A phase field model for
the electromigration of intergranular voids

JOHN W. BARRETT[†]

Department of Mathematics, Imperial College, London, SW7 2AZ, U.K.

HARALD GARCKE[‡]

NWF I – Mathematik, Universität Regensburg, 93040 Regensburg, Germany

AND

ROBERT NÜRNBERG[§]

Department of Mathematics, Imperial College, London, SW7 2AZ, U.K.

We propose a degenerate Allen–Cahn/Cahn–Hilliard system coupled to a quasi-static diffusion equation to model the motion of intergranular voids. The system can be viewed as a phase field system with an interfacial parameter γ. In the limit $\gamma \to 0$, the phase field system models the evolution of voids by surface diffusion and electromigration in an electrically conducting solid with a grain boundary. We introduce a finite element approximation for the proposed system, show stability bounds, prove convergence, and hence existence of a weak solution to this non-linear degenerate parabolic system in two space dimensions. An iterative scheme for solving the resulting non-linear discrete system at each time level is introduced and analysed, and some numerical experiments are presented. In the Appendix we discuss the sharp interface limit of the above degenerate system as the interfacial parameter γ tends to zero.

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 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 [6]). The electric field can cause a so called "electron wind" force and this leads to the transport of atoms which results in the 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 come

[†] E-mail: j.barrett@imperial.ac.uk

[‡] E-mail: harald.garcke@mathematik.uni-regensburg.de

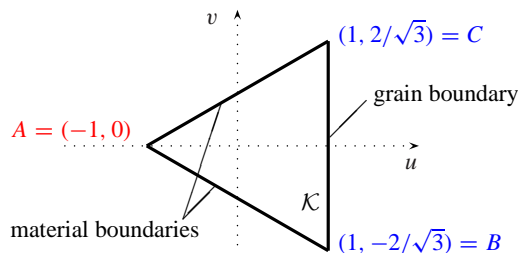[§] E-mail: robert.nurnberg@imperial.ac.uk

FIG. 1. The (u, v) space $\mathcal{K} = \triangle ABC$. Here A represents the void phase, B grain I, and C grain II.

into contact with another grain (i.e. a region with a different lattice orientation). The modelling and computation 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 [1] 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 [1] quite severe symmetry conditions are assumed.

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. [22]). 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), where 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 ($v = 2/\sqrt{3}$) or in grain II ($v = -2/\sqrt{3}$). If $u = -1$ it makes no sense to distinguish between the grains and we set $v = 0$. This means that the three points $A = (-1, 0)$, $B = (1, -2/\sqrt{3})$ and $C = (1, 2/\sqrt{3})$ for (u, v) are relevant to distinguish between void, grain I and grain II (see Figure 1). We choose $\pm 2/\sqrt{3}$ as values for v in the grains because this makes the closed triangular domain $\mathcal{K}$ 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 in [29], and studied later in [10, 9], to include grain boundaries. We make use also of ideas in [14, 15, 16], where a degenerate Allen–Cahn/Cahn–Hilliard equation was studied.

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

$$\mathcal{E}(u, v) := \int_{\Omega} \left[\frac{\gamma}{2} |\nabla u|^2 + \frac{\gamma}{2} |\nabla v|^2 + \gamma^{-1} \Psi(u, v) \right] dx,$$

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. [2, 11, 23] 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, as $\gamma \to 0$, $\mathcal{E}$ leads to a sharp interface free energy with surface energy densities (sometimes also called surface tensions)

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

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 [2, 11, 23] for more details.

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

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

where $\delta\mathcal{E}/\delta u$ and $\delta\mathcal{E}/\delta v$ are the L^2-representations of the variational derivatives of $\mathcal{E}$ with respect to u and v respectively. The equations for w and z are stated here for a smooth Ψ. Later, we will use a non-differentiable Ψ and then we have to solve a variational inequality to compute w and z (see (1.6c) below). The potential w is the chemical potential for the diffusion of atoms in the void-material interfacial layer and z acts as the driving force for the grain boundary motion. Taking into account that diffusion of atoms is also caused by the electrical field $-\nabla\phi$ (see e.g. [29]), we propose that the mass flux $\mathcal{J}$ for the diffusion of atoms is given as

$$\mathcal{J} = -\gamma^{-1} b(u) \nabla[w + \alpha\phi],$$

where α is a constant, $b(u) := 1 - u^2$ is a degenerate mobility and the factor γ^{-1} takes into account that the diffusion is enhanced in the interface, $|u| < 1$. The evolution equation for u now follows from the mass balance law $\partial u/\partial t + \nabla.\mathcal{J} = 0$. The unknown v models the grain boundary and as usual we take a gradient flow dynamics, i.e. $\partial v/\partial t$ points in the direction of the negative gradient of $\mathcal{E}$ with respect to v (see e.g. [22, 23]). Altogether we obtain the following set of evolution equations:

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

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

which are coupled to the equation for the electric potential ϕ,

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

Here $\ell(\gamma)$ is a non-negative coefficient, and later we will use the scalings $\ell(\gamma) := \beta\gamma$ and $\ell(\gamma) := \beta\gamma^2$, where $\beta \in \mathbb{R}_{>0}$. 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 in [13], 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 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 from a similar set of equations introduced by Cahn and Novick-Cohen [14], where the Allen–Cahn equation was also degenerate. For this system, which is a model for simultaneous order-disorder and phase separation, it was shown in [32, 33] that, under an appropriate scaling and under certain assumptions on the interface geometry, one obtains mean curvature flow and surface diffusion coupled at triple junctions in the sharp interface limit. We obtain a similar sharp interface limit also for our system with a non-degenerate Allen–Cahn equation. However, our sharp interface limit is different in some aspects, and leads to some interesting new effects. For example, we derive a model where the grain boundary motion is quasi-static, and another model where viscous effects appear in the surface diffusion equation. Analogous results would follow for the doubly degenerate system derived in [14], if one would work with the scaling used in this paper. We will discuss 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. Here we discuss, for ease of exposition, only the case when no coupling to the electric field is present.

For the scaling $\ell(\gamma) := \beta\gamma^2$ we find that the interfaces which bound the void move by surface diffusion, i.e.

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

where $\mathcal{V}$ is the normal velocity of the interface, κ is the (mean) curvature of the interface, Δ_s is the surface Laplacian, and M and σ are constants, whose precise definitions can be found in the Appendix. For a grain boundary, we show that its mean curvature is zero. These evolution laws are coupled at triple junctions, where 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

$$\mathcal{V} = \frac{M}{4}\Delta_s(-\sigma\kappa + \beta\omega\mathcal{V}), \qquad (1.3)$$

where ω is a constant. This law has been derived in [37], and has been studied in [21]. It links the fourth order surface diffusion flow to a second order flow, which is called motion by averaged mean curvature (see [37, 34] for details). In this context we refer also to work in [31] on the viscous Cahn–Hilliard equation, for which a degenerate variant would lead to (1.3) in the sharp interface limit. 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 [32, 33] a singular limit of an Allen–Cahn/Cahn–Hilliard system has also been analyzed using different scalings.

In the recent paper [9], the following phase field model for void electromigration was considered:

$$\gamma\frac{\partial u}{\partial t} - \nabla.(b(u)\nabla[w + \alpha\phi]) = 0, \qquad w = -\gamma\Delta u + \gamma^{-1}\overline{\Psi}'(u), \qquad \nabla.(c(u)\nabla\phi) = 0 \qquad (1.4)$$
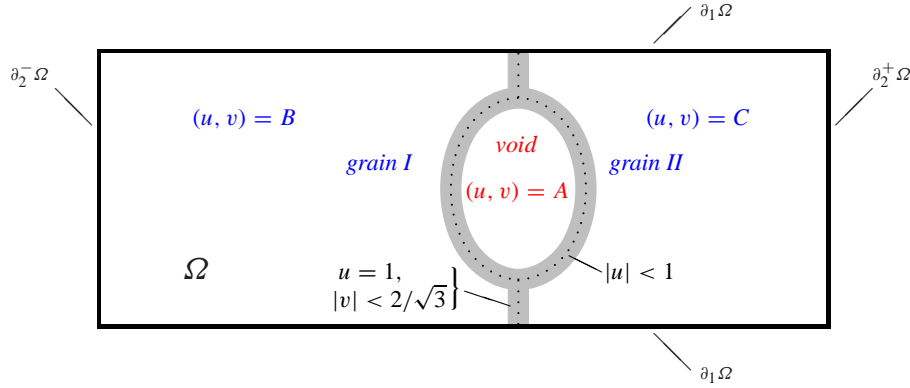
FIG. 2. The order parameters for a typical intergranular void. Note that $A = (-1, 0)$, $B = (1, -2/\sqrt{3})$ and $C = (1, 2/\sqrt{3})$.

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; and hence the second equation in (1.4) is in fact a variational inequality (see [9] for details). While, as in (1.2a–c), $\phi(\cdot, t)$ is the electric potential, $\gamma \in \mathbb{R}_{>0}$ is the interfacial parameter, $\alpha \in \mathbb{R}_{\geqslant 0}$ 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 [13] to show that the zero level sets of u_γ, the solution to (1.4) for a fixed $\gamma > 0$, converge as $\gamma \to 0$ to an interface $\Gamma(t)$ with unit normal n_Γ, evolving with normal velocity

$$\mathcal{V} = -\frac{\pi^2}{16} \Delta_s \kappa + \alpha \frac{\pi}{4} \Delta_s \overline{\phi} \quad \text{on } \Gamma(t), \tag{1.5a}$$

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

$$\Delta \overline{\phi} = 0 \quad \text{in } \Omega^+(t) := \Omega \setminus \overline{\Omega^-(t)}, \quad \frac{\partial \overline{\phi}}{\partial n_\Gamma} = 0 \quad \text{on } \Gamma(t), \tag{1.5b}$$

where $\Omega^-(t)$ is the void with boundary $\Gamma(t)$. For a discussion of different approaches to approximating (1.5a,b), see [9]. For further details on void electromigration see e.g. [38, 19] and the references therein.

