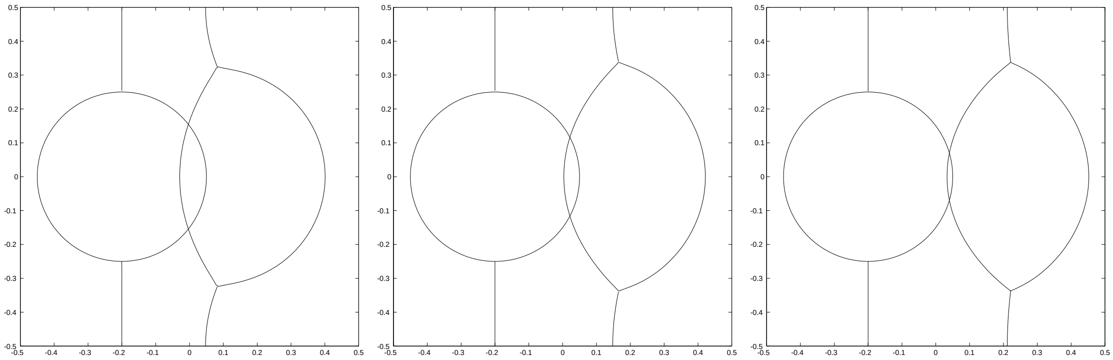

Figure 4: ($\alpha = 5\,\pi$) Solution $\{U_\varepsilon, V_\varepsilon\}$ at times $t = 0,\ T = 4 \times 10^{-3}$ for $\gamma = \frac{1}{12\pi}$, $\gamma = \frac{1}{24\pi}$ and $\gamma = \frac{1}{48\pi}$.

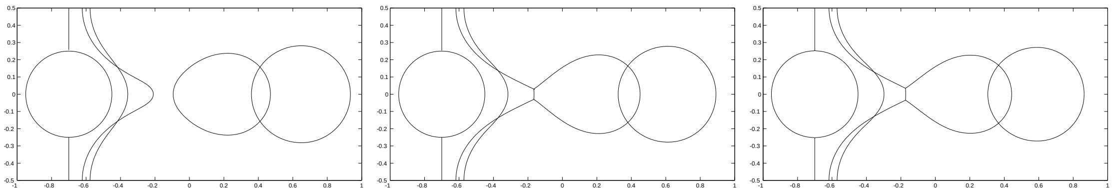

Figure 5: ($\alpha = 5\,\pi$) Solution $\{U_\varepsilon, V_\varepsilon\}$ at times $t = 0,\ 0.04,\ T = 0.056$ for $\gamma = \frac{1}{12\pi}$, $\gamma = \frac{1}{24\pi}$ and $\gamma = \frac{1}{48\pi}$.

$48\,\gamma\,\pi \times 10^{-5}$ and used the appropriate refinement parameters $N_f = \frac{32}{3}\,\frac{1}{\gamma\pi}$ and $N_c = \frac{N_f}{8}$. Considerations using formal asymptotic expansions, see (A.41), yield that in the sharp interface limit the grain boundaries have zero curvature and a 90° degree contact angle with the boundary. This can be observed in the convergence experiment, where for γ getting smaller the grain boundaries get closer and closer to straight lines. See Figure 4, where we plot the results for $\gamma = \frac{1}{12\pi}$, $\gamma = \frac{1}{24\pi}$ and $\gamma = \frac{1}{48\pi}$.

The same experiment for the scaling $\ell(\gamma) := \gamma$ leads to a dramatically different evolution, as this now models surface diffusion combined with surface attachment limited kinetics (SALK), see (A.42). For the new scaling, we repeated the previous experiment on a slightly larger domain Ω in order to see more of the ensuing evolution. We used the following parameters: $L_1 = 1$, $L_2 = 0.5$, $T = 0.056$, $\tau_n = \tau = 1152\,(\gamma\,\pi)^2 \times 10^{-7}$, $\varepsilon = 48\,\gamma\,\pi \times 10^{-5}$ and used the appropriate refinement parameters $N_f = \frac{32}{3}\,\frac{1}{\gamma\pi}$, $N_c = \frac{N_f}{8}$. In Figure 5 one can see that the void detaches from the grain boundary. Note also the very good agreement between the results as γ is decreased.

In a further experiment, we investigated the evolution of a circular void when it attaches to a vertical grain boundary. To this end, we set the following parameters for $(\widetilde{\mathrm{P}}^{h,\tau}_{\varepsilon})$: $L_1 = 1$, $L_2 = 0.5$, $\gamma = \frac{1}{24\pi}$, $\alpha = 5\,\pi$, $T = 0.012$, $\tau_n = \tau = 5 \times 10^{-8}$, $\varepsilon = 10^{-5}$. For the initial profile we chose (5.1) and (5.2) with $i = 1$, $y = \{0, 0\}$, $R = 0.25$. The refinement parameters were $N_f = 256$ and $N_c = 32$. The evolution is shown in Figure 6. We can observe that once the void has attached to the grain boundary, it settles into a steady shape inside the grain boundary, which then drifts through the conductor.

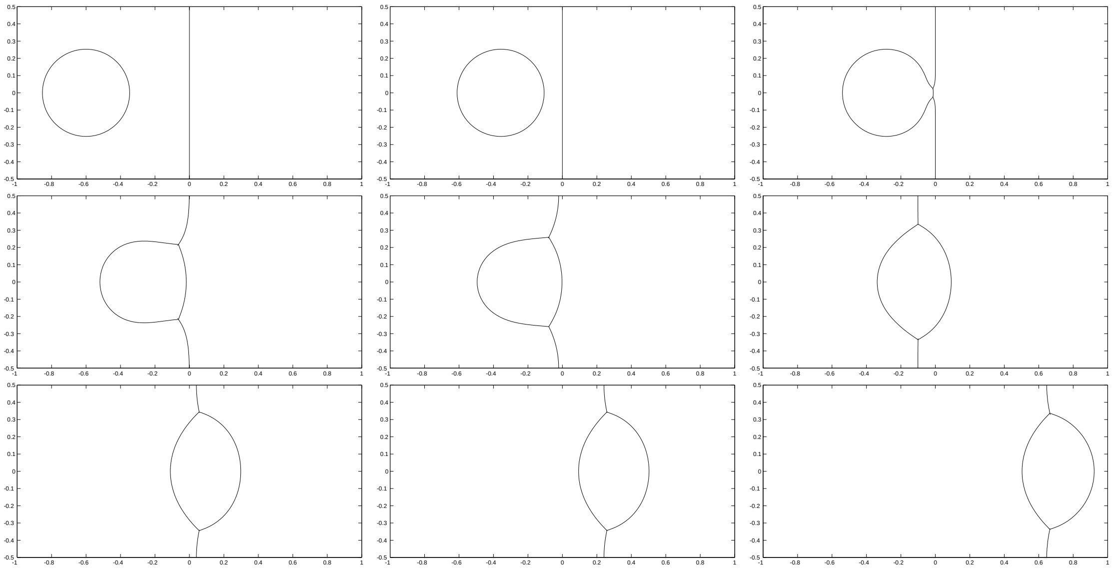

Figure 6: ($\gamma = \frac{1}{24\pi}$, $\alpha = 5\,\pi$) Solution $\{U_\varepsilon, V_\varepsilon\}$ at times $t = 0,\ 2\times10^{-3},\ 2.6\times10^{-3},\ 2.8\times10^{-3},\ 3\times10^{-3},\ 4\times10^{-3},\ 6\times10^{-3},\ 8\times10^{-3},\ 0.012$.

We include also an experiment that produces a travelling wave solution in the absence of electromigration, first mentioned in Mullins (1958) (see also Kanel, Novick-Cohen, and Vilenkin (2004)). We used the following parameters for $(\widetilde{\mathrm{P}}^{h,\tau}_{\varepsilon})$: $L_1 = 1$, $L_2 = 0.5$, $\gamma = \frac{1}{24\pi}$, $\alpha = 0$, $T = 2.6\times10^{-3}$, $\tau_n = \tau = 5\times10^{-8}$, $\varepsilon = 10^{-5}$. For the initial profile we chose a straight horizontal line for u^0, as described by ρ_{l_2} in (5.2) with $y = \{0,0\}$, and a straight line with a segment of a circle for v^0, i.e. (5.2) with ρ_{l_i} replaced by

$$\rho_q(y, R;\, x) := \begin{cases} \rho_c(y + \{0, R\}, R;\, x) & x_1 < y_1, \\ \rho_{l_2}(y;\, x) & x_1 \geq y_1 \end{cases}$$

with $y = \{-0.3, -0.3\}$, $R = 0.25$. The refinement parameters were $N_f = 256$ and $N_c = 2$. The evolution is shown in Figure 7.

5.1 Different contact angles

In this subsection, we report on contact angles for the triple junction that are different from the symmetric case $\frac{2\pi}{3}$. Since different contact angles are observed in practice, this is an important and desirable feature of our phase field model. In order to achieve different triple junction angles, we have to choose the obstacle potential Ψ, see (1.6), such that the grain and material boundaries have different surface energies. To this end, we use (1.7) with $\mu \neq 0$.