   The present paper extends the phase field model (1.4) to take into account grain boundaries. In summary, the evolution of intergranular voids is described by the following non-linear 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 $(\eta_1(x, t), \eta_2(x, t)) \in \mathcal{K}$,

$$\gamma \frac{\partial u}{\partial t} - \nabla.(b(u)\nabla[w + \alpha\phi]) = 0 \quad \text{in } \Omega_T, \tag{1.6a}$$

$$\ell(\gamma)\frac{\partial v}{\partial t} + z = 0 \qquad\qquad\qquad\qquad \text{in } \Omega_T, \qquad\qquad (1.6\text{b})$$

$$(-\gamma\,\Delta u + \gamma^{-1}\psi_{,u}(u,v) - w)(\eta_1 - u)$$
$$+ (-\gamma\,\Delta v + \gamma^{-1}\psi_{,v}(u,v) - z)(\eta_2 - v) \geqslant 0 \qquad \text{in } \Omega_T, \qquad (1.6\text{c})$$

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

$$\frac{\partial u}{\partial \nu} = b(u)\frac{\partial[w + \alpha\phi]}{\partial \nu} = \frac{\partial v}{\partial \nu} = 0 \qquad\qquad \text{on } \partial\Omega \times (0, T], \qquad (1.6\text{d})$$

$$\nabla.(c(u)\nabla\phi) = 0 \qquad\qquad\qquad\qquad \text{in } \Omega_T, \qquad\qquad (1.6\text{e})$$

$$c(u)\frac{\partial\phi}{\partial \nu} = 0 \quad \text{on } \partial_1\Omega \times (0, T], \quad c(u)\frac{\partial\phi}{\partial \nu} + \phi = g^\pm := x_1 \pm 2 \quad \text{on } \partial_2^\pm\Omega \times (0, T], \quad (1.6\text{f})$$

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 $c(u)\partial\phi/\partial\nu = \pm 2$ 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 $\partial_2^\pm\Omega$, (1.6f), has the added advantage that one obtains an immediate $L^2(\partial_2\Omega)$ bound on ϕ for the degenerate elliptic equation (1.6e). In (1.6a–d), γ, $\ell(\gamma) \in \mathbb{R}_{>0}$ and $\alpha \in \mathbb{R}_{\geqslant 0}$ are given constants and

$$\Psi(r,s) := \begin{cases} \psi(r,s) & \text{if } (r,s) \in \mathcal{K}, \\ \infty & \text{if } (r,s) \notin \mathcal{K}, \end{cases} \quad \text{with } \psi \in C^2(\mathcal{K}), \qquad (1.7)$$

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

$$\psi(r,s) := \frac{8}{9} - \frac{1}{2}\left[\left(r - \frac{1}{3}\right)^2 + (1 - \mu)s^2 + \frac{2}{3}\mu(r + 1)\right], \qquad (1.8)$$

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

$$c(s) := 1 + s, \quad b(s) := 1 - s^2 = c(s)c(-s) \quad \forall s \in -[1, 1]. \qquad (1.9)$$

The variational inequality (1.6c) is obtained by taking the first variation of $\mathcal{E}$ with respect to u and v, and taking the non-smooth character of Ψ into account. The first variation can be computed in the context of subdifferentials. We choose a characterization of w and z in terms of variational inequalities, as it simplifies the formulation of the discrete problem. We refer to [20, 28] for further details on subdifferentials and variational inequalities.

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 (1.6a) by $\gamma^{-1}w = \gamma^{-1}\delta\mathcal{E}/\delta u$ and

(1.6b) by $[\ell(\gamma)]^{-1}z = [\ell(\gamma)]^{-1}\delta\mathcal{E}/\delta v$ yields, after integration of the sum of the two terms, the following free energy identity:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(u,v) + \int_\Omega [\gamma^{-1}b(u)|\nabla w|^2 + [\ell(\gamma)]^{-1}z^2]\,\mathrm{d}x = -\gamma^{-1}\alpha\int_\Omega b(u)\nabla w.\nabla\phi\,\mathrm{d}x. \qquad (1.10)$$

If we multiply (1.6e) 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 (1.10), leading to H^1-estimates in space for the phase field (u,v). Relating F to c and G to b by the identities

$$c(s)F''(s) = 1 \quad \text{and} \quad b(s)G''(s) = 1, \qquad (1.11)$$

and testing (1.6a) with $G'(u)$, (1.6b) with $-\Delta v$, and adding leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left[\gamma G(u) + \frac{l(\gamma)}{2}|\nabla v|^2\right]\mathrm{d}x + \gamma\int_\Omega [|\Delta u|^2 + |\Delta v|^2]\,\mathrm{d}x$$
$$= -\int_\Omega \nabla u.\nabla(\gamma^{-1}\Psi_{,u} + \alpha\phi)\,\mathrm{d}x - \gamma^{-1}\int_\Omega \nabla v.\nabla\Psi_{,v}\,\mathrm{d}x, \qquad (1.12)$$

where the term $\int_\Omega \nabla u.\nabla\phi\,\mathrm{d}x$ can be controlled if we test (1.6e) 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 order to derive a discrete analogue of the energy estimate (1.12), we need to extend a technique introduced in [39, 25] 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 in [14] can be found in [3]. However due to the lack of a corresponding entropy bound, convergence of that approximation was only established in one space dimension. Finally, we remark that no uniqueness results for degenerate fourth order parabolic equations are known, and hence no error estimates for this finite element approximation can be expected at present.

This paper is organised as follows. In Section 2 we formulate a finite element approximation of the degenerate system (P) and derive important discrete analogues of the energy estimates (1.10) and (1.12). In Section 3 we prove convergence, and hence existence of a weak solution to the system (P) in two space dimensions. In both of the above sections we need to substantially extend the techniques in [9], since e.g. the discrete analogue of (1.12) and the convergence of w on $\{|u| < 1\}$ are both far more difficult to establish; see Lemmas 2.4 and 3.2, respectively, below. This is due to the fact that the geometry of $\mathcal{K}$ leads to a far more complicated variational inequality compared to the one studied in [9]. In Section 4 we introduce and prove convergence of a "Gauss–Seidel type" iterative scheme for solving the non-linear 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 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 $\langle\cdot,\cdot\rangle$ denotes the standard L^2 inner product over Ω. In addition, we define $\underline{m}(\Omega)$ as the measure of Ω and $\fint \eta := [\underline{m}(\Omega)]^{-1}\langle\eta, 1\rangle$ 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

$$\{\eta \in L^2(0, T; X_1) : \partial\eta/\partial t \in L^2(0, T; X_3)\} \hookrightarrow L^2(0, T; X_2), \tag{1.13a}$$

$$\{\eta \in L^\infty(0, T; X_1) : \partial\eta/\partial t \in L^2(0, T; X_3)\} \hookrightarrow C([0, T]; X_2). \tag{1.13b}$$

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

$$\langle\nabla[\mathcal{G}\eta_1], \nabla\eta_2\rangle = \langle\eta_1, \eta_2\rangle_* \quad \forall\eta_2 \in H^1(\Omega), \tag{1.14}$$

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

$$|\eta|_0 \leqslant C(|\eta|_1 + |\langle\eta, 1\rangle|) \quad \forall\eta \in H^1(\Omega). \tag{1.15}$$

We note also for future reference Young's inequality

$$rs \leqslant \frac{\theta}{2}r^2 + \frac{1}{2\theta}s^2 \quad \forall r, s \in \mathbb{R}, \ \theta \in \mathbb{R}_{>0}. \tag{1.16}$$

Throughout C denotes a generic constant independent of h, τ and ε, the mesh and temporal discretization parameters and the regularization parameter. In addition $C(a_1, \ldots, a_I)$ 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 triangles σ with $h_\sigma := \mathrm{diam}(\sigma)$ and $h := \max_{\sigma \in \mathcal{T}^h} h_\sigma$, so that $\overline{\Omega} = \bigcup_{\sigma \in \mathcal{T}^h} \overline{\sigma}$. In addition, it is assumed that all triangles $\sigma \in \mathcal{T}^h$ are right-angled.

We note that the right-angle assumption is not a severe constraint, as there exist adaptive finite element codes that satisfy this requirement (see e.g. [35]).

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

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

We introduce also

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

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

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

We introduce $\pi^h : C(\overline{\Omega}) \to S^h$, the Lagrange 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

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

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

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

On recalling (1.9) and (1.11), we define functions F and G such that $c(\eta)\nabla[F'(\eta)] = \nabla\eta$ and $b(\eta)\nabla[G'(\eta)] = \nabla\eta$; that is,

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

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

$$F(s) = (1+s)\log\left(\frac{1+s}{2}\right) + (1-s) \quad \text{and} \quad G(s) = \tfrac{1}{2}[F(s) + F(-s)]; \tag{2.5}$$

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

$$F_\varepsilon(s) := \begin{cases} F(\varepsilon-1) + (s-\varepsilon+1)F'(\varepsilon-1) + \tfrac{1}{2}(s-\varepsilon+1)^2 F''(\varepsilon-1), & s \leqslant \varepsilon-1, \\ F(s) & s \geqslant \varepsilon-1, \end{cases}$$

$$G_\varepsilon(s), := \tfrac{1}{2}[F_\varepsilon(s) + F_\varepsilon(-s)]. \tag{2.6}$$

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

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

Similarly to the approach in [39, 25], 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,

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

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

The introduction of the matrices Λ_ε and Ξ_ε (see below) is crucial for the construction of our finite element approximation. Using an appropriate "standard" approach to approximate $c(u)$ and $b(u)$ would yield a finite element approximation satisfying a discrete analogue of (1.10), but not of (1.12). Without (1.12), one does not obtain the spatial continuity of the limit function u in two space dimensions and hence it is not possible to prove convergence of the resulting finite element approximation.

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 jth 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 triangle 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, 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, $k = 1 \to 2$,

$$
[\widehat{\Lambda}_\varepsilon(\widehat{\eta}^h)|_{\widehat{\sigma}}]_{kk}
$$
$$
:= \begin{cases} \dfrac{\widehat{\eta}^h(\widehat{p}_k) - \widehat{\eta}^h(\widehat{p}_0)}{F_\varepsilon'(\widehat{\eta}^h(\widehat{p}_k)) - F_\varepsilon'(\widehat{\eta}^h(\widehat{p}_0))} \equiv \dfrac{\eta^h(p_{j_k}) - \eta^h(p_{j_0})}{F_\varepsilon'(\eta^h(p_{j_k})) - F_\varepsilon'(\eta^h(p_{j_0}))} & \text{if } \eta^h(p_{j_k}) \neq \eta^h(p_{j_0}), \\[4mm] \dfrac{1}{F_\varepsilon''(\widehat{\eta}^h(\widehat{p}_0))} \equiv \dfrac{1}{F_\varepsilon''(\eta^h(p_{j_0}))} & \text{if } \eta^h(p_{j_k}) = \eta^h(p_{j_0}). \end{cases} \tag{2.10}
$$

As $R_\sigma^T \equiv R_\sigma^{-1}$, $\nabla \eta^h \equiv R_\sigma \widehat{\nabla} \widehat{\eta}^h$, where $x \equiv (x_1, x_2)^T$, $\nabla \equiv (\partial/\partial x_1, \partial/\partial x_2)^T$, $\widehat{x} \equiv (\widehat{x}_1, \widehat{x}_2)^T$ and $\widehat{\nabla} \equiv (\partial/\partial \widehat{x}_1, \partial/\partial \widehat{x}_2)^T$, 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 $\mathcal{T}^h$. Without this constraint, we could loose the property (2.8a) and hence the positivity of the $|\nabla w|^2$ term in the discrete analogue of (1.10). Another consequence of this constraint is that

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

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 $\eta^h \in S^h$,

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

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

In addition to $\mathcal{T}^h$, let $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ 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):

$(P_\varepsilon^{h,\tau})$ For $n \geqslant 1$ 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

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

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

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

$$\gamma\langle \nabla U_\varepsilon^n, \nabla[\eta_1^h - U_\varepsilon^n] \rangle + \gamma\langle \nabla V_\varepsilon^n, \nabla[\eta_2^h - V_\varepsilon^n] \rangle$$
$$\geqslant \langle W_\varepsilon^n - \gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \eta_1^h - U_\varepsilon^n \rangle^h$$
$$+ \langle Z_\varepsilon^n - \gamma^{-1}\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \eta_2^h - V_\varepsilon^n \rangle^h \quad \forall(\eta_1^h, \eta_2^h) \in K^h, \tag{2.14d}$$

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 for V_ε^0.

Below we recall some well-known results (see e.g. [17]) 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 $q \in (2, \infty]$:

$$|(I - \pi^h)\eta|_m \leqslant Ch^{2-m}|\eta|_2 \quad \forall \eta \in H^2(\Omega); \tag{2.15}$$

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

$$\int_\sigma \chi^2 \, dx \leqslant \int_\sigma \pi^h[\chi^2] \, dx \leqslant 4\int_\sigma \chi^2 \, dx; \tag{2.17}$$

$$\left| \int_\sigma (I - \pi^h)(\chi\eta^h) \, dx \right| \leqslant |(I - \pi^h)(\chi\eta^h)|_{0,1,\sigma} \leqslant Ch^{1+m}|\chi|_{m,\sigma}|\eta^h|_{1,\sigma}. \tag{2.18}$$

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

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

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

$$\langle \Delta^h\eta^h, \chi \rangle^h = -\langle \nabla\eta^h, \nabla\chi \rangle \quad \forall \chi \in S^h. \tag{2.20}$$

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

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

Then the following two lemmas follow immediately from the construction of Λ_ε and Ξ_ε (see [9, 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, Ξ_ε : $S^h \to [L^\infty(\Omega)]^{2\times 2}$ satisfy for all $\eta^h \in K^h$, $\xi \in \mathbb{R}^2$ and for all $\sigma \in \mathcal{T}^h$,

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

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

$$\xi^T \Xi_\varepsilon(\eta^h)|_\sigma \xi \leqslant 2 \xi^T \Lambda_\varepsilon(\eta^h)|_\sigma \xi. \quad (2.22\text{c})$$

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 $\sigma \in \mathcal{T}^h$,

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

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

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

We now derive discrete analogues of the energy estimates (1.10) and (1.12) in Lemmas 2.3 and 2.4 which follow.

LEMMA 2.3   Let the assumptions (A) hold and $(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}) \in K^h$. Then for all $\varepsilon \in (0, 1)$ and for all h, $\tau_n > 0$ there exists a solution $(\Phi_\varepsilon^n, U_\varepsilon^n, V_\varepsilon^n, W_\varepsilon^n, Z_\varepsilon^n)$ to the n-th step of $(P_\varepsilon^{h,\tau})$ with $\fint U_\varepsilon^n = \fint U_\varepsilon^{n-1}$. The 4-tuple $(\Phi_\varepsilon^n, U_\varepsilon^n, V_\varepsilon^n, Z_\varepsilon^n)$ is unique. In addition, W_ε^n is unique if there exists $j \in J$ such that $U_\varepsilon^n(p_j) \in (-1, 1)$. Moreover,

$$\langle \Lambda_\varepsilon(U_\varepsilon^{n-1}) \nabla \Phi_\varepsilon^n, \nabla \Phi_\varepsilon^n \rangle + \tfrac{1}{2} |\Phi_\varepsilon^n|_{0,\partial_2\Omega}^2 \leqslant \tfrac{1}{2} |g|_{0,\partial_2\Omega}^2, \quad (2.24)$$

$$|\langle \nabla \Phi_\varepsilon^n, \nabla U_\varepsilon^{n-1} \rangle| \leqslant 2 |g|_{0,\partial_2\Omega} |\pi^h[F_\varepsilon'(U_\varepsilon^{n-1})]|_{0,\partial_2\Omega} \quad (2.25)$$

and

$$\mathcal{E}(U_\varepsilon^n, V_\varepsilon^n) + \tfrac{1}{2}[\gamma |U_\varepsilon^n - U_\varepsilon^{n-1}|_1^2 + \gamma |V_\varepsilon^n - V_\varepsilon^{n-1}|_1^2] + [\ell(\gamma)]^{-1} \tau_n |Z_\varepsilon^n|_h^2$$
$$+ \tfrac{1}{2} \gamma^{-1} \tau_n |[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{1/2} \nabla W_\varepsilon^n|_0^2 \leqslant \mathcal{E}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}) + \tfrac{1}{2} \alpha^2 \gamma^{-1} \tau_n |g|_{0,\partial_2\Omega}^2, \quad (2.26\text{a})$$

where

$$\mathcal{E}(U_\varepsilon^n, V_\varepsilon^n) := \tfrac{1}{2}[\gamma |U_\varepsilon^n|_1^2 + \gamma |V_\varepsilon^n|_1^2] + \gamma^{-1} \langle \psi(U_\varepsilon^n, V_\varepsilon^n), 1 \rangle^h. \quad (2.26\text{b})$$

Proof. The derivation of the existence of a unique solution $\Phi_\varepsilon^n \in S^h$ to (2.14a), and the bounds (2.24) and (2.25), is straightforward, and can be found in [9, Lemma 2.4]. In order to prove existence of a solution $((U_\varepsilon^n, V_\varepsilon^n), W_\varepsilon^n, Z_\varepsilon^n) \in K^h \times [S^h]^2$ to (2.14b–d), we introduce, similarly to (1.14), for $q^h \in K^h$ the discrete anisotropic Green's operator $\mathcal{G}_{q^h}^h : Y_2^h \to Y_2^h$ such that

$$\langle \Xi_\varepsilon(q^h) \nabla [\mathcal{G}_{q^h}^h \eta^h], \nabla \chi \rangle = \langle \eta^h, \chi \rangle^h \quad \forall \chi \in S^h. \quad (2.27)$$

It follows immediately from (2.22b) and (1.15) that $\mathcal{G}_{q^h}^h$ is well-posed. It follows from (2.14b) and (2.27) that

$$W_\varepsilon^n \equiv -\alpha \Phi_\varepsilon^n - \gamma \mathcal{G}_{U_\varepsilon^{n-1}}^h \left[\frac{U_\varepsilon^n - U_\varepsilon^{n-1}}{\tau_n} \right] + \lambda^n, \quad (2.28)$$

where $\lambda^n \in \mathbb{R}$. Hence (2.14b–d) can be restated as follows: Find $(U_\varepsilon^n, V_\varepsilon^n) \in K^h(U_\varepsilon^{n-1}) :=$ $\{(\eta_1^h, \eta_2^h) \in K^h : \eta_1^h - U_\varepsilon^{n-1} \in Y_2^h\}$ and a Lagrange multiplier $\lambda^n \in \mathbb{R}$ such that for all $(\eta_1^h, \eta_2^h) \in K^h$,

$$\gamma \langle \nabla U_\varepsilon^n, \nabla(\eta_1^h - U_\varepsilon^n) \rangle + \gamma \langle \nabla V_\varepsilon^n, \nabla(\eta_2^h - V_\varepsilon^n) \rangle$$
$$+ \gamma \left\langle \mathcal{G}_{U_\varepsilon^{n-1}}^h \left[\frac{U_\varepsilon^n - U_\varepsilon^{n-1}}{\tau_n} \right], \eta_1^h - U_\varepsilon^n \right\rangle^h + \ell(\gamma) \left\langle \frac{V_\varepsilon^n - V_\varepsilon^{n-1}}{\tau_n}, \eta_2^h - V_\varepsilon^n \right\rangle^h$$
$$\geqslant \langle -\gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}) - \alpha\Phi_\varepsilon^n + \lambda^n, \eta_1^h - U_\varepsilon^n \rangle^h - \gamma^{-1} \langle \psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \eta_2^h - V_\varepsilon^n \rangle^h.$$
(2.29)

It follows from (2.29) that $(U_\varepsilon^n, V_\varepsilon^n) \in K^h(U_\varepsilon^{n-1})$ is such that for all $(\eta_1^h, \eta_2^h) \in K^h(U_\varepsilon^{n-1})$,

$$\gamma \langle \nabla U_\varepsilon^n, \nabla(\eta_1^h - U_\varepsilon^n) \rangle + \gamma \langle \nabla V_\varepsilon^n, \nabla(\eta_2^h - V_\varepsilon^n) \rangle$$
$$+ \gamma \left\langle \mathcal{G}_{U_\varepsilon^{n-1}}^h \left[\frac{U_\varepsilon^n - U_\varepsilon^{n-1}}{\tau_n} \right], \eta_1^h - U_\varepsilon^n \right\rangle^h + \ell(\gamma) \left\langle \frac{V_\varepsilon^n - V_\varepsilon^{n-1}}{\tau_n}, \eta_2^h - V_\varepsilon^n \right\rangle^h$$
$$\geqslant \langle -\gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}) - \alpha\Phi_\varepsilon^n, \eta_1^h - U_\varepsilon^n \rangle^h - \gamma^{-1} \langle \psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \eta_2^h - V_\varepsilon^n \rangle^h.$$
(2.30)

There exists a unique $(U_\varepsilon^n, V_\varepsilon^n) \in K^h(U_\varepsilon^{n-1})$ solving (2.30) since, on noting (2.27), this is the Euler–Lagrange variational inequality of the strictly convex minimization problem

$$\min_{(\eta_1^h, \eta_2^h) \in K^h(U_\varepsilon^{n-1})} \left\{ \frac{\gamma}{2}|\eta_1^h|_1^2 + \frac{\gamma}{2}|\eta_2^h|_1^2 + \frac{\gamma}{2\tau_n}|[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{1/2}\nabla\mathcal{G}_{U_\varepsilon^{n-1}}^h(\eta_1^h - U_\varepsilon^{n-1})|_0^2 \right.$$
$$\left. + \frac{\ell(\gamma)}{2\tau_n}|\eta_2^h - V_\varepsilon^{n-1}|_0^2 + \langle \gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}) + \alpha\Phi_\varepsilon^n, \eta_1^h \rangle^h + \langle \gamma^{-1}\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \eta_2^h \rangle^h \right\}.$$

Existence of the Lagrange multiplier λ^n in (2.29) then follows from standard optimisation theory (see e.g. [18]). Hence we have existence of a solution $((U_\varepsilon^n, V_\varepsilon^n), W_\varepsilon^n, Z_\varepsilon^n) \in K^h \times [S^h]^2$ to (2.14b–d). If $|U_\varepsilon^n(p_j)| < 1$ for some $j \in J$ then $\pi^h[1 - (U_\varepsilon^n)^2] \not\equiv 0$ and choosing $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n \pm \delta\pi^h[1 - (U_\varepsilon^n)^2], V_\varepsilon^n \pm (\delta/\sqrt{3})\pi^h[1 - (U_\varepsilon^n)^2])$ in (2.29) for $\delta > 0$ sufficiently small yields uniqueness of λ^n and, on noting (2.28), uniqueness of W_ε. Furthermore, choosing $\chi \equiv 1$ in (2.14b) yields $\fint U_\varepsilon^n = \fint U_\varepsilon^{n-1}$.