Assume we are given the ratio of the surface energies for the grain and material boundaries $\frac{\sigma^A}{\sigma^B}$, where we have adopted the notation of the Appendix, see (A.37). Then this angle law, where $\sigma^B = \sigma^C$ is the surface energy of the material boundary and σ^A is

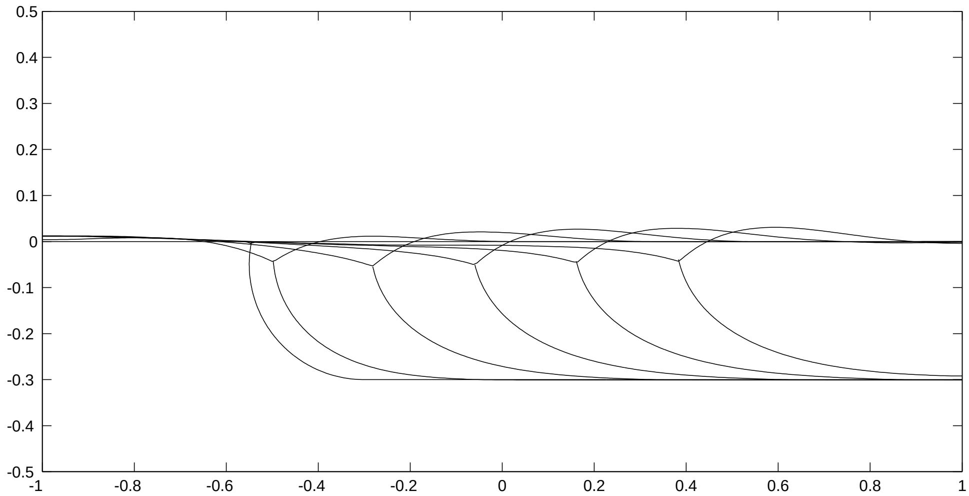

Figure 7: ($\gamma = \frac{1}{24\pi}$, $\alpha = 0$) Solution $\{U_\varepsilon, V_\varepsilon\}$ at times $t = 0,\ 2\times10^{-4},\ 8\times10^{-4},\ 1.4\times10^{-3},\ 2\times10^{-3},\ 2.6\times10^{-3}$.

the surface energy of the grain boundary, yields that

$$\theta^A = 2\ \arccos(\tfrac{1}{2}\,\tfrac{\sigma_{grain}}{\sigma_{mat}}).$$

Using (A.17) we compute for $\mu \in (-2, \frac{4}{7})$ that

$$\frac{\sigma_{grain}}{\sigma_{mat}} = \frac{\frac{2}{3}\,\pi\,(1-\mu)^{\frac{1}{2}}}{\frac{2}{3}\,\pi\,(1-\frac{\mu}{4})^{\frac{1}{2}}} = 2\,(\tfrac{1-\mu}{4-\mu})^{\frac{1}{2}}\,. \tag{5.4}$$

In the derivation of (A.17) it is assumed that the first order solution to the variational inequality (A.13) leads after a suitable rescaling to a minimiser in (1.1). However it is not straightforward to establish this rigorously. In any case, one can also compute the above ratio numerically. To this end, one splits the domain Ω into two pure phases $i, j \in \{A, B, C\}$, with a vertically or horizontally aligned straight phase boundary between them. Using this setup for the initial profiles of $\{u^0, v^0\}$, one computes the evolution of $(\widetilde{\mathrm{P}}^{h,\tau}_{\varepsilon})$ until a steady state has been reached. This resulting standing wave will then approximate the energy minimizing profile in (1.1), and hence provides a numerical value for the energy density σ.

For the case $\mu = \frac{1}{2}$, we computed the different surface energies for the grain and material boundaries in this way and obtained a ratio $\frac{\sigma_{grain}}{\sigma_{mat}} \approx 0.758$, i.e. almost exactly the value $\frac{2}{\sqrt{7}}$ derived from (5.4). This suggests a triple junction with angles 135° and twice 112.5°, which is confirmed by the numerical results shown in Figure 8, where we have used the same parameters as for Figure 3. Note that the true steady state solution is again defined by (5.3).

Next, we computed the different surface energies for the grain and material boundaries numerically for the case $\mu = -1$ and obtained a ratio $\frac{\sigma_{grain}}{\sigma_{mat}} \approx 1.26$, i.e. almost exactly the value $(\frac{8}{5})^{\frac{1}{2}}$ derived from (5.4). This suggests a triple junction with degrees 102° and twice 129°. This is confirmed by the numerical results shown in Figure 8, where we used the same parameters for $(\widetilde{\mathrm{P}}^{h,\tau}_{\varepsilon})$ as before, except $\gamma = \frac{\sqrt{2}}{24\pi}$.

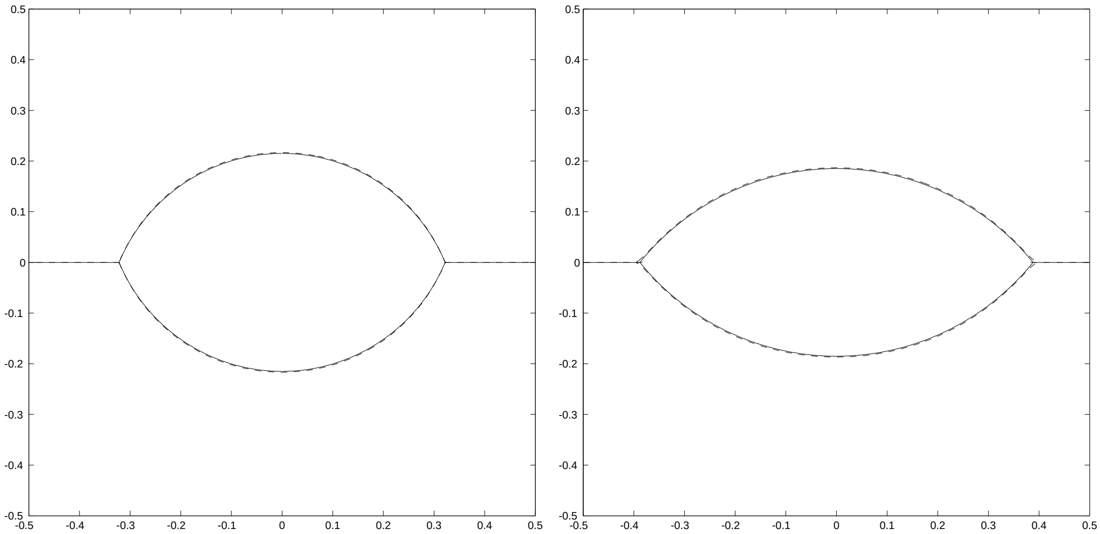

Figure 8: ($\gamma = \frac{1}{24\pi}$, $\mu = \frac{1}{2}$ (left) and $\gamma = \frac{\sqrt{2}}{24\pi}$, $\mu = -1$ (right)) Comparison between computed solution (dashed) and true solution (solid) for $\theta^A = 135°$ and $\theta^A = 102°$, respectively.

A Formal asymptotic expansions

The method of formally matched asymptotic expansions for systems with triple junctions is by now well-established, see e.g. Bronsard and Reitich (1993), Garcke and Novick-Cohen (2000), Novick-Cohen (2000) and Garcke, Nestler, and Stoth (1998). Hence we will only sketch the asymptotics where arguments are similar to other papers, and only present the new features in more detail. Three different types of expansions will be used. In regions where either a grain or the void is present we use an outer expansion. Close to interfaces separating either a void and a grain or two grains an inner expansion is used. A third type of expansion has to be performed at a triple junction. All these expansions have to be matched.

The equations for the outer expansion imply that the vector (u, v) attains one of the values A, B, C. That is, in the sharp interface limit (u, v) will be either A, B or C and there are interfaces separating these regions. For the electric potential ϕ we obtain that it solves Laplace's equation in the regions where (u, v) is either B or C. We note that as no confusion with the L^2 inner product can arise, we will use the round bracket notation for vectors throughout this Appendix.

Now the inner expansion has to be used to determine the governing equations on the interface. In the following we derive the governing equations, evolution laws, for the sharp interface limit. There are three interfaces (curves in two dimensions) for which we seek these laws. Let $\Gamma^{ij} = (\Gamma^{ij}(t))_{t\geq 0}$ with either $(i, j) = (A, B)$, (B, C) or (C, A) be a smooth evolving curve, an interface between regions occupied by i and j. Let $X^{ij}(s, t)$ be a parameterization of Γ^{ij}, where s is an arc-length parameter; see Gurtin (1993) for more information on evolving curves. We define the unit tangent $\tau_\Gamma^{ij} := \partial_s X^{ij}$ and the unit normal n_Γ^{ij} such that $(n_\Gamma^{ij}, \tau_\Gamma^{ij})$ is positively orientated, i.e. $\tau_\Gamma^{ij} = R\, n_\Gamma^{ij}$, where R is the clockwise rotation through $\frac{\pi}{2}$. We define also the direction of increasing s such that n_Γ^{ij} points into the region occupied by j. From now on we will suppress the region superscripts, when no confusion can arise. The curvature κ is defined to be positive if Γ is curved in the direction of the normal. With this choice the Frenet formulas read as

$$\partial_s n_\Gamma = -\kappa\, \tau_\Gamma, \qquad \partial_s \tau_\Gamma = \kappa\, n_\Gamma. \tag{A.1}$$