Choosing $\chi \equiv W_\varepsilon^n$ in (2.14b), $\chi \equiv Z_\varepsilon^n$ in (2.14c) and $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^{n-1}, V_\varepsilon^{n-1})$ in (2.14d) yields

$$\gamma \langle U_\varepsilon^n - U_\varepsilon^{n-1}, W_\varepsilon^n \rangle^h + \tau_n \langle \Xi_\varepsilon(U_\varepsilon^{n-1})\nabla[W_\varepsilon^n + \alpha\Phi_\varepsilon^n], \nabla W_\varepsilon^n \rangle = 0, \tag{2.31a}$$

$$\ell(\gamma)\langle V_\varepsilon^n - V_\varepsilon^{n-1}, Z_\varepsilon^n \rangle^h + \tau_n \langle Z_\varepsilon^n, Z_\varepsilon^n \rangle^h = 0, \tag{2.31b}$$

$$\gamma \langle \nabla U_\varepsilon^n, \nabla[U_\varepsilon^{n-1} - U_\varepsilon^n] \rangle + \gamma \langle \nabla V_\varepsilon^n, \nabla[V_\varepsilon^{n-1} - V_\varepsilon^n] \rangle$$
$$\geqslant \langle W_\varepsilon^n - \gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), U_\varepsilon^{n-1} - U_\varepsilon^n \rangle^h$$
$$+ \langle Z_\varepsilon^n - \gamma^{-1}\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), V_\varepsilon^{n-1} - V_\varepsilon^n \rangle^h. \tag{2.31c}$$

On noting the elementary identity

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

it follows from (2.31a–c), (1.16), (2.22c) and the convexity of $-\psi$ (recall (1.7)) that

$$
\begin{aligned}
\tfrac{1}{2}\gamma[|U_\varepsilon^n|_1^2 &+ |U_\varepsilon^n - U_\varepsilon^{n-1}|_1^2 - |U_\varepsilon^{n-1}|_1^2 + |V_\varepsilon^n|_1^2 + |V_\varepsilon^n - V_\varepsilon^{n-1}|_1^2 - |V_\varepsilon^{n-1}|_1^2] \\
&+ \gamma^{-1}\tau_n|[\varXi_\varepsilon(U_\varepsilon^{n-1})]^{1/2}\nabla W_\varepsilon^n|_0^2 + [\ell(\gamma)]^{-1}\tau_n|Z_\varepsilon^n|_h^2 \\
&\leqslant -\gamma^{-1}\langle\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), U_\varepsilon^n - U_\varepsilon^{n-1}\rangle^h - \gamma^{-1}\langle\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), V_\varepsilon^n - V_\varepsilon^{n-1}\rangle^h \\
&\quad - \alpha\gamma^{-1}\tau_n\langle\varXi_\varepsilon(U_\varepsilon^{n-1})\nabla\varPhi_\varepsilon^n, \nabla W_\varepsilon^n\rangle \\
&\leqslant \gamma^{-1}\langle\psi(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}) - \psi(U_\varepsilon^n, V_\varepsilon^n), 1\rangle^h \\
&\quad + \tfrac{1}{2}\gamma^{-1}\tau_n[|[\varXi_\varepsilon(U_\varepsilon^{n-1})]^{1/2}|\nabla W_\varepsilon^n|_0^2 + 2\alpha^2|[\varLambda_\varepsilon(U_\varepsilon^{n-1})]^{1/2}\nabla\varPhi_\varepsilon^n|_0^2].
\end{aligned}
\tag{2.32}
$$

Hence the desired result (2.26a) follows from (2.32), (2.26b) and (2.24). $\qquad\square$

LEMMA 2.4 Let the assumptions of Lemma 2.3 hold. Then

$$
\begin{aligned}
\gamma\langle G_\varepsilon(U_\varepsilon^n) &- G_\varepsilon(U_\varepsilon^{n-1}), 1\rangle^h + \gamma\tau_n|\Delta^h U_\varepsilon^n|_h^2 \\
&+ \tfrac{1}{2}\ell(\gamma)[|V_\varepsilon^n|_1^2 + |V_\varepsilon^n - V_\varepsilon^{n-1}|_1^2 - |V_\varepsilon^{n-1}|_1^2] + \gamma\tau_n|\Delta^h V_\varepsilon^n|_h^2 \\
&\leqslant \varepsilon^{-1}\gamma|U_\varepsilon^n - U_\varepsilon^{n-1}|_h^2 + \tau_n[\langle\nabla W_\varepsilon^n, \nabla[U_\varepsilon^n - U_\varepsilon^{n-1}]\rangle - \alpha\langle\nabla\varPhi_\varepsilon^n, \nabla U_\varepsilon^n\rangle \\
&\quad - \gamma^{-1}\langle\nabla\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \nabla U_\varepsilon^n\rangle - \gamma^{-1}\langle\nabla\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \nabla V_\varepsilon^n\rangle].
\end{aligned}
\tag{2.33}
$$

Proof. Choosing $\chi \equiv \pi^h[G_\varepsilon'(U_\varepsilon^{n-1})]$ in (2.14b) and noting (2.12b) yields

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

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

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

We now extend an argument in [5, 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 $v_b := 2/\sqrt{3}$.

  (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) $U_\varepsilon^n(p_j) \in (-1, 1), \ V_\varepsilon^n(p_j) = \tfrac{1}{2}v_b(U_\varepsilon^n(p_j) + 1)$,
 (iv) $U_\varepsilon^n(p_j) \in (-1, 1), \ V_\varepsilon^n(p_j) = -\tfrac{1}{2}v_b(U_\varepsilon^n(p_j) + 1)$,
  (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 $(\eta_1^h, \eta_2^h) \in \mathcal{K}$ can be chosen in (2.14d). In case (i), on choosing $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n \pm \delta\chi_j, V_\varepsilon^n)$ and $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n, V_\varepsilon^n \pm \delta\chi_j)$, respectively, we have

$$
A_j^U := \gamma\langle\nabla U_\varepsilon^n, \nabla\chi_j\rangle - \langle W_\varepsilon^n - \gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \chi_j\rangle^h = 0,
\tag{2.36a}
$$

$$
A_j^V := \gamma\langle\nabla V_\varepsilon^n, \nabla\chi_j\rangle - \langle Z_\varepsilon^n - \gamma^{-1}\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \chi_j\rangle^h = 0.
\tag{2.36b}
$$

In case (ii), on choosing $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n, V_\varepsilon^n \pm \delta\chi_j)$ and $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n - \delta\chi_j, V_\varepsilon^n)$, respectively, we have $A_j^U \leqslant 0$ and $A_j^V = 0$. For case (iii) we choose $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n \pm \delta\chi_j, V_\varepsilon^n \pm \frac{1}{2}v_b\delta\chi_j)$ and $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n, V_\varepsilon^n - \delta\chi_j)$, respectively, so that $A_j^U + \frac{1}{2}v_b A_j^V = 0$ and $A_j^V \leqslant 0$. Similarly, for case (iv), on choosing $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n \pm \delta\chi_j, V_\varepsilon^n \mp \frac{1}{2}v_b\delta\chi_j)$ and $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n, V_\varepsilon^n + \delta\chi_j)$, respectively, we obtain $A_j^U - \frac{1}{2}v_b A_j^V = 0$ and $A_j^V \geqslant 0$. In case (v) we obtain $A_j^V \leqslant 0$ and $A_j^U + \frac{1}{2}v_b A_j^V \leqslant 0$, on choosing $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n, V_\varepsilon^n - \delta\chi_j)$ and $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n - \delta\chi_j, V_\varepsilon^n - \frac{1}{2}v_b\delta\chi_j)$, respectively. Similarly, for case (vi) we have $A_j^V \geqslant 0$ and $A_j^U - \frac{1}{2}v_b A_j^V \leqslant 0$, on choosing $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n, V_\varepsilon^n + \delta\chi_j)$ and $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n - \delta\chi_j, V_\varepsilon^n + \frac{1}{2}v_b\delta\chi_j)$, respectively. Finally, in case (vii) we that $A_j^U \pm \frac{1}{2}v_b A_j^V \geqslant 0$, on choosing $(\eta_1^h, \eta_2^h) \equiv (U_\varepsilon^n + \delta\chi_j, V_\varepsilon^n \pm \frac{1}{2}v_b\delta\chi_j)$.

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

$$U_\varepsilon^n(p_j) = 1 \implies U_\varepsilon^n(p_j) \geqslant U_\varepsilon^n(p_i) \; \forall i \in J \implies \Delta^h U_\varepsilon^n(p_j) \leqslant 0. \tag{2.37a}$$

Similarly, in cases (iii), (v) and (vii),

$$\Delta^h(U_\varepsilon^n - \sqrt{3}V_\varepsilon)(p_j) \geqslant 0, \tag{2.37b}$$

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

$$\Delta^h(U_\varepsilon^n + \sqrt{3}V_\varepsilon)(p_j) \geqslant 0. \tag{2.37c}$$

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

$$-[A_j^U \Delta^h U_\varepsilon^n(p_j) + A_j^V \Delta^h V_\varepsilon^n(p_j)] \leqslant 0. \tag{2.38}$$

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

$$\gamma|\Delta^h U_\varepsilon^n|_h^2 + \gamma|\Delta^h V_\varepsilon^n|_h^2$$
$$\leqslant -\langle W_\varepsilon^n - \gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \Delta^h U_\varepsilon^n\rangle^h - \langle Z_\varepsilon^n - \gamma^{-1}\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \Delta^h V_\varepsilon^n\rangle^h$$
$$= \langle \nabla[W_\varepsilon^n - \gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1})], \nabla U_\varepsilon^n\rangle + \langle \nabla[Z_\varepsilon^n - \gamma^{-1}\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1})], \nabla V_\varepsilon^n\rangle. \tag{2.39}$$