Since Γ is smooth, there exist functions $s(x,t)$ and $d(x,t)$ defined in a neighbourhood of Γ such that

$$x = X(s(x,t),t) + d(x,t)\, n_\Gamma(s(x,t),t); \tag{A.2}$$

see e.g. Gilbarg and Trudinger (1983, §14.6). The quantity $d(x,t)$ is the distance of the point x to $\Gamma(t)$ (note that $(x - X(s,t))\,.\,\tau_\Gamma(s,t) = 0$). In the following we will make use of the coordinate change

$$(x,t) \mapsto (\rho(x,t), s(x,t), t)\,,$$

where $\rho(x,t) = \gamma^{-1} d(x,t)$ is the re-scaled distance to Γ. This change of variables is a diffeomorphism flattening the interface Γ. Computing the spatial derivatives of (A.2), we obtain that

$$I = \partial_s X\, (\nabla_x s)^T + n_\Gamma\, (\nabla_x d)^T + d\, \partial_s n_\Gamma\, (\nabla_x s)^T = (1 - d\,\kappa)\, \tau_\Gamma\, (\nabla_x s)^T + n_\Gamma\, (\nabla_x d)^T. \tag{A.3}$$

This implies that

$$n_\Gamma(s(x,t),t) = \nabla_x d(x,t)\,, \qquad \tau_\Gamma(s(x,t),t) = (1 - d(x,t)\,\kappa(s(x,t),t))\, \nabla_x s(x,t)\,; \tag{A.4}$$

where these identities follow if we multiply (A.3) from the left by n_Γ^T and τ_Γ^T, respectively. Hence we have that

$$\nabla_x s\,.\,\nabla_x d = 0\,, \qquad |\nabla_x d| = 1 \qquad \text{and} \qquad |\nabla_x s| = \tfrac{1}{1-d\,\kappa}\,. \tag{A.5}$$

Taking the time derivative of (A.2) gives

$$0 = \partial_t X + \partial_s X\, \partial_t s + \partial_t d\, n_\Gamma + d\, \tfrac{d}{dt} n_\Gamma(s(x,t),t).$$

Dotting this identity with $n_\Gamma(s(x,t),t)$ yields that

$$0 = \partial_t X\,.\,n_\Gamma + \partial_t d + d\, \tfrac{1}{2}\, \tfrac{d}{dt}\, |n_\Gamma(t, s(x,t))|^2 = \partial_t X\,.\,n_\Gamma + \partial_t d\,,$$

where we have noted that $\frac{d}{dt}\,|n_\Gamma(t,s(x,t))|^2 = 0$. Defining the normal velocity of Γ as $\mathcal{V} := \partial_t X\,.\,n_\Gamma$, we obtain that

$$\mathcal{V} = -\partial_t d\,. \tag{A.6}$$

Using the new coordinates (ρ, s, t), we obtain the following identities for a scalar quantity $a(x,t) \equiv \widehat{a}(\rho(x,t), s(x,t), t)$

$$\nabla_x a(x,t) = \gamma^{-1} \partial_\rho \widehat{a}\, \nabla_x d + \partial_s \widehat{a}\, \nabla_x s \quad \text{and} \quad \partial_t a(x,t) = \gamma^{-1} \partial_\rho \widehat{a}\, \partial_t d + \partial_s \widehat{a}\, \partial_t s + \partial_t \widehat{a}\,. \tag{A.7}$$

For a vector function $\underline{a}(x,t) \equiv \underline{\widehat{a}}(\rho(x,t),s(x,t),t)$ we obtain that

$$\nabla_x \,.\, \underline{a}(x,t) = \gamma^{-1}\partial_\rho \underline{\widehat{a}} \,.\, \nabla_x d + \partial_s \underline{\widehat{a}} \,.\, \nabla_x s\,. \tag{A.8}$$

Combining (A.7) and (A.8) yields, on noting (A.5), that

$$\Delta_x a(x,t) = \gamma^{-2}\partial_{\rho\rho}\widehat{a} + \gamma^{-1}\partial_\rho\widehat{a}\,\Delta_x d + \partial_{ss}\widehat{a}\,|\nabla_x s|^2 + \partial_s\widehat{a}\,\Delta_x s\,. \tag{A.9}$$

Applying (A.8) to the right hand side of the identity $\nabla_x d(x,t) = n_\Gamma(s(x,t),t)$, and noting (A.4) and (A.1), we obtain that

$$\Delta_x d(x,t) = \partial_s n_\Gamma(s(x,t),t) \,.\, \nabla_x s = -\kappa\,\tau_\Gamma \,.\, (\tfrac{1}{1-d\,\kappa})\tau_\Gamma = -\tfrac{\kappa}{1-d\,\kappa}. \tag{A.10}$$

From (A.4), (A.8) and (A.1), it follows that

$$\begin{aligned}\Delta_x s(x,t) &= \nabla_x \,.\, (\tfrac{1}{1-\gamma\,\rho\,\kappa(s(x,t),t)}\,\tau_\Gamma(s(x,t),t)) = \partial_s(\tfrac{1}{1-\gamma\,\rho\,\kappa}\tau_\Gamma) \,.\, \nabla_x s \\ &= \partial_s(\tfrac{1}{1-\gamma\,\rho\,\kappa})\,\tau_\Gamma \,.\, (\tfrac{1}{1-d\,\kappa})\,\tau_\Gamma = \tfrac{\gamma\,\partial_s(\rho\,\kappa)}{(1-\gamma\,\rho\,\kappa)^3}\,.\end{aligned} \tag{A.11}$$

Combining (A.9), (A.10), (A.11) and (A.5) yields the following representation of Δ_x in the new coordinates

$$\begin{aligned}\Delta_x a(x,t) &= \gamma^{-2}\partial_{\rho\rho}\widehat{a} - \gamma^{-1}\partial_\rho\widehat{a}\,\tfrac{\kappa}{1-\gamma\,\rho\,\kappa} + \partial_{ss}\widehat{a}\,|\nabla_x s|^2 + \partial_s\widehat{a}\,\tfrac{\gamma\,\partial_s(\rho\,\kappa)}{(1-\gamma\,\rho\,\kappa)^3} \\ &= \gamma^{-2}\partial_{\rho\rho}\widehat{a} - \gamma^{-1}\partial_\rho\widehat{a}\,\tfrac{\kappa}{1-\gamma\,\rho\,\kappa} + \partial_{ss}\widehat{a} + \mathcal{O}(\gamma)\,.\end{aligned} \tag{A.12}$$

We now assume that there exist expansions of u, v, w, z and ϕ in these new variables, i.e. for example

$$u(x,t) = \widehat{u}(\rho,s,t) = \widehat{u}_0(\rho,s,t) + \gamma\,\widehat{u}_1(\rho,s,t) + \ldots\,.$$

In the following we drop the $\widehat{}$ for notational convenience and consider the potential (1.6) with Ψ_0 given by (1.7), where $\mu \in (-2, \frac{4}{7})$. Considering (1.5c) to leading order we obtain that $(u_0,v_0) : \mathbb{R} \to \mathcal{K}$ has to solve for all $(\eta_1,\eta_2) : \mathbb{R} \to \mathcal{K}$ the inequality

$$(-\partial_{\rho\rho}u_0 + \Psi_{,u}(u_0,v_0))\,(\eta_1 - u_0) + (-\partial_{\rho\rho}v_0 + \Psi_{,v}(u_0,v_0))\,(\eta_2 - v_0) \geq 0\,. \tag{A.13}$$

This variational inequality has the following solutions. At a grain boundary with

$$\lim_{\rho\to-\infty}(u_0,v_0)(\rho) = B = (1, -\tfrac{2}{\sqrt{3}}) \qquad \text{and} \qquad \lim_{\rho\to\infty}(u_0,v_0)(\rho) = C = (1, \tfrac{2}{\sqrt{3}})\,,$$

we obtain that $(u_0,v_0) = (1,\bar{v})$ with

$$\bar{v}(\rho) = \tfrac{2}{\sqrt{3}} \begin{cases} 1 & \text{if} \quad \rho > \rho_g := \frac{\pi}{2\sqrt{1-\mu}}\,, \\ \sin(\frac{\pi}{2}\,\frac{\rho}{\rho_g}) & \text{if} \quad |\rho| \leq \rho_g\,, \\ -1 & \text{if} \quad \rho < -\rho_g \end{cases} \tag{A.14}$$

is a solution, since $\mu \in (-2, \frac{4}{7})$. Similarly, at a material boundary with

$$\lim_{\rho\to-\infty}(u_0,v_0)(\rho) = A = (-1,0) \qquad \text{and} \qquad \lim_{\rho\to\infty}(u_0,v_0)(\rho) = B = (1, -\tfrac{2}{\sqrt{3}})\,,$$

we obtain that $(u_0, v_0) = (\bar{u}, -\frac{1+\bar{u}}{\sqrt{3}})$ with

$$\bar{u}(\rho) = \begin{cases} 1 & \text{if} \quad \rho > \rho_m := \frac{\pi}{\sqrt{4-\mu}}\,, \\ \sin(\frac{\pi}{2}\,\frac{\rho}{\rho_m}) & \text{if} \quad |\rho| \leq \rho_m\,, \\ -1 & \text{if} \quad \rho < -\rho_m \end{cases} \tag{A.15}$$

is a solution of the variational inequality (A.13). The solution of the material boundary CA is then given, through symmetry, as $(u_0, v_0)(\rho) = (\bar{u}, \frac{1+\bar{u}}{\sqrt{3}})(-\rho)$.

For later use we compute the interfacial energy

$$\begin{aligned} \sigma &= \int_{-\infty}^{\infty} \left[\, \tfrac{1}{2}((\partial_\rho u_0)^2 + (\partial_\rho v_0)^2) + \Psi(u_0, v_0) \,\right] \mathrm{d}\rho = \int_{-\infty}^{\infty} \left[\, (\partial_\rho u_0)^2 + (\partial_\rho v_0)^2 \,\right] \mathrm{d}\rho \\ &= 2 \int_{-\infty}^{\infty} \sqrt{(\partial_\rho u_0)^2 + (\partial_\rho v_0)^2}\, \sqrt{\tfrac{1}{2}\,\Psi(u_0, v_0)}\, \mathrm{d}\rho \end{aligned} \tag{A.16}$$

of the solutions (u_0, v_0) above. The formula (A.16) coincides with σ^{ij} in (1.1) if (u_0, v_0), upon rescaling, is in fact the minimum in (1.1), see Sternberg (1991). Numerical computations indicate that (u_0, v_0) is indeed the minimizer in (1.1). For the solutions $(u_0, v_0) = (1, \bar{v})$ at the grain boundary, and $(u_0, v_0) = (\bar{u}, \pm\frac{1+\bar{u}}{\sqrt{3}})$ at the material boundary we obtain that

$$\sigma_{grain} = \tfrac{2}{3}\,\pi\,(1-\mu)^{\frac{1}{2}} \qquad \text{and} \qquad \sigma_{mat} = \tfrac{2}{3}\,\pi\,(1-\tfrac{\mu}{4})^{\frac{1}{2}}\,, \tag{A.17}$$

respectively.

The solutions (u_0, v_0) connect the corners of the triangle $\mathcal{K}$ via paths which lie entirely on an edge of $\mathcal{K}$. For arbitrary potentials, solutions of this form in general do not exist, see Garcke, Haas, and Stinner (2005). In what follows we will always assume that such solutions to (A.13) exist, which for Ψ as in (1.6)–(1.7) is guaranteed if $\mu \in (-2, \frac{4}{7})$, and that to leading order (u_0, v_0) is the solution of (1.5c).

Next we derive an equation for the grain boundary in the sharp interface limit. First of all we require that

$$(u_0, v_0) + \gamma\,(u_1, v_1) + \gamma^2\,(u_2, v_2) + \ldots \in \mathcal{K} \tag{A.18}$$

to all orders. Since $(u_0, v_0) = (1, \overline{v})$ we obtain to the order $\mathcal{O}(\gamma)$ that $u_1 \leq 0$. In addition we obtain that

$$-u_1 \pm \sqrt{3}\,v_1 \leq 0 \quad \text{if} \quad (u_0, v_0) = (1, \pm\tfrac{2}{\sqrt{3}})\,.$$

This ensures that (A.18) is fulfilled if (u_0, v_0) lies in a corner. Above and in what follows we will always consider the two void/grain interfaces in combination. If a choice has to be made for the sign we always take the upper sign for the CA interface and the lower sign for the AB interface.

We now plug the asymptotic ansatz for u, v, w and z into the variational inequality (1.5c) and require that it holds for all

$$(\eta_1, \eta_2) = (\eta_{10}, \eta_{20}) + \gamma\,(\eta_{11}, \eta_{21}) + \gamma^2\,(\eta_{12}, \eta_{22}) + \ldots\,,$$

which are assumed to have, to all orders, values in $\mathcal{K}$. To the order $\mathcal{O}(1)$ we obtain, on noting (A.12), that $(u_1, v_1) : \mathbb{R} \to \mathbb{R}^2$ has to fulfil $(u_0, v_0) + \gamma\,(u_1, v_1) \in \mathcal{K}$ to the order $\mathcal{O}(\gamma)$ and

$$\begin{aligned}(-\partial_{\rho\rho} u_1 + \kappa\,\partial_\rho u_0 + \Psi_{,uu}(u_0, v_0)\,u_1 + \Psi_{,uv}(u_0, v_0)\,v_1 - w_0)\,(\eta_{10} - u_0)\\ + (-\partial_{\rho\rho} v_1 + \kappa\,\partial_\rho v_0 + \Psi_{,uv}(u_0, v_0)\,u_1 + \Psi_{,vv}(u_0, v_0)\,v_1 - z_0)\,(\eta_{20} - v_0)\\ + (-\partial_{\rho\rho} u_0 + \Psi_{,u}(u_0, v_0))\,(\eta_{11} - u_1) + (-\partial_{\rho\rho} v_0 + \Psi_{,v}(u_0, v_0))\,(\eta_{21} - v_1) \geq 0 \quad \text{(A.19)}\end{aligned}$$

for all $(\eta_{10}, \eta_{20}) : \mathbb{R} \to \mathcal{K}$ and $(\eta_{11}, \eta_{21}) : \mathbb{R} \to \mathbb{R}^2$ which fulfil $(\eta_{10}, \eta_{20}) + \gamma\,(\eta_{11}, \eta_{21}) \in \mathcal{K}$ to the order $\mathcal{O}(\gamma)$. Choosing $(\eta_{10}, \eta_{20}) = (u_0, v_0)$, and as $u_0 = 1$, we obtain that

$$\Psi_{,u}(1, v_0)\,(\eta_{11} - u_1) + (-\partial_{\rho\rho} v_0 + \Psi_{,v}(1, v_0))\,(\eta_{21} - v_1) \geq 0 \quad \text{(A.20)}$$

for all $(\eta_{11}, \eta_{21}) : \mathbb{R} \to \mathbb{R}^2$ with $\eta_{11} \leq 0$. In addition, we have to impose that

$$-\eta_{11} \pm \sqrt{3}\,\eta_{21} \leq 0 \quad \text{if} \quad (u_0, v_0) = (1, \pm\tfrac{2}{\sqrt{3}})\,. \quad \text{(A.21)}$$

If $|v_0| < \frac{2}{\sqrt{3}}$, we have that $-\partial_{\rho\rho} v_0 + \Psi_{,v}(u_0, v_0) = 0$ and it follows from (A.20) that

$$0 \leq \Psi_{,u}(1, v_0)\,(\eta_{11} - u_1) = -\tfrac{1}{3}\,(2 + \mu)\,(\eta_{11} - u_1) \qquad \forall\ \eta_{11} \leq 0;$$

which implies that $u_1 = 0$ as $\mu > -2$. In the interior of the set $\{|v_0| = \frac{2}{\sqrt{3}}\}$, we obtain from (A.20) that

$$\begin{aligned}0 &\leq \Psi_{,u}(1, \pm\tfrac{2}{\sqrt{3}})\,(\eta_{11} - u_1) + \Psi_{,v}(1, \pm\tfrac{2}{\sqrt{3}})\,(\eta_{21} - v_1)\\ &= -\tfrac{1}{3}\,(2 + \mu)\,(\eta_{11} - u_1) - (1 - \mu)\,(\pm\tfrac{2}{\sqrt{3}})\,(\eta_{21} - v_1)\end{aligned}$$

for all (η_{11}, η_{21}) that fulfil $\eta_{11} \leq 0$ and (A.21). We seek a solution (u_1, v_1) of this variational inequality in the cone $\{(u_1, v_1) : u_1 \leq 0, -u_1 \pm \sqrt{3}\,v_1 \leq 0\}$, where this constraint on (u_1, v_1) follows from the $\mathcal{O}(\gamma)$ condition in (A.18). It is easily deduced that only the trivial solution $(0, 0)$ exists if $\mu \in (-2, \frac{4}{7})$. Hence we obtain that $(u_1, v_1) = (0, 0)$ if $(u_0, v_0) = (1, \pm\frac{2}{\sqrt{3}})$.

For points in the set $\{|v_0| < \frac{2}{\sqrt{3}}\}$ we now choose $(\eta_{11}, \eta_{21}) = (0, 0)$ and $(\eta_{10}, \eta_{20}) = (u_0, v_0 + \delta)$ with some small $\delta \in \mathbb{R}$ in the variational inequality (A.19). This yields that

$$-\partial_{\rho\rho} v_1 + \kappa\,\partial_\rho v_0 + \Psi_{,vv}(u_0, v_0)\,v_1 - z_0 = 0 \quad \text{in} \quad \{|v_0| < \tfrac{2}{\sqrt{3}}\}\,.$$

Multiplying this identity by $\partial_\rho v_0$, leads after integration, and integration by parts, to

$$\kappa \int_{-\infty}^{\infty} (\partial_\rho v_0)^2 \,\mathrm{d}\rho = \int_{-\infty}^{\infty} z_0\,\partial_\rho v_0 \,\mathrm{d}\rho\,; \quad \text{(A.22)}$$

where we have used the facts that $\partial_\rho u_0 = 0$, that $-\partial_{\rho\rho} v_0 + \Psi_{,v}(u_0, v_0) = 0$ and that $v_1 = 0$ on $\{|v_0| = \frac{2}{\sqrt{3}}\}$. On noting (A.7) and (A.6), equation (1.5b) to the order $\mathcal{O}(1)$ now gives

$$z_0 = 0 \quad \text{if} \quad \ell(\gamma) := \beta\,\gamma^2 \qquad \text{and} \qquad z_0 = \beta\,\mathcal{V}\,\partial_\rho v_0 \quad \text{if} \quad \ell(\gamma) := \beta\,\gamma\,. \quad \text{(A.23)}$$

Therefore depending on the scaling in (1.5b), we obtain on the grain boundary that

$$\kappa = 0 \quad \text{if} \quad \ell(\gamma) := \beta\,\gamma^2 \qquad \text{and} \qquad \beta\,\omega\,\mathcal{V} = \sigma\,\kappa \quad \text{if} \quad \ell(\gamma) := \beta\,\gamma\,; \tag{A.24}$$

where $\omega := \int_{-\infty}^{\infty} (\partial_\rho v_0)^2 \,\mathrm{d}\rho = \sigma_{grain} = \frac{2}{3}\,\pi\,(1-\mu)^{\frac{1}{2}}$, on recalling (A.16)–(A.17). Obviously, the factors ω and σ cancel in (A.24). However, for later developments, concerning triple junctions, we do not remove them.