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

$$\gamma\langle G_\varepsilon(U_\varepsilon^n) - G_\varepsilon(U_\varepsilon^{n-1}), 1\rangle^h + \gamma\tau_n|\Delta^h U_\varepsilon^n|_h^2$$
$$+ \tfrac{1}{2}\ell(\gamma)[|V_\varepsilon^n|_1^2 + |V_\varepsilon^n - V_\varepsilon^{n-1}|_1^2 - |V_\varepsilon^{n-1}|_1^2] + \gamma\tau_n|\Delta^h V_\varepsilon^n|_h^2$$
$$\leqslant \gamma\langle U_\varepsilon^n - U_\varepsilon^{n-1}, G_\varepsilon'(U_\varepsilon^n)\rangle^h + \tau_n[\langle \nabla[W_\varepsilon^n - \gamma^{-1}\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1})], \nabla U_\varepsilon^n\rangle$$
$$- \gamma^{-1}\langle \nabla\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \nabla V_\varepsilon^n\rangle]$$
$$\leqslant \gamma\langle U_\varepsilon^n - U_\varepsilon^{n-1}, G_\varepsilon'(U_\varepsilon^n) - G_\varepsilon'(U_\varepsilon^{n-1})\rangle^h$$
$$+ \tau_n[\langle \nabla W_\varepsilon^n, \nabla[U_\varepsilon^n - U_\varepsilon^{n-1}]\rangle - \alpha\langle \nabla\Phi_\varepsilon^n, \nabla U_\varepsilon^{n-1}\rangle$$
$$- \gamma^{-1}\langle \nabla\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \nabla U_\varepsilon^n\rangle - \gamma^{-1}\langle \nabla\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \nabla V_\varepsilon^n\rangle]$$
$$\leqslant \varepsilon^{-1}\gamma|U_\varepsilon^n - U_\varepsilon^{n-1}|_h^2 + \tau_n[\langle \nabla W_\varepsilon^n, \nabla[U_\varepsilon^n - U_\varepsilon^{n-1}]\rangle - \alpha\langle \nabla\Phi_\varepsilon^n, \nabla U_\varepsilon^{n-1}\rangle$$
$$- \gamma^{-1}\langle \nabla\psi_{,u}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \nabla U_\varepsilon^n\rangle - \gamma^{-1}\langle \nabla\psi_{,v}(U_\varepsilon^{n-1}, V_\varepsilon^{n-1}), \nabla V_\varepsilon^n\rangle]$$

and hence the desired result (2.33). $\qquad\square$

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

THEOREM 2.1   Let the assumptions (A) hold and $(U_\varepsilon^0, V_\varepsilon^0) \in K^h$. Then for all $\varepsilon \in (0, 1)$, $h > 0$ and for all time partitions $\{\tau_n\}_{n=1}^N$, the solution $\{(\Phi_\varepsilon^n, U_\varepsilon^n, V_\varepsilon^n, W_\varepsilon^n, Z_\varepsilon^n)\}_{n=1}^N$ to $(P_\varepsilon^{h,\tau})$ is such that $\fint U_\varepsilon^n = \fint U_\varepsilon^0$, $n = 1 \to N$, and

$$
\gamma \max_{n=1 \to N} \|U_\varepsilon^n\|_1^2 + \gamma \max_{n=1 \to N} \|V_\varepsilon^n\|_1^2 + \gamma \sum_{n=1}^N [|U_\varepsilon^n - U_\varepsilon^{n-1}|_1^2 + |V_\varepsilon^n - V_\varepsilon^{n-1}|_1^2]
$$
$$
+ \sum_{n=1}^N \tau_n \left[\gamma^{-1}|[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{1/2} \nabla W_\varepsilon^n|_0^2 + [\ell(\gamma)]^{-1}|Z_\varepsilon^n|_h^2 + \ell(\gamma) \left| \frac{V_\varepsilon^n - V_\varepsilon^{n-1}}{\tau_n} \right|_h^2 \right]
$$
$$
\leqslant C[\gamma \|U_\varepsilon^0\|_1^2 + \gamma \|V_\varepsilon^0\|_1^2 + \gamma^{-1}(1 + T|g|_{0,\partial_2\Omega}^2)]. \quad (2.40)
$$

In addition

$$
\gamma \sum_{n=1}^N \tau_n \left| \mathcal{G}\left[\frac{U_\varepsilon^n - U_\varepsilon^{n-1}}{\tau_n} \right] \right|_1^2 + \gamma \tau^{-1/2} \sum_{n=1}^N |U_\varepsilon^n - U_\varepsilon^{n-1}|_0^2
$$
$$
\leqslant C[\gamma \|U_\varepsilon^0\|_1^2 + \gamma \|V_\varepsilon^0\|_1^2 + \gamma^{-1}(1 + T|g|_{0,\partial_2\Omega}^2)] \quad (2.41)
$$

and

$$
\gamma \max_{n=1 \to N} \langle G_\varepsilon(U_\varepsilon^n), 1 \rangle^h + \gamma \sum_{n=1}^N \tau_n |\Delta^h U_\varepsilon^n|_h^2 + \gamma \sum_{n=1}^N \tau_n |\Delta^h V_\varepsilon^n|_h^2
$$
$$
\leqslant \gamma \langle G_\varepsilon(U_\varepsilon^0), 1 \rangle^h + \alpha^2 \sum_{n=1}^N \tau_n |\pi^h[F_\varepsilon'(U_\varepsilon^{n-1})]|_{0,\partial_2\Omega}^2
$$
$$
+ C(T)[1 + \gamma^{-2} + \varepsilon^{-1}\tau^{1/2}][\gamma \|U_\varepsilon^0\|_1^2 + \gamma \|V_\varepsilon^0\|_1^2 + \gamma^{-1}(1 + T|g|_{0,\partial_2\Omega}^2)]. \quad (2.42)
$$

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

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

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 (1.14), (2.3), (2.14b), (2.22b,c), (2.19), (2.24) and (2.40); see [9, Theorem 2.6] for details.

Finally, summing (2.33) from $n = 1$ to k and noting (1.7), (2.2), (2.17) and (2.22b) implies for any $k \leqslant N$ that

$$\gamma \langle G_\varepsilon(U_\varepsilon^k), 1\rangle^h + \gamma \sum_{n=1}^{k} \tau_n |\Delta^h U_\varepsilon^n|_h^2 + \gamma \sum_{n=1}^{k} \tau_n |\Delta^h V_\varepsilon^n|_h^2$$

$$\leqslant \gamma \langle G_\varepsilon(U_\varepsilon^0), 1\rangle^h + \sum_{n=1}^{k} [\varepsilon^{-1}\gamma |U_\varepsilon^n - U_\varepsilon^{n-1}|_0^2 + \tau_n |\alpha\langle \nabla\Phi_\varepsilon^n, \nabla U_\varepsilon^{n-1}\rangle|]$$

$$+ \gamma^{-1} t_k |\psi|_{2,\infty,\mathcal{K}} [\max_{n=0 \to k} \|U_\varepsilon^n\|_1^2 + \max_{n=0 \to k} \|V_\varepsilon^n\|_1^2]$$

$$+ \Big[\varepsilon^{-1}\sum_{n=1}^{k} \tau_n |[\Xi_\varepsilon(U_\varepsilon^{n-1})]^{1/2}\nabla W_\varepsilon^n|_0^2\Big]^{1/2} \Big[\sum_{n=1}^{k} \tau_n |U_\varepsilon^n - U_\varepsilon^{n-1}|_1^2\Big]^{1/2}. \qquad (2.44)$$

The desired result (2.42) then follows from (2.44), (2.25), (1.16), (2.40) and (2.41). $\qquad \square$

LEMMA 2.5 Let $(u^0, v^0) \in K \cap [W^{1,p}(\Omega)]^2$ with $p > 2$, and let 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 $h > 0$,

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

Proof. This follows from (2.16), (2.6) and (2.5). $\qquad \square$

REMARK 2.1 The approximation $(P_\varepsilon^{h,\tau})$ 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 [9] and introduce the following approximation of (P). Adopting the notation (2.9) and (2.10), let $\widetilde{\Lambda}_\varepsilon, \widetilde{\Xi}_\varepsilon : S^h \to [L^\infty(\Omega)]^{2\times 2}$ be 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

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

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

We note that the key identities, with $\Lambda_\varepsilon(\eta^h)$ 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 $(\widetilde{P}_\varepsilon^{h,\tau})$ of (P), which is the same as $(P_\varepsilon^{h,\tau})$ but with $\Lambda_\varepsilon(U_\varepsilon^{n-1})$ in (2.14a) replaced by $\widetilde{\Lambda}_\varepsilon(U_\varepsilon^{n-1})$ and $\Xi_\varepsilon(U_\varepsilon^{n-1})$ in (2.14b) replaced by $\widetilde{\Xi}_\varepsilon(U_\varepsilon^{n-1})$. As $\widetilde{\Lambda}_\varepsilon(\cdot)$ and $\widetilde{\Xi}_\varepsilon(\cdot)$ are now degenerate, existence of a solution $(\Phi_\varepsilon^n, U_\varepsilon^n, V_\varepsilon^n, W_\varepsilon^n, Z_\varepsilon^n)$ to $(\widetilde{P}_\varepsilon^{h,\tau})$ does not appear to be trivial. However, this can easily be established by splitting the nodes into passive and active sets (see e.g. [4]). Moreover, one can show that $(U_\varepsilon^n, V_\varepsilon^n, Z_\varepsilon^n)$ 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 $(\widetilde{P}_\varepsilon^{h,\tau})$.

3. Convergence

Let

$$U_\varepsilon(t) := \frac{t - t_{n-1}}{\tau_n} U_\varepsilon^n + \frac{t_n - t}{\tau_n} U_\varepsilon^{n-1}, \quad t \in [t_{n-1}, t_n], \ n \geqslant 1, \qquad (3.1a)$$

$$U_\varepsilon^+(t) := U_\varepsilon^n, \quad U_\varepsilon^-(t) := U_\varepsilon^{n-1}, \quad t \in (t_{n-1}, t_n], \ n \geqslant 1. \qquad (3.1b)$$

We note for future reference that

$$U_\varepsilon - U_\varepsilon^\pm = (t - t_n^\pm)\frac{\partial U_\varepsilon}{\partial t}, \quad t \in (t_{n-1}, t_n),\ n \geqslant 1, \tag{3.2}$$

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

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

Using the above notation, and introducing analogous notation for V_ε, W_ε^+, Z_ε^+ and Φ_ε^+, we can restate $(\mathrm{P}_\varepsilon^{h,\tau})$ 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 $\chi \in L^\infty(0, T; S^h)$,

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

$$\int_0^T \left[\gamma\left\langle \frac{\partial U_\varepsilon}{\partial t}, \chi \right\rangle^h + \langle \Xi_\varepsilon(U_\varepsilon^-)\nabla[W_\varepsilon^+ + \alpha\Phi_\varepsilon^+], \nabla\chi \rangle \right] \mathrm{d}t = 0, \tag{3.4b}$$

$$\int_0^T \left[\ell(\gamma)\left\langle \frac{\partial V_\varepsilon}{\partial t}, \chi \right\rangle^h + \langle Z_\varepsilon^+, \chi \rangle^h \right] \mathrm{d}t = 0, \tag{3.4c}$$

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

$$\gamma[\langle \nabla U_\varepsilon^+, \nabla[\eta_1^h - U_\varepsilon^+] \rangle + \langle \nabla V_\varepsilon^+, \nabla[\eta_2^h - V_\varepsilon^+] \rangle]$$
$$\geqslant [\langle W_\varepsilon^+ - \gamma^{-1}\psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \eta_1^h - U_\varepsilon^+ \rangle^h + \langle Z_\varepsilon^+ - \gamma^{-1}\psi_{,v}(U_\varepsilon^-, V_\varepsilon^-), \eta_2^h - V_\varepsilon^+ \rangle^h]$$
$$\forall (\eta_1^h, \eta_2^h) \in K^h. \tag{3.4d}$$

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 assumptions (A), $\varepsilon \in (0, 1)$ with $\varepsilon \to 0$ as $h \to 0$, and $\tau_n \leqslant C\tau_{n-1} \leqslant C\varepsilon^2$, $n = 2 \to N$.