Let us remark on the scaling $\ell(\gamma) := \beta\,\gamma^2$. In order to derive an asymptotic expansion around a sharp interface solution we require zero curvature, $\kappa = 0$, of the grain boundaries. Finally we point out that (1.5a) degenerates on grain boundaries, i.e. we obtain $\frac{\partial u}{\partial t} = 0$, and (1.5e) has no interfacial structure on grain boundaries since $c(u_0)$ is constant.

Deriving the governing equation for the void boundaries is more involved. From (A.7), (A.8) and (A.5) we obtain, on dropping the $\widehat{\ }$ notation,

$$\nabla_x \,.\, (b(u)\nabla_x w) = \gamma^{-2}\,\partial_\rho(b(u)\,\partial_\rho w) + \gamma^{-1}\,b(u)\,\partial_\rho w\,\partial_s(\nabla_x d) \,.\, \nabla_x s + \partial_s(b(u)\,\partial_s w\,\nabla_x s) \,.\, \nabla_x s\,. \tag{A.25}$$

Similar expressions can be obtained for $\nabla \,.\, (b(u)\,\nabla\phi)$ and $\nabla \,.\, (c(u)\,\nabla\phi)$. Hence on noting (A.25) and (A.7), the equations (1.5a) and (1.5e) to the order $\mathcal{O}(\gamma^{-2})$ imply on integrating with respect to ρ and matching that

$$\partial_\rho(w_0 + \alpha\,\phi_0) = 0 \qquad \text{and} \qquad \partial_\rho\phi_0 = 0\,.$$

As $\partial_\rho w_0 = 0$, similarly we obtain to the order $\mathcal{O}(\gamma^{-1})$ that

$$\partial_\rho(w_1 + \alpha\,\phi_1) = 0 \qquad \text{and} \qquad \partial_\rho\phi_1 = 0\,.$$

To the order $\mathcal{O}(1)$ we obtain from (1.5a), (A.7), (A.6), (A.25) and (A.5), since u_0 does not depend on s, that

$$-\mathcal{V}\,\partial_\rho u_0 = \partial_\rho(b(u_0)\,\partial_\rho(w_2 + \alpha\,\phi_2)) + b(u_0)\,\partial_{ss}(w_0 + \alpha\,\phi_0)\,.$$

After integration with respect to ρ we obtain

$$-\mathcal{V}\,[u_0]_i^j = M\,\partial_{ss}(w_0 + \alpha\,\phi_0) \tag{A.26}$$

where $[u_0]_i^j$ denotes the jump across the interface Γ^{ij} (the value for $\rho \to \infty$ minus the value for $\rho \to -\infty$) and $M := \int_{-\infty}^{\infty} b(u_0(\rho))\,\mathrm{d}\rho = \rho_m = \pi\,(4-\mu)^{-\frac{1}{2}}$.

It remains to exploit (A.19) at a void interface. At a void interface we have

$$-u_0 \pm \sqrt{3}\,v_0 = 1\,. \tag{A.27}$$

Let us first consider points such that $|u_0| < 1$. In order to fulfil (A.18) to the order $\mathcal{O}(\gamma)$ we need to have

$$-u_1 \pm \sqrt{3}\,v_1 \leq 0\,. \tag{A.28}$$

Choosing $(\eta_{10}, \eta_{20}) = (u_0, v_0) = (\bar{u}, \pm\frac{1+\bar{u}}{\sqrt{3}})$ in (A.19) we obtain that

$$(-\partial_{\rho\rho}u_0 + \Psi_{,u}(u_0,v_0))\,(\eta_{11} - u_1) + (-\partial_{\rho\rho}v_0 + \Psi_{,v}(u_0,v_0))\,(\eta_{21} - v_1) \geq 0 \tag{A.29}$$

for all $(\eta_{11}, \eta_{21}) : \mathbb{R} \to \mathbb{R}^2$ with

$$-\eta_{11} \pm \sqrt{3}\,\eta_{21} \leq 0$$

which in addition fulfil $\eta_{11} \leq 0$ if $(u_0, v_0) = (1, \pm\frac{2}{\sqrt{3}})$ or $\eta_{11} \geq 0$ if $(u_0, v_0) = (-1, 0)$. The variational inequality (A.13), recall (A.15), implies that

$$-\partial_{\rho\rho}(\sqrt{3}\,u_0 \pm v_0) + \sqrt{3}\,\Psi_{,u}(u_0, v_0) \pm \Psi_{,v}(u_0, v_0) = 0 \quad \text{in} \quad \{|u_0| < 1\}\,. \tag{A.30}$$

Taking second derivatives in (A.27) we obtain from (A.30) after solving a linear system for $(\partial_{\rho\rho}u_0, \partial_{\rho\rho}v_0)$ that

$$\partial_{\rho\rho}u_0 = \pm\sqrt{3}\,\partial_{\rho\rho}v_0 = \tfrac{1}{4}(3\,\Psi_{,u} \pm \sqrt{3}\,\Psi_{,v})\,. \tag{A.31}$$

Hence (A.29) yields if $|u_0| < 1$ that

$$(\Psi_{,u}(u_0, v_0) \mp \sqrt{3}\,\Psi_{,v}(u_0, v_0))\,(\eta_{11} - u_1) + (\mp\sqrt{3}\,\Psi_{,u} + 3\,\Psi_{,v})\,(\eta_{21} - v_1) \geq 0 \tag{A.32}$$

for all (η_{11}, η_{21}) with $-\eta_{11} + \sqrt{3}\,\eta_{21} \leq 0$. Now we represent (u_1, v_1) as

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \omega_1 \begin{pmatrix} -1 \\ \pm\sqrt{3} \end{pmatrix} + \omega_2 \begin{pmatrix} \sqrt{3} \\ \pm 1 \end{pmatrix}\,,$$

and note that (A.28) implies that $\omega_1 \leq 0$. Choosing $(\eta_{11}, \eta_{21}) = \omega_2\,(\sqrt{3}, \pm 1)$ in (A.32) gives

$$\begin{aligned} 0 &\geq \omega_1\,(-\Psi_{,u} \pm \sqrt{3}\,\Psi_{,v} - 3\,\Psi_{,u} \pm 3\sqrt{3}\,\Psi_{,v})(u_0, v_0) \\ &= 4\,\omega_1\,(-\Psi_{,u} \pm \sqrt{3}\,\Psi_{,v})(u_0, v_0) = 4\,\omega_1\,(-\tfrac{4}{3}(1 - \mu) + \mu\,u_0)\,. \end{aligned}$$

The term in the last bracket is always negative provided that $\mu < \frac{4}{7}$. This implies that $\omega_1 \geq 0$, and hence $\omega_1 = 0$, which in turn leads to $-u_1 \pm \sqrt{3}\,v_1 = 0$.

For points that lie in the interior of the set $\{(u_0, v_0) = (1, \pm\frac{2}{\sqrt{3}})\}$ we can argue as in the case of a grain boundary to obtain that $(u_1, v_1) = (0, 0)$. Now we consider points that lie in the interior of the set $\{(u_0, v_0) = (-1, 0)\}$. For these points the inequality (A.19) yields on choosing $(\eta_{10}, \eta_{20}) = (u_0, v_0)$

$$0 \leq \Psi_{,u}(-1, 0)\,(\eta_{11} - u_1) + \Psi_{,v}(-1, 0)\,(\eta_{21} - v_1) = \tfrac{1}{3}(4 - \mu)\,(\eta_{11} - u_1)$$

which has to hold for all (η_{11}, η_{21}) fulfilling $\sqrt{3}\,|\eta_{21}| \leq \eta_{11}$. Since by (A.18) the solution (u_1, v_1) has to satisfy $\sqrt{3}\,|v_1| \leq u_1$, we obtain that $(u_1, v_1) = (0, 0)$ is the only solution to the above variational inequality.

For points in the set $\{|u_0| < 1\}$ we now choose $(\eta_{11}, \eta_{21}) = (0, 0)$ and $(\eta_{10}, \eta_{20}) = (u_0 + \sqrt{3}\,\delta, v_0 \pm \delta)$ with some small $\delta \in \mathbb{R}$ in the variational inequality (A.19). This yields that

$$\begin{aligned} &-\partial_{\rho\rho}(\sqrt{3}\,u_1 \pm v_1) + \kappa\,\partial_\rho(\sqrt{3}\,u_0 \pm v_0) + (\sqrt{3}\,\Psi_{,uu}(u_0, v_0)\,u_1 + \sqrt{3}\,\Psi_{,uv}(u_0, v_0)\,v_1 \\ &\qquad \pm \Psi_{,vu}(u_0, v_0)\,u_1 \pm \Psi_{,vv}(u_0, v_0)\,v_1) - \sqrt{3}\,w_0 \mp z_0 = 0\,. \end{aligned}$$

As $-u_0 \pm \sqrt{3}\, v_0 = 1$ and $-u_1 \pm \sqrt{3}\, v_1 = 0$, it follows from the above that

$$-4\,\partial_{\rho\rho} u_1 + 4\,\kappa\,\partial_\rho u_0 - 3\,u_1 - (1-\mu)\,u_1 - 3\,w_0 \mp \sqrt{3}\,z_0 = 0\,.$$

Similarly to (A.22), on multiplying the above identity by $\partial_\rho u_0$, integrating, performing integration by parts; we obtain, on noting (A.15) and (A.31) that

$$\begin{aligned}
4\,\kappa \int_{-\infty}^{\infty} (\partial_\rho u_0)^2 \,\mathrm{d}\rho - 3 \int_{-\infty}^{\infty} w_0\,\partial_\rho u_0 \,\mathrm{d}\rho \mp \sqrt{3} \int_{-\infty}^{\infty} z_0\,\partial_\rho u_0 \,\mathrm{d}\rho \\
= 3\,\kappa \int_{-\infty}^{\infty} [(\partial_\rho u_0)^2 + (\partial_\rho v_0)^2] \,\mathrm{d}\rho - 3\,w_0\,[u_0]_i^j \mp \sqrt{3} \int_{-\infty}^{\infty} z_0\,\partial_\rho u_0 \,\mathrm{d}\rho = 0\,.
\end{aligned}$$

Equation (1.5b) gives to the order $\mathcal{O}(1)$ the identities (A.23) and hence we get, on recalling (A.16),

$$\sigma\,\kappa = [u_0]_i^j\,w_0 \quad \text{if} \quad \ell(\gamma) := \beta\,\gamma^2 \qquad \text{and} \qquad \sigma\,\kappa = [u_0]_i^j\,w_0 + \beta\,\omega\,\mathcal{V} \quad \text{if} \quad \ell(\gamma) := \beta\,\gamma\,; \tag{A.33}$$

where $\omega := \int_{-\infty}^{\infty} (\partial_\rho v_0)^2 \,\mathrm{d}\rho = \frac{1}{4}\,\sigma_{mat} = \frac{\pi}{6}\,(1 - \frac{\mu}{4})^{\frac{1}{2}}$.

For the material interfaces $(i,j) = (A,B)$, (C,A), and the grain interface (B,C), we derive from (A.24), (A.26) and (A.33) for the scaling $\ell(\gamma) := \beta\,\gamma$ that

$$-2\,\mathcal{V}^{AB} = M^{AB}\,\partial_{ss}(w_0^{AB} + \alpha\,\phi_0^{AB}) \quad \text{and} \quad 2\,w_0^{AB} + \beta\,\omega^{AB}\,\mathcal{V}^{AB} = \sigma^{AB}\,\kappa^{AB}, \tag{A.34a}$$

$$2\,\mathcal{V}^{CA} = M^{CA}\,\partial_{ss}(w_0^{CA} + \alpha\,\phi_0^{CA}) \quad \text{and} \quad -2\,w_0^{CA} + \beta\,\omega^{CA}\,\mathcal{V}^{CA} = \sigma^{CA}\,\kappa^{CA}, \tag{A.34b}$$

$$\beta\,\omega^{BC}\,\mathcal{V}^{BC} = \sigma^{BC}\,\kappa^{BC}; \tag{A.34c}$$

where, on recalling (A.17), we have that

$$\begin{aligned}
\omega^{BC} = \sigma^{BC} = \sigma_{grain} = \tfrac{2}{3}\,\pi\,(1-\mu)^{\frac{1}{2}}\,, \qquad M^{AB} = M^{CA} = \pi\,(4-\mu)^{-\frac{1}{2}}\,, \\
4\,\omega^{AB} = 4\,\omega^{CA} = \sigma^{AB} = \sigma^{CA} = \sigma_{mat} = \tfrac{2}{3}\,\pi\,(1 - \tfrac{\mu}{4})^{\frac{1}{2}}\,.
\end{aligned}$$

The evolution laws (A.34a,b) for the material interfaces combine surface diffusion and surface attachment limited kinetics (SALK), which was discussed in Taylor and Cahn (1994); see also Elliott and Garcke (1997).

If we choose the scaling $\ell(\gamma) := \beta\,\gamma^2$ instead of $\ell(\gamma) := \beta\,\gamma$ in the evolution equation (1.5b) we derive from (A.26), (A.24) and (A.33) that

$$\begin{aligned}
&\mathcal{V}^{AB} = -\tfrac{M^{AB}}{2}\,\partial_{ss}(\tfrac{\sigma^{AB}}{2}\kappa^{AB} + \alpha\,\phi_0^{AB}), \qquad \mathcal{V}^{CA} = -\tfrac{M^{CA}}{2}\,\partial_{ss}(\tfrac{\sigma^{CA}}{2}\kappa^{CA} - \alpha\,\phi_0^{CA}), \\
&\sigma^{BC}\,\kappa^{BC} = 0.
\end{aligned}$$

Therefore under this scaling the evolution of the void surface is given by surface diffusion, see Cahn, Elliott, and Novick-Cohen (1996), whereas the grain boundaries have zero mean curvature, i.e. they are in equilibrium.

It remains to derive the equations at a triple junction. From now on, we will always denote by superscripts A, B and C quantities that are defined on the interfaces BC, CA

and AB, respectively. In particular we have that the normals n_Γ^A, n_Γ^B and n_Γ^C are such that n_Γ^A points into C, n_Γ^B points into A and n_Γ^C points into B. At a triple junction $m(t)$ we choose at a fixed time t a triangle T_γ, whose midpoint coincides with the triple junction. In addition it is assumed that the sides of the triangle intersect the interfaces to leading order perpendicularly and have to leading order a length which is proportional to $\gamma^{\frac{1}{2}}$. We now introduce the stretched variable $y = \gamma^{-1}(x - m(t))$, and make the asymptotic ansatz

$$(u, v)(x, t) = (U_0, V_0)(y, t) + \gamma\, (U_1, V_1)(y, t) + \dots.$$

Then (1.5c) gives to leading order that the following variational inequality has to hold almost everywhere on $\widetilde{T}_\gamma := \{y \in \mathbb{R}^2 \,|\, m(t) + \gamma\, y \in T_\gamma\}$:

$$(-\Delta_y U_0 + \Psi_{,u}(U_0, V_0))\,(\eta_1 - U_0) + (-\Delta_y V_0 + \Psi_{,u}(U_0, V_0))\,(\eta_2 - V_0) \geq 0 \qquad \text{(A.36)}$$

for all $(\eta_1, \eta_2) : \widetilde{T}_\gamma \to \mathcal{K}$. We now want to derive a solvability condition for (A.36), which will lead to an angle condition at the triple junction. The ansatz $(\eta_1, \eta_2) = (U_0, V_0) \pm \delta(\partial_{y_l} U_0, \partial_{y_l} V_0)$, $l = 1$, 2, leads to values in $\mathcal{K}$ for small δ in the following cases. If $(U_0, V_0)(y, t)$ lies in the interior of $\mathcal{K}$, then this is obviously true. If $(U_0, V_0)(y, t)$ lies in the interior of one of the sets $\{(u, v) = i\}$ with $i \in \{A, B, C\}$, we obtain $\nabla_y U_0 = 0$ and hence $(\eta_1, \eta_2) = (U_0, V_0)$. In the case of points that lie in the interior of one of the three sets $\{(u, v) : -u \pm \sqrt{3}\, v = 1 \text{ and } |u| < 1\}$ and $\{(u, v) : u = 1 \text{ and } |v| < \frac{2}{\sqrt{3}}\}$, we obtain also that $(\eta_1, \eta_2) \in \mathcal{K}$ for small δ. For example, if $|U_0| < 1$ and $-U_0 \pm \sqrt{3}\, V_0 = 1$ in a neighbourhood of (y, t), then we obtain $-\partial_{y_l} U_0 \pm \sqrt{3}\, \partial_{y_l} V_0 = 0$ and hence $(U_0, V_0) \pm \delta(\partial_{y_l} U_0, \partial_{y_l} V_0) \in \mathcal{K}$ if δ is sufficiently small.