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

$$u \in L^\infty(0, T; H^1(\Omega)) \cap H^1(0, T; (H^1(\Omega))'), \tag{3.5a}$$

$$v \in L^\infty(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega)), \quad z \in L^2(\Omega_T), \tag{3.5b}$$

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

$$U_\varepsilon, U_\varepsilon^\pm \to u \quad \text{and} \quad V_\varepsilon, V_\varepsilon^\pm \to v \qquad \text{weak-}* \text{ in } L^\infty(0, T; H^1(\Omega)), \tag{3.6a}$$

$$\mathcal{G}\frac{\partial U_\varepsilon}{\partial t} \to \mathcal{G}\frac{\partial u}{\partial t} \qquad\qquad\qquad \text{weakly in } L^2(0, T; H^1(\Omega)), \tag{3.6b}$$

$$\frac{\partial V_\varepsilon}{\partial t} \to \frac{\partial v}{\partial t} \quad \text{and} \quad Z_\varepsilon^+ \to z \qquad \text{weakly in } L^2(\Omega_T), \tag{3.6c}$$

$$U_\varepsilon, U_\varepsilon^\pm \to u \quad \text{and} \quad V_\varepsilon,\ V_\varepsilon^\pm \to v \qquad \text{strongly in } L^2(0, T; L^s(\Omega)), \tag{3.7a}$$

$$\Xi_\varepsilon(U_\varepsilon^-) \to b(u)\mathcal{I} \quad \text{and} \quad \Lambda_\varepsilon(U_\varepsilon^-) \to c(u)\mathcal{I} \qquad \text{strongly in } L^2(0, T; L^s(\Omega)), \tag{3.7b}$$

for all $s \in [2, \infty)$. If in addition $u^0 \in H^2(\Omega)$ with $\partial u^0 / \partial \nu = 0$ on $\partial \Omega$ and

$$\alpha^2 \int_0^T |\pi^h[F'_\varepsilon(U^-_\varepsilon)]|^2_{0, \partial_2 \Omega} \, dt \leqslant C, \tag{3.8}$$

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 $h \to 0$,

$$\Delta^h U_\varepsilon, \; \Delta^h U^\pm_\varepsilon \to \Delta u \quad \text{weakly in } L^2(\Omega_T), \tag{3.10a}$$

$$U_\varepsilon, U^\pm_\varepsilon \to u \quad \text{weakly in } L^2(0, T; W^{1,s}(\Omega)), \text{ for any } s \in [2, \infty), \tag{3.10b}$$

$$U_\varepsilon, U^\pm_\varepsilon \to u \quad \text{strongly in } L^2(0, T; C^{0,\zeta}(\overline{\Omega})), \text{ for any } \zeta \in (0, 1), \tag{3.10c}$$

$$U_\varepsilon, U^\pm_\varepsilon \to u \quad \text{strongly in } L^2(0, T; H^1(\Omega)). \tag{3.10d}$$

Similarly, if in addition $v^0 \in H^2(\Omega)$ with $\partial v^0 / \partial \nu = 0$ 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 $h \to 0$,

$$\Delta^h V_\varepsilon, \; \Delta^h V^\pm_\varepsilon \to \Delta v \quad \text{weakly in } L^2(\Omega_T), \tag{3.12a}$$

$$V_\varepsilon, V^\pm_\varepsilon \to v \quad \text{weakly in } L^2(0, T; W^{1,s}(\Omega)), \text{ for any } s \in [2, \infty), \tag{3.12b}$$

$$V_\varepsilon, V^\pm_\varepsilon \to v \quad \text{strongly in } L^2(0, T; C^{0,\zeta}(\overline{\Omega})), \text{ for any } \zeta \in (0, 1), \tag{3.12c}$$

$$V_\varepsilon, V^\pm_\varepsilon \to v \quad \text{strongly in } L^2(0, T; H^1(\Omega)). \tag{3.12d}$$

Proof. Noting the definitions (3.1a,b), (3.3), the bounds in (2.24), (2.40), (2.41) and (2.42) together with (1.15), (2.45) and our assumption (i) imply that

$$\|[\Lambda_\varepsilon(U^-_\varepsilon)]^{1/2} \nabla \Phi^+_\varepsilon\|^2_{L^2(\Omega_T)} + \|\Phi^+_\varepsilon\|^2_{L^2(0,T;L^2(\partial_2\Omega))} + \|U^{(\pm)}_\varepsilon\|^2_{L^\infty(0,T;H^1(\Omega))} + \|V^{(\pm)}_\varepsilon\|^2_{L^\infty(0,T;H^1(\Omega))}$$

$$+ \left\|\bar{\tau}^{1/2} \frac{\partial U_\varepsilon}{\partial t}\right\|^2_{L^2(0,T;H^1(\Omega))} + \left\|\bar{\tau}^{1/2} \frac{\partial V_\varepsilon}{\partial t}\right\|^2_{L^2(0,T;H^1(\Omega))} + \|[\Xi_\varepsilon(U^-_\varepsilon)]^{1/2} \nabla W^+_\varepsilon\|^2_{L^2(\Omega_T)}$$

$$+ \left\|\mathcal{G} \frac{\partial U_\varepsilon}{\partial t}\right\|^2_{L^2(0,T;H^1(\Omega))} + \tau^{-1/2} \left\|\bar{\tau}^{1/2} \frac{\partial U_\varepsilon}{\partial t}\right\|^2_{L^2(\Omega_T)} + \left\|\frac{\partial V_\varepsilon}{\partial t}\right\|^2_{L^2(\Omega_T)} + \|Z^+_\varepsilon\|^2_{L^2(\Omega_T)} \leqslant C, \tag{3.13a}$$

and

$$\|\Delta^h U^+_\varepsilon\|^2_{L^2(\Omega_T)} + \|\Delta^h V^+_\varepsilon\|^2_{L^2(\Omega_T)} \leqslant C. \tag{3.13b}$$

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

$$\|U_\varepsilon - U^\pm_\varepsilon\|^2_{L^2(0,T;H^1(\Omega))} + \|V_\varepsilon - V^\pm_\varepsilon\|^2_{L^2(0,T;H^1(\Omega))}$$

$$\leqslant \left\|\bar{\tau} \frac{\partial U_\varepsilon}{\partial t}\right\|^2_{L^2(0,T;H^1(\Omega))} + \left\|\bar{\tau} \frac{\partial V_\varepsilon}{\partial t}\right\|^2_{L^2(0,T;H^1(\Omega))} \leqslant C\tau. \tag{3.14}$$

Hence on noting (3.13a), (3.14), $(U_\varepsilon(\cdot, t), V_\varepsilon(\cdot, t)) \in K^h$, and (1.13a), with e.g. $X_1 \equiv H^1(\Omega)$, $X_2 \equiv L^s(\Omega)$ and $X_3 \equiv H^{-1}(\Omega)$, 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 imply, on noting (1.13b), 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 [9]. Moreover, the proofs of the results (3.9)–(3.10b) and the result on U_ε in (3.10c) are also in [9, 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 $\zeta \in (0, 1)$, $s \in (2/(1 - \zeta), \infty)$ and any $\bar{s} \in (2/(1 - \zeta), s)$, we have, on noting the compact embedding $W^{1,\bar{s}}(\Omega) \hookrightarrow C^{0,\zeta}(\overline{\Omega})$, (3.14) and (3.10b),

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

where $q = 2(s - \bar{s})/(s - 2)\bar{s} \in (0, 1)$. 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

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

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

$$
\begin{aligned}
\left| \int_{\Omega_T} \nabla(U_\varepsilon^+ - \pi^h u).\nabla U_\varepsilon^+ \, dx \, dt \right| &= \left| -\int_0^T \langle \Delta^h U_\varepsilon^+, U_\varepsilon^+ - \pi^h u \rangle^h \, dt \right| \\
&\leqslant C\|\Delta^h U_\varepsilon^+\|_{L^2(\Omega_T)} \|U_\varepsilon^+ - \pi^h u\|_{L^2(\Omega_T)}.
\end{aligned}
\tag{3.16b}
$$

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 $\partial u^0/\partial \nu = 0$ 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 [8, 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 $(\widetilde{P}_\varepsilon^{h,\tau})$ (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 [4]). Finally, the assumption (ii) yields no real restriction on the time step size τ in terms of h, as the requirement $\varepsilon \to 0$ when $h \to 0$ is very weak. Note also that the positive constant C in (ii) can be chosen arbitrarily.

From (3.13a), (2.22a,b), (2.21), (1.9) 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

of 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. $t \in (0, T)$,

$$B_\delta(t) := \{x \in \overline{\Omega} : |u(x, t)| < 1 - \delta\} \subset D_\delta(t) := \{x \in \overline{\Omega} : -1 + \delta < u(x, t)\}, \quad (3.17a)$$

$$B_{\delta, I}(t) := \{x \in B_\delta(t) : |v(x, t)| < \tfrac{1}{\sqrt{3}}(1 + u(x, t) - \delta)\}, \quad (3.17b)$$

$$B_{\delta, +}(t) := \{x \in B_\delta(t) : v(x, t) - \tfrac{1}{\sqrt{3}}(1 + u(x, t)) \in [-\delta/\sqrt{3}, 0]\}, \quad (3.17c)$$

$$B_{\delta, -}(t) := \{x \in B_\delta(t) : v(x, t) + \tfrac{1}{\sqrt{3}}(1 + u(x, t)) \in [0, \delta/\sqrt{3}]\}. \quad (3.17d)$$

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

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

As $\fint u(\cdot, t) = \fint u^0 \in (-1, 1)$ for a.a. $t \in (0, T)$, it follows that there exists a $\delta_0 \in (0, 1 - |\fint u^0|)$ 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 $\delta_1, \delta_2 \in (0, \delta_0)$ with $\delta_1 > \delta_2$ that if

$$\text{either} \quad y_1 \in B_{\delta_1}(t) \text{ and } y_2 \in \partial B_{\delta_2}(t) \quad \text{or} \quad y_1 \in D_{\delta_1}(t) \text{ and } y_2 \in \partial D_{\delta_2}(t) \text{ with } y_2 \notin \partial\Omega$$

then

$$C_x(t)|y_1 - y_2|^\zeta \geqslant |u(y_1, t) - u(y_2, t)| > \delta_1 - \delta_2, \quad (3.19a)$$

and if $y_1 \in B_{\delta_1, I}(t)$ and $y_2 \in \partial B_{\delta_2, I}(t)$ with $y_2 \notin \partial\Omega$ then

$$C_x(t)|y_1 - y_2|^\zeta \geqslant |u(y_1, t) - u(y_2, t)| + |v(y_1, t) - v(y_2, t)| > \tfrac{1}{\sqrt{3}}(\delta_1 - \delta_2), \quad (3.19b)$$

where $\partial B_\delta(t)$, $\partial D_\delta(t)$ and $\partial B_{\delta, I}(t)$ are the boundaries of the respective sets. 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 $h \leqslant h_0(\delta, t)$ there exist collections of triangles $\mathcal{T}_{B, \delta, I}^h(t) \subset \mathcal{T}_{B, \delta}^h(t) \subset \mathcal{T}_{D, \delta}^h(t) \subset \mathcal{T}^h$ such that

$$B_\delta(t) \subset B_\delta^h(t) := \bigcup_{\sigma \in \mathcal{T}_{B, \delta}^h(t)} \overline{\sigma} \subset B_{\delta/2}(t), \quad D_\delta(t) \subset D_\delta^h(t) := \bigcup_{\sigma \in \mathcal{T}_{D, \delta}^h(t)} \overline{\sigma} \subset D_{\delta/2}(t), \quad (3.20a)$$

$$B_{\delta, I}(t) \subset B_{\delta, I}^h(t) := \bigcup_{\sigma \in \mathcal{T}_{B, \delta, I}^h(t)} \overline{\sigma} \subset B_{\delta/2, I}(t). \quad (3.20b)$$