Assuming that the complement of the sets considered above has measure zero, which is supported by numerical experiments, we obtain from (A.36) with $(\eta_1, \eta_2) = (u_0, v_0) \pm \delta\,(\partial_{y_l} u_0, \partial_{y_l} v_0)$, $l = 1$, 2, that

$$(\nabla_y U_0)^T(-\Delta_y U_0 + \Psi_{,u}(U_0, V_0)) + (\nabla_y V_0)^T(-\Delta_y V_0 + \Psi_{,v}(U_0, V_0)) = 0$$

almost everywhere; where $\nabla_y \cdot = (\partial_{y_1} \cdot, \partial_{y_2} \cdot)^T$. Defining $\Lambda_0 := (U_0, V_0)^T$ and using the identity

$$-(\nabla_y \Lambda_0)^T(\Delta_y \Lambda_0) = -\nabla_y \,.\, (\,(\nabla_y \Lambda_0)^T(\nabla_y \Lambda_0)) + \tfrac{1}{2}\,(\nabla_y[\,|\nabla_y \Lambda_0|^2]) \,;$$

we obtain, after integration over $\widetilde{T}_\gamma$, that the following identity holds

$$\begin{aligned} 0 &= \int_{\widetilde{T}_\gamma} \left[-(\nabla_y \Lambda_0)^T(\Delta_y \Lambda_0) + (\nabla_y \Lambda_0)^T[(\Psi_{,u}, \Psi_{,v})(\Lambda_0)]^T\,\right] \,\mathrm{d}y \\ &= \int_{\widetilde{T}_\gamma} \left[-\nabla_y \,.\, (\,(\nabla_y \Lambda_0)^T(\nabla_y \Lambda_0)) + \nabla_y(\tfrac{1}{2}\,|\nabla_y \Lambda_0|^2 + \Psi(\Lambda_0))\right] \,\mathrm{d}y \\ &= -\int_{\partial\widetilde{T}_\gamma} (\nabla_y \Lambda_0)^T\,(\nabla_y \Lambda_0)\, n_{\partial T} \,\mathrm{d}s_T + \int_{\partial\widetilde{T}_\gamma} (\tfrac{1}{2}\,|\nabla_y \Lambda_0|^2 + \Psi(\Lambda_0))\, n_{\partial T} \,\mathrm{d}s_T \,; \end{aligned}$$

where we have applied the Gauss theorem to obtain the last identity. Moreover, $n_{\partial T}$ is the outer unit normal to $\partial\widetilde{T}_\gamma$. Since we chose the triangle T_γ such that ∂T_γ intersects the interfaces asymptotically perpendicularly, we obtain that the term $\int_{\partial\widetilde{T}_\gamma} (\nabla_y \Lambda_0)^T(\nabla_y \Lambda_0)\, n_{\partial T} \,\mathrm{d}s_T$

vanishes asymptotically. Recalling (A.16), matching Λ_0 and the standing wave (u_0, v_0), and noting that asymptotically $n_{\partial T}$ equals τ_Γ^i along the different sides for $i \in \{A, B, C\}$, we obtain that

$$0 = \sum_{i \in \{A,B,C\}} \sigma^i \, \tau_\Gamma^i \,.$$

This is the force balance at the triple junction and a simple computation shows that the above identity is equivalent to Young's law,

$$\frac{\sin \theta^A}{\sigma^A} = \frac{\sin \theta^B}{\sigma^B} = \frac{\sin \theta^C}{\sigma^C}, \tag{A.37}$$

where θ^A, θ^B and θ^C are the angles that the regions A, B and C form at the triple junction (see also Bronsard and Reitich (1993), Bronsard, Garcke, and Stoth (1998), and Garcke and Novick-Cohen (2000)).

To obtain a flux balance condition we consider the mass balance (1.5a). We observe that only the second term on the left hand side of (1.5a) in (P) gives a contribution to leading order. Integrating the leading order term over $\widetilde{T}_\gamma$, we obtain that

$$0 = \int_{\widetilde{T}_\gamma} \nabla_y \,.\, (\, b(U_0)\, \nabla_y [W_0 + \alpha\, \Phi_0])\; \mathrm{d}y = \int_{\partial \widetilde{T}_\gamma} b(U_0)\, \nabla_y [W_0 + \alpha\, \Phi_0] \,.\, n_{\partial T}\; \mathrm{d}s_T .$$

The right hand side gives a contribution only if $b(U_0) \neq 0$ which means only on the material interfaces, AB and CA. Matching with the inner solutions, using (A.7) and $\nabla_x s = n_{\partial T} + \mathcal{O}(\gamma)$, we obtain that

$$\left[\int_{-\infty}^{\infty} b(u_0^C(\rho))\; \mathrm{d}\rho\right] \partial_s (w_0^C + \alpha\, \phi_0^C) + \left[\int_{-\infty}^{\infty} b(u_0^B(\rho))\; \mathrm{d}\rho\right] \partial_s (w_0^B + \alpha\, \phi_0^B) = 0,$$

where u_0^C, w_0^C, ϕ_0^C and u_0^B, w_0^B, ϕ_0^B are the inner leading order solutions at the interfaces AB and CA, respectively. Altogether at the triple junction we obtain the flux balance condition

$$M^C\, \partial_s (w_0^C + \alpha\, \phi_0^C) + \, M^B\, \partial_s (w_0^B + \alpha\, \phi_0^B) = 0 \,. \tag{A.38}$$

It remains to determine an additional condition at the triple junction, which is related to the fact that the chemical potential is continuous. Neglecting lower order terms in (1.5a), we obtain close to the triple junction that

$$\begin{aligned} 0 &= \int_{\widetilde{T}_\gamma} [\nabla_y \,.\, (b(U_0)\, \nabla_y (W_0 + \alpha\, \Phi_0))\,]\, (W_0 + \alpha\, \Phi_0)\, \mathrm{d}y \\ &= - \int_{\widetilde{T}_\gamma} b(U_0)\, |\nabla_y (W_0 + \alpha\, \Phi_0)|^2\, \mathrm{d}y + \int_{\partial \widetilde{T}_\gamma} b(U_0)\, (W_0 + \alpha\, \Phi_0)\, \nabla_y (W_0 + \alpha\, \Phi_0) \,.\, n_{\partial T}\; \mathrm{d}s_T \,. \end{aligned}$$

The choice of $\widetilde{T}_\gamma$ yields that $\nabla_y (W_0 + \alpha\, \Phi_0) \,.\, n_{\partial T}$ results in a partial derivative along the interface. Since the s-variable in the inner expansion is scaled in a different way we obtain from matching the inner expansion to the triple junction expansion that $\nabla_y (W_0 +$

$\alpha\,\Phi_0)\,.\,n_{\partial T}$ has to vanish to leading order. Hence to leading order at the triple junction $W_0 + \alpha\,\Phi_0$ is constant on the support of $b(U_0)$ which is assumed to be connected. By matching the solution close to the triple junction with the inner solution we obtain that the limit for $w_0 + \alpha\,\phi_0$ coming from the AB interface has to equal that coming from the CA interface. Assuming that the ϕ equation has a continuous solution up to the boundary, we obtain that at the triple junction

$$w_0^C = w_0^B. \tag{A.39}$$

We remark that the choice of scaling $\ell(\gamma) := \beta\,\gamma$ or $\beta\,\gamma^2$ does not effect the conditions (A.37), (A.38) and (A.39) at the triple junction, as the equation (1.5b) was not used to derive them. Of course, under the scaling $\ell(\gamma) := \beta\,\gamma^2$ we deduce from (A.33) and (A.39) that at the triple junction

$$\sigma^C\,\kappa^C = -\sigma^B\,\kappa^B. \tag{A.40}$$

Finally, when an interface meets the external boundary, further boundary conditions have to hold which can be derived as in Garcke and Novick-Cohen (2000) and Novick-Cohen (2000). We include these conditions in the summary below. Let us also point out that the ideas presented above can also be used to handle more general potentials Ψ. In particular the approach used to derive the triple junction conditions can be applied to the setting in Garcke and Novick-Cohen (2000) and Novick-Cohen (2000).

To summarize we obtain, depending on the scaling in (1.5b), the following two sharp interface problems. In both cases we obtain that at a triple junction the identities (A.37), (A.38) and (A.39) have to hold for w and ϕ. When an interface meets an external boundary a 90° angle condition has to hold. In addition, at a material boundary, $\partial_s(w^i + \alpha\,\phi^i) = 0$ for $i \in \{B, C\}$. Firstly, the scaling $\ell(\gamma) := \beta\,\gamma^2$ leads to

$$\begin{array}{rcll} \mathcal{V}^C &=& -\frac{M^C}{2}\,\partial_{ss}(\frac{\sigma^C}{2}\,\kappa^C + \alpha\,\phi^C) & \text{on } \Gamma^C, \\ \mathcal{V}^B &=& -\frac{M^B}{2}\,\partial_{ss}(\frac{\sigma^B}{2}\,\kappa^B - \alpha\,\phi^B) & \text{on } \Gamma^B, \\ 0 &=& \kappa^A & \text{on } \Gamma^A. \end{array} \tag{A.41}$$

Whilst for the scaling $\ell(\gamma) := \beta\,\gamma$, we obtain that

$$\begin{array}{rcll} \pm 2\,\mathcal{V}^i &=& M^i\,\partial_{ss}\,(w^i + \alpha\,\phi^i) & \text{on } \Gamma^i \text{ for } i \in \{B, C\}, \\ \beta\,\omega^i\,\mathcal{V}^i &=& \sigma^i\,\kappa^i \pm 2\,w^i & \text{on } \Gamma^i \text{ for } i \in \{B, C\}, \\ \beta\,\omega^A\,\mathcal{V}^A &=& \sigma^A\,\kappa^A & \text{on } \Gamma^A; \end{array} \tag{A.42}$$

where in the $\pm$ option we take the top for $i = B$ and the bottom for $i = C$. Furthermore, the limiting electric potential satisfies

$$\Delta\phi = 0 \quad \text{in} \quad \Omega \setminus \overline{\Omega^A(t)}\,, \qquad \frac{\partial\phi}{\partial n_\Gamma} = 0 \quad \text{on} \quad \Gamma^B \cup \Gamma^C$$

where $\Omega^A(t)$ is the void with boundary $\Gamma^B \cup \Gamma^C$.