Clearly, from (3.17a,b) we have

$$\delta_2 < \delta_1 < \delta_0 \implies h_0(\delta_2, t) \leqslant h_0(\delta_1, t).$$

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

$$1 - 2\delta \leqslant |U_\varepsilon^\pm(x, t)| \quad \forall x \notin B_\delta(t), \quad |U_\varepsilon^\pm(x, t)| < 1 - \delta/2 \quad \forall x \in B_\delta(t), \quad (3.21a)$$

$$U_\varepsilon^\pm(x, t) \leqslant -1 + 2\delta \quad \forall x \notin D_\delta(t), \quad -1 + \delta/2 < U_\varepsilon^\pm(x, t) \quad \forall x \in D_\delta(t); \quad (3.21b)$$

$$|V_\varepsilon^\pm(x,t)| < \tfrac{1}{\sqrt{3}}(1 + U_\varepsilon^\pm(x,t)) \qquad\qquad \forall x \in B_{\delta,I}(t), \qquad (3.22a)$$

$$V_\varepsilon^\pm(x,t) - \tfrac{1}{\sqrt{3}}(1 + U_\varepsilon^\pm(x,t)) \in [-2\delta/\sqrt{3}, 0] \quad \forall x \in B_{\delta,+}(t), \qquad (3.22b)$$

$$V_\varepsilon^\pm(x,t) + \tfrac{1}{\sqrt{3}}(1 + U_\varepsilon^\pm(x,t)) \in [0, 2\delta/\sqrt{3}] \qquad \forall x \in B_{\delta,-}(t); \qquad (3.22c)$$

and

$$\varepsilon \leqslant \delta. \qquad (3.23)$$

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

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

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, as $h \to 0$ we have

$$\Lambda_\varepsilon(U_\varepsilon^-)\nabla\Phi_\varepsilon^+ \to \mathcal{H}_{\{u > -1\}}c(u)\nabla\phi \quad \text{weakly in } L^2(\Omega_T), \qquad (3.25a)$$

$$\Xi_\varepsilon(U_\varepsilon^-)\nabla\Phi_\varepsilon^+ \to \mathcal{H}_{\{|u| < 1\}}b(u)\nabla\phi \quad \text{weakly in } L^2(\Omega_T), \qquad (3.25b)$$

$$\Xi_\varepsilon(U_\varepsilon^-)\nabla W_\varepsilon^+ \to \mathcal{H}_{\{|u| < 1\}}b(u)\nabla w \quad \text{weakly in } L^2(\Omega_T), \qquad (3.25c)$$

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

$$\int_{\{|u(\cdot,t)|<1\}} [\gamma\nabla u.\nabla(\eta_1 - u) + (\gamma^{-1}\psi_{,u}(u,v) - w)(\eta_1 - u)]\,\mathrm{d}x$$

$$+ \int_\Omega [\gamma\nabla v.\nabla(\eta_2 - v) + (\gamma^{-1}\psi_{,v}(u,v) - z)(\eta_2 - v)]\,\mathrm{d}x \geqslant 0$$

$$\forall(\eta_1, \eta_2) \in K \text{ with } \mathrm{supp}(\eta_1 - u) \subset \{|u(\cdot, t)| < 1\}. \qquad (3.26)$$

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

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

it follows that

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

Proof. This lemma is a generalisation of Lemma 3.4 in [9]. 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 (1.7), 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

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

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

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

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

$$\Xi_\varepsilon(U_\varepsilon^-)\nabla W_\varepsilon^+ \to f \quad \text{weakly in } L^2(\Omega_T) \text{ as } h \to 0. \tag{3.32}$$

We now identify the function f.

First, we consider a fixed $\delta \in (0, \widehat{\delta_0})$. It follows from (1.9), (2.21), (2.22b), (3.21a) and (3.13a) that for a.a. $t \in (0, T)$ and for all $h \leqslant \widehat{h}_0(\delta, t)$,

$$\delta(1 - \delta/4)|\nabla W_\varepsilon^+(\cdot, t)|_{0, B_\delta(t)}^2 = b(1 - \delta/2)|\nabla W_\varepsilon^+(\cdot, t)|_{0, B_\delta(t)}^2 \leqslant b_\varepsilon(1 - \delta/2)|\nabla W_\varepsilon^+(\cdot, t)|_{0, B_\delta(t)}^2$$
$$\leqslant |([\Xi_\varepsilon(U_\varepsilon^-)]^{1/2}\nabla W_\varepsilon^+)(\cdot, t)|_0^2 \leqslant C(t). \tag{3.33}$$

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

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

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

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

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

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

see [9, (3.18)] for details. Recalling the notation (2.36a,b), we see from cases (i) $A_j^U = A_j^V = 0$, (iii) $A_j^U + \tfrac{1}{\sqrt{3}}A_j^V = 0$ and (iv) $A_j^U - \tfrac{1}{\sqrt{3}}A_j^V = 0$ in the proof of Lemma 2.4, on noting (2.20), (3.1b), (3.22a–c) and (3.20b), that for a.a. $t \in (0, T)$ and for all $h \leqslant \widehat{h}_0(\delta/2, t)$,

$$W_\varepsilon^+(\cdot, t) \equiv -\gamma \Delta^h U_\varepsilon^+(\cdot, t) + \gamma^{-1}\psi_{,u}(U_\varepsilon^-(\cdot, t), V_\varepsilon^-(\cdot, t)),$$
$$Z_\varepsilon^+(\cdot, t) \equiv -\gamma \Delta^h V_\varepsilon^+(\cdot, t) + \gamma^{-1}\psi_{,v}(U_\varepsilon^-(\cdot, t), V_\varepsilon^-(\cdot, t)) \quad \text{on } B_{\delta, I}(t); \tag{3.37a}$$
$$W_\varepsilon^+(\cdot, t) \pm \tfrac{1}{\sqrt{3}}Z_\varepsilon^+(\cdot, t) \equiv [-\gamma \Delta^h U_\varepsilon^+(\cdot, t) + \gamma^{-1}\psi_{,u}(U_\varepsilon^-(\cdot, t), V_\varepsilon^-(\cdot, t))]$$
$$\pm \tfrac{1}{\sqrt{3}}[-\gamma \Delta^h V_\varepsilon^+(\cdot, t) + \gamma^{-1}\psi_{,v}(U_\varepsilon^-(\cdot, t), V_\varepsilon^-(\cdot, t))] \quad \text{on } B_{\delta, \pm}(t). \tag{3.37b}$$

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

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

This together with (3.29) and (3.33) implies 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) shows for a.a. $t \in (0, T)$ that as $h \to 0$,

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

We now work towards 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 $\operatorname{supp} \widetilde{\xi} \subset \{x \in \widetilde{\Omega} : |\widetilde{u}(x, t)| < 1 - 3\delta\}$, 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 $\widetilde{\xi}$, we see that there exist $C_0^\infty(\mathbb{R}^2)$ functions whose restrictions to $\overline{\Omega}$ satisfy

$$(\eta_1^{(\ell)}, \eta_2^{(\ell)}) \in K \cap [C^\infty(\overline{\Omega})]^2, \quad \eta_1^{(\ell)}(\cdot) \equiv u^{(\ell)}(\cdot, t) + \xi^{(\ell)}(\cdot) \quad \text{in } C^\infty(\overline{\Omega}),$$

$$\eta_i^{(\ell)} \to \eta_i, \quad u^{(\ell)}(\cdot, t) \to u(\cdot, t), \quad \xi^{(\ell)} \to \xi \quad \text{strongly in } H^1(\Omega) \text{ as } \ell \to \infty. \tag{3.39}$$

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

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

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

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

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

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

and

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

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

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

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

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

$$|\pi^h \eta_2^{(\ell)} - \mathcal{R}_{U_\varepsilon^+}^2(\pi^h \eta_2^{(\ell)})|_h \leqslant \frac{1}{\sqrt{3}}[|\pi^h u^{(\ell)}(\cdot, t) - U_\varepsilon^+(\cdot, t)|_h + |\pi^h \xi^{(\ell)} - \mathcal{R}_{U_\varepsilon^+}^1(\pi^h \xi^{(\ell)})|_h]$$

$$\leqslant \frac{2}{\sqrt{3}} |\pi^h u^{(\ell)}(\cdot, t) - U_\varepsilon^+(\cdot, t)|_h. \tag{3.44b}$$

We now choose $(\eta_1^h, \eta_2^h) \equiv (\chi_1^{(\ell)}, \chi_2^{(\ell)})$ in (3.4d) and analyse the subsequent terms. First, we infer from (2.20), (3.35), (3.44a) and (2.17) that $U_\varepsilon^+(\cdot, t)$, $\chi_1^{(\ell)}(\cdot)$ and $\xi^{(\ell)}(\cdot)$ satisfy

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

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

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

Next, it follows from (3.43), (2.17), (3.44a), (1.7), (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

$$|\langle W_\varepsilon^+ - \gamma^{-1}\psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \chi_1^{(\ell)} - U_\varepsilon^+\rangle^h - \langle W_\varepsilon^+ - \gamma^{-1}\psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \pi^h \xi^{(\ell)}\rangle|$$
$$\leqslant |\langle W_\varepsilon^+ - \gamma^{-1}\psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), (I - \mathcal{R}_{U_\varepsilon^+}^1)(\pi^h \xi^{(\ell)})\rangle^h|$$
$$+ |\langle W_\varepsilon^+ - \gamma^{-1}\psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \pi^h \xi^{(\ell)}\rangle - \langle W_\varepsilon^+ - \gamma^{-1}\psi_{,u}(U_\varepsilon^-, V_\varepsilon^-), \pi^h \xi^{(\ell)}\rangle^h|$$
$$\leqslant C[1 + |W_\varepsilon|_{0, B_\delta(t)}][|\pi^h u^{(\ell)}(\cdot, t) - U_\varepsilon^+(\cdot, t)|_0 + h|\pi^h \xi^{(\ell)}|_1]$$
$$\leqslant C(t)[|\pi^h u^{(\ell)}(\cdot, t) - U_\varepsilon^+(\cdot, t)|_0 + h|\xi^{(\ell)}|_2]. \qquad (3.47)$$

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

$$|\langle Z_\varepsilon^+ - \gamma^{-1}\psi_{,v}(U_\varepsilon^-, V_\varepsilon^-), \chi_2^{(\ell)} - V_\varepsilon^+\rangle^h - \langle Z_\varepsilon^+ - \gamma^{-1}\psi_{,v}(U_\varepsilon^-, V_\varepsilon^-), \pi^h \eta_2^{(\ell)} - V_\varepsilon^+\rangle|$$
$$\leqslant C(t)[|\pi^h u^{(\ell)}(\cdot, t) - U_\varepsilon^+(\cdot, t)|_0 + h(1 + |\eta_2^{(\ell)}|_2)]. \qquad (3.48)$$