The total area occupied by the void (and hence also the total area occupied by the material) is conserved by the above flows. Let $a(t)$ be the total volume of the void at time t. By using a transport theorem for area, see e.g. Gurtin (1993), we obtain that

$$\begin{aligned}\tfrac{d}{dt}a(t) &= -\int_{\Gamma^B(t)} \mathcal{V}^B \,\mathrm{d}s + \int_{\Gamma^C(t)} \mathcal{V}^C \,\mathrm{d}s \\ &= -\tfrac{1}{2}\int_{\Gamma^B(t)} M^B\, \partial_{ss}(w^B + \alpha\,\phi^B)\,\mathrm{d}s - \tfrac{1}{2}\int_{\Gamma^C(t)} M^C\, \partial_{ss}(w^C + \alpha\,\phi^C)\,\mathrm{d}s = 0\end{aligned}$$

by the flux condition (A.38) and the no flux condition at the outer boundary.

The surface energy of the system at time t is given by $E^s(t) := \sum_{i\in\{A,B,C\}} \int_{\Gamma^i} \sigma^i \,\mathrm{d}s$. We now want to show that in the absence of an electric field, i.e. $\alpha = 0$, the total surface energy is a Lyapunov functional. Firstly, we consider the case of $\ell(\gamma) := \beta\,\gamma$. Using a transport theorem for integrals over the interface, see e.g. Gurtin (1993), (A.34a,b), (A.37) and the angle condition at the outer boundary we obtain (for more details in a related situation see Garcke and Novick-Cohen (2000) and Novick-Cohen (2000))

$$\begin{aligned}\tfrac{d}{dt}E^s(t) &= -\sum_{i\in\{A,B,C\}} \int_{\Gamma^i(t)} \sigma^i\,\kappa^i\,\mathcal{V}^i\,\mathrm{d}s \\ &= -\beta \sum_{i\in\{A,B,C\}} \omega^i \int_{\Gamma^i(t)} [\mathcal{V}^i]^2\,\mathrm{d}s + 2\left[\int_{\Gamma^B(t)} w^B\,\mathcal{V}^B\,\mathrm{d}s - \int_{\Gamma^C(t)} w^C\,\mathcal{V}^C\,\mathrm{d}s\right] \\ &\leq -\sum_{i\in\{B,C\}} \int_{\Gamma^i(t)} M^i\,|\partial_s\,w^i|^2\,\mathrm{d}s \leq 0. \end{aligned} \qquad \text{(A.43)}$$

We note in the above that terms $[\,M^B\,w^B\,\partial_s\,w^B + M^C\,w^C\,\partial_s\,w^C\,]$ resulting from the integration by parts vanish at the outer boundary and at the triple junction due to the flux condition (A.38), the continuity condition (A.39) and the no flux condition at the outer boundary. The argument in (A.43) is easily adapted to the scaling $\ell(\gamma) := \beta\,\gamma^2$. In conclusion we have for both motions that $\frac{d}{dt}E^s(t) \leq 0$.

Acknowledgement This work was partly supported by the RTN Programme Fronts-Singularities, HPRN-CT-2002-00274.

References

Averbuch, A., M. Israeli, and I. Ravve (2003). Electromigration of intergranular voids in metal films for microelectronic interconnects. *J. Comput. Phys. 186*, 481–502.

Baldo, S. (1990). Minimal interface criterion for phase transitions in mixtures of Cahn–Hilliard fluids. *Ann. Inst. H. Poincaré Anal. Non Linéaire 7*(2), 67–90.

Barrett, J. W. and J. F. Blowey (2001). Finite element approximation of a degenerate Allen–Cahn/Cahn–Hilliard system. *SIAM J. Numer. Anal. 39*, 1598–1624.

Barrett, J. W., J. F. Blowey, and H. Garcke (1999). Finite element approximation of the Cahn–Hilliard equation with degenerate mobility. *SIAM J. Numer. Anal. 37*, 286–318.

Barrett, J. W., J. F. Blowey, and H. Garcke (2001). On fully practical finite element approximations of degenerate Cahn–Hilliard systems. *M2AN Math. Model. Numer. Anal. 35*, 713–748.

Barrett, J. W., H. Garcke, and R. Nürnberg (2005). Finite element approximation of a phase field model for surface diffusion of voids in a stressed solid. *Math. Comp.* (to appear).

Barrett, J. W. and R. Nürnberg (2004). Convergence of a finite element approximation of surfactant spreading on a thin film in the presence of van der Waals forces. *IMA J. Numer. Anal. 24*, 323–363.

Barrett, J. W., R. Nürnberg, and V. Styles (2004). Finite element approximation of a phase field model for void electromigration. *SIAM J. Numer. Anal. 42*, 738–772.

Bhate, D. N., A. Kumar, and A. F. Bower (2000). Diffuse interface model for electromigration and stress voiding. *J. Appl. Phys. 87*, 1712–1721.

Bronsard, L., H. Garcke, and B. Stoth (1998). A multi-phase Mullins-Sekerka system: matched asymptotic expansions and an implicit time discretisation for the geometric evolution problem. *Proc. Roy. Soc. Edinburgh Sect. A 128*(3), 481–506.

Bronsard, L. and F. Reitich (1993). On three-phase boundary motion and the singular limit of a vector-valued Ginzburg–Landau equation. *Arch. Rational Mech. Anal. 124*(4), 355–379.

Cahn, J. W., C. M. Elliott, and A. Novick-Cohen (1996). The Cahn–Hilliard equation with a concentration dependent mobility: motion by minus the Laplacian of the mean curvature. *European J. Appl. Math. 7*, 287–301.

Cahn, J. W. and A. Novick-Cohen (1994). Evolution equations for phase separation and ordering in binary alloys. *J. Statist. Phys. 76*, 877–909.

Cahn, J. W. and A. Novick-Cohen (1996). Limiting motion for an Allen–Cahn/Cahn–Hilliard system. In *Free boundary problems, theory and applications (Zakopane, 1995)*, Volume 363 of *Pitman Res. Notes Math. Ser.*, pp. 89–97. Harlow: Longman.

Cahn, J. W. and A. Novick-Cohen (2000). Motion by curvature and impurity drag: Resolution of a mobility paradox. *Acta mater. 48*, 3425–3440.

Ciarlet, P. G. (1988). *Numerical Linear Algebra and Optimisation.* Cambridge: Cambridge University Press.

Cummings, L. J., G. Richardson, and M. B. Amar (2001). Models of void electromigration. *European J. Appl. Math. 12*, 97–134.

Elliott, C. M. and H. Garcke (1997). Existence results for diffusive surface motion laws. *Adv. Math. Sci. Appl. 7*, 465–488.

Garcke, H., R. Haas, and B. Stinner (2005). On Ginzburg–Landau free energies for multi-phase systems. (in preparation).

Garcke, H. and B. Nestler (2000). A mathematical model for grain growth in thin metallic films. *Math. Models Methods Appl. Sci. 10*, 895–921.

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

Garcke, H. and A. Novick-Cohen (2000). A singular limit for a system of degenerate Cahn–Hilliard equations. *Adv. Differential Equations 5*, 401–434.

Gilbarg, D. and N. S. Trudinger (1983). *Elliptic partial differential equations of second order. Second edition.* Berlin: Springer.

Grün, G. and M. Rumpf (2000). Nonnegativity preserving numerical schemes for the thin film equation. *Numer. Math. 87*, 113–152.

Gurtin, M. E. (1993). *Thermomechanics of Evolving Phase Boundaries in the Plane.* New York: Oxford University Press.

Ito, K. and Y. Kohsaka (2001). Three-phase boundary motion by surface diffusion: stability of a mirror symmetric stationary solution. *Interfaces Free Bound. 3*, 45–80.

Kanel, J., A. Novick-Cohen, and A. Vilenkin (2004). Coupled surface and grain boundary motion: nonclassical traveling wave solutions. *Adv. Differential Equations 9*, 299–327.

Mahadevan, M. and R. M. Bradley (1999). Phase field model of surface electromigration in single crystal metal thin films. *Phys. D 126*, 201–213.

Mullins, W. W. (1958). The effect of thermal grooving on grain boundary motion. *Acta Metall. 6*, 414–427.

Novick-Cohen, A. (2000). Triple-junction motion for an Allen–Cahn/Cahn–Hilliard system. *Phys. D 137*, 1–24.

Novick-Cohen, A. and L. P. Hari (2005). Geometric motion for a degenerate Allen–Cahn/Cahn–Hilliard system: The partial wetting case. *Phys. D 209*, 205–235.

Schmidt, A. and K. G. Siebert (2004). *Design of Adaptive Finite Element Software: The Finite Element Toolbox ALBERTA*, Volume 42 of *Lecture Notes in Computational Science and Engineering.* Berlin: Springer.

Sternberg, P. (1991). Vector-valued local minimizers of nonconvex variational problems. Current directions in nonlinear partial differential equations (Provo, UT, 1987). *Rocky Mountain J. Math. 21*(2), 799–807.

Taylor, J. E. and J. W. Cahn (1994). Linking anisotropic sharp and diffuse surface motion laws via gradient flows. *J. Statist. Phys. 77*, 183–197.

Xia, L., A. F. Bower, Z. Suo, and C. Shih (1997). A finite element analysis of the motion and evolution of voids due to strain and electromigration induced surface diffusion. *J. Mech. Phys. Solids 45*, 1473–1493.

Zhornitskaya, L. and A. L. Bertozzi (2000). Positivity preserving numerical schemes for lubrication-type equations. *SIAM J. Numer. Anal. 37*, 523–555.