Combining (3.45)–(3.48), noting (3.4d), (3.10d), (3.12d), (3.38), (3.6c), (2.15) and letting $h \to 0$, we find, 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

$$\int_{B_\delta(t)} [\gamma \nabla u . \nabla(\eta_1^{(\ell)} - u^{(\ell)}) + (\gamma^{-1}\psi_{,u}(u, v) - w)(\eta_1^{(\ell)} - u^{(\ell)})]\, dx$$
$$+ \int_\Omega [\gamma \nabla v . \nabla(\eta_2^{(\ell)} - v) + (\gamma^{-1}\psi_{,v}(u, v) - z)(\eta_2^{(\ell)} - v)]\, dx \geqslant r^{(\ell)}(t), \qquad (3.49)$$

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

$$\int_{B_\delta(t)} [\gamma \nabla u . \nabla(\eta_1 - u) + (\gamma^{-1}\psi_{,u}(u, v) - w)(\eta_1 - u)]\, dx$$
$$+ \int_\Omega [\gamma \nabla v . \nabla(\eta_2 - v) + (\gamma^{-1}\psi_{,v}(u, v) - z)(\eta_2 - v)]\, dx \geqslant 0. \qquad (3.50)$$

Repeating (3.33), (3.34) and (3.37a)–(3.50) for all $\delta \in (0, \widehat{\delta}_0)$ shows, 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 (P$_\varepsilon^{h,\tau}$), and functions (ϕ, 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, $u(\cdot, 0) = u^0(\cdot)$, $v(\cdot, 0) = v^0(\cdot)$ in $L^2(\Omega)$ and $\fint u(\cdot, t) = \fint u^0$ for a.a. $t \in (0, T)$. Moreover, for all $\eta \in L^2(0, T; H^1(\Omega))$,

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

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

$$\ell(\gamma) \int_0^T \left\langle \frac{\partial v}{\partial t}, \eta \right\rangle \mathrm{d}t + \int_0^T \langle z, \eta \rangle \,\mathrm{d}t = 0; \qquad (3.51c)$$

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

$$\int_{\{|u(\cdot,t)|<1\}} [\gamma\nabla u.\nabla(\eta_1 - u) + (\gamma^{-1}\psi_{,u}(u, v) - w)(\eta_1 - u)] \,\mathrm{d}x$$

$$+ \int_\Omega [\gamma\nabla v.\nabla(\eta_2 - v) + (\gamma^{-1}\psi_{,v}(u, v) - z)(\eta_2 - v)] \,\mathrm{d}x \geqslant 0$$

$$\forall(\eta_1, \eta_2) \in K \text{ with supp } (\eta_1 - u) \subset \{|u(\cdot, t)| < 1\}. \qquad (3.52)$$

Proof. Only (3.51a–c) need to be established, as (3.52) was shown in Lemma 3.2 above. The proof of (3.51a,b) can be found in [9, 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 (P$_\varepsilon^{h,\tau}$). As (2.14a) in (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. With 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

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

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

$$(\underline{\chi}_1 - \underline{U}_\varepsilon^n)^T(\gamma\mathcal{B}\underline{U}_\varepsilon^n - \mathcal{M}\underline{W}_\varepsilon^n) + (\underline{\chi}_2 - \underline{V}_\varepsilon^n)^T(\gamma\mathcal{B}\underline{V}_\varepsilon^n - \mathcal{M}\underline{Z}_\varepsilon^n)$$

$$\geqslant (\underline{\chi}_1 - \underline{U}_\varepsilon^n)^T\underline{s}_1 + (\underline{\chi}_2 - \underline{V}_\varepsilon^n)^T\underline{s}_2 \quad \forall(\underline{\chi}_1, \underline{\chi}_2) \in \mathcal{K}^{\mathcal{J}}, \qquad (4.1c)$$

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

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

We note for later use, as $\sum_{j=1}^{\mathcal{J}} \chi_j = 1$, that $\sum_{j=1}^{\mathcal{J}} \mathcal{B}_{ij} = \sum_{j=1}^{\mathcal{J}} \mathcal{A}_{ij}^{n-1} = 0$. In addition,

$$\underline{r}_1 := \gamma \mathcal{M} \underline{U}_\varepsilon^{n-1} - \alpha \tau_n \mathcal{A}^{n-1} \underline{\Phi}_\varepsilon^n \in \mathbb{R}^{\mathcal{J}}, \qquad \underline{r}_2 := \ell(\gamma) \mathcal{M} \underline{V}_\varepsilon^{n-1} \in \mathbb{R}^{\mathcal{J}},$$

$$\underline{s}_1 := -\gamma^{-1} \mathcal{M} \underline{\psi}_{,u}(\underline{U}_\varepsilon^{n-1}, \underline{V}_\varepsilon^{n-1}) \in \mathbb{R}^{\mathcal{J}}, \qquad \underline{s}_2 := -\gamma^{-1} \mathcal{M} \underline{\psi}_{,v}(\underline{U}_\varepsilon^{n-1}, \underline{V}_\varepsilon^{n-1}) \in \mathbb{R}^{\mathcal{J}},$$

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 $\mathcal{A}^{n-1}$, and 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 $k \geqslant 1$ 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

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

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

$$(\underline{\chi}_1 - \underline{U}_\varepsilon^{n,k})^T (\gamma (\mathcal{B}_D - \mathcal{B}_L) \underline{U}_\varepsilon^{n,k} - \mathcal{M} \underline{W}_\varepsilon^{n,k}) + (\underline{\chi}_2 - \underline{V}_\varepsilon^{n,k})^T (\gamma (\mathcal{B}_D - \mathcal{B}_L) \underline{V}_\varepsilon^{n,k} - \mathcal{M} \underline{Z}_\varepsilon^{n,k})$$

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

The above is the natural extension of the iterative method in [9] for solving the corresponding non-linear algebraic system arising from the corresponding finite element approximation of (1.4). Below, we prove convergence of (4.2a–c) for our non-linear 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 $((U_\varepsilon^{n,k}, V_\varepsilon^{n,k}), W_\varepsilon^{n,k}, Z_\varepsilon^{n,k})_{k \geqslant 0}$ generated by the algorithm (4.2a–c) satisfies

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

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

Proof. The proof is similar to the proof of Theorem 4.1 in [9]. Let $\underline{E}^k := \underline{U}_\varepsilon^n - \underline{U}_\varepsilon^{n,k}$, $\underline{F}^k := \underline{V}_\varepsilon^n - \underline{V}_\varepsilon^{n,k}$, $\underline{P}^k := \underline{W}_\varepsilon^n - \underline{W}_\varepsilon^{n,k}$ and $\underline{Q}^k := \underline{Z}_\varepsilon^n - \underline{Z}_\varepsilon^{n,k}$. Now subtracting (4.2a) from (4.1a) and testing the resulting equation with $\underline{P}^k$ yields

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

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

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

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

$$- \gamma [[\underline{E}^k]^T (\mathcal{B}_D - \mathcal{B}_L) \underline{E}^k + [\underline{F}^k]^T (\mathcal{B}_D - \mathcal{B}_L) \underline{F}^k] + [\underline{E}^k]^T \mathcal{M} \underline{P}^k + [\underline{F}^k]^T \mathcal{M} \underline{Q}^k$$

$$\geqslant -\gamma [[\underline{E}^k]^T \mathcal{B}_L^T \underline{E}^{k-1} + [\underline{F}^k]^T \mathcal{B}_L^T \underline{F}^{k-1}]. \tag{4.5}$$

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

$$\gamma^2 [[\underline{E}^k]^T (\mathcal{B}_D - \mathcal{B}_L) \underline{E}^k + [\underline{F}^k]^T (\mathcal{B}_D - \mathcal{B}_L) \underline{F}^k]$$

$$+ \tau_n [[\underline{P}^k]^T (\mathcal{A}_D - \mathcal{A}_L) \underline{P}^k + [\ell(\gamma)]^{-1} \gamma [\underline{Q}^k]^T \mathcal{M} \underline{Q}^k]$$

$$\leqslant \gamma^2 [[\underline{E}^k]^T \mathcal{B}_L^T \underline{E}^{k-1} + [\underline{F}^k]^T \mathcal{B}_L^T \underline{F}^{k-1}] + \tau_n [\underline{P}^k]^T \mathcal{A}_L^T \underline{P}^{k-1}. \tag{4.6}$$

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

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

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

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

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

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

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

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

where we have again used the fact that $\{\tau_n [\underline{P}^k]^T \mathcal{A}_{D_1}\underline{P}^k\}_{k\geqslant 0}$ has a limit. Combining (4.9), (4.10), (2.2) and (1.15) yields the desired result (4.3a). Similarly, multiplying (4.2b) with $\underline{1}^T := (1, \ldots, 1)$ yields, on noting (4.9),

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

The desired result (4.3b) then follows from (4.11), (4.1b), (4.9), (2.2) and (1.15). $\qquad\square$

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

$$
(\eta_1^h - U_j)(C_1 U_j - b_1) + (\eta_2^h - V_j)(C_2 V_j - b_2) \geqslant 0 \quad \forall(\eta_1^h, \eta_2^h) \in \mathcal{K},
\tag{4.12}
$$

where

$$
C_1 := \gamma\left(\mathcal{B}_{jj} + \frac{[\mathcal{M}_{jj}]^2}{\tau_n \mathcal{A}_{jj}^{n-1}}\right), \quad C_2 := \gamma \mathcal{B}_{jj} + \frac{\ell(\gamma)\mathcal{M}_{jj}}{\tau_n}, \quad b_1 := \widehat{\underline{s}}_j^1 + \frac{\mathcal{M}_{jj}\widehat{\underline{r}}_j^1}{\tau_n \mathcal{A}_{jj}^{n-1}}, \quad b_2 := \widehat{\underline{s}}_j^2 + \frac{\widehat{\underline{r}}_j^2}{\tau_n}.
$$

Clearly, the unique solution to (4.12) is

$$(U_j, V_j) \equiv \mathrm{P}^C_{\mathcal{K}}(b_1/C_1, b_2/C_2),$$

where $\mathrm{P}^C_{\mathcal{K}}(x_1, x_2)$ is the orthogonal projection of the point $\underline{x} = (x_1, x_2)^T \in \mathbb{R}^2$ 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 $C = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}$. The projection $\underline{y} = \mathrm{P}^C_{\mathcal{K}}(\underline{x})$ can be computed as follows:

1. If $\underline{x} \in \mathcal{K}$, then $\underline{y} = \underline{x}$, else

2.   If $x_1 \geqslant 1$ then $\underline{y} := (1, \max\{-2/\sqrt{3}, \min\{x_2, 2/\sqrt{3}\}\})^T$, else

3.   If $x_2 \geqslant 0$ then $\underline{v} := (2, 2/\sqrt{3})^T$, else $\underline{v} := (2, -2/\sqrt{3})^T$.

4. $\alpha := \langle \underline{x} + (1, 0)^T, \underline{v} \rangle_C / \|\underline{v}\|^2_C$.

5. $\underline{y} := (-1, 0)^T + \min\{\max\{\alpha, 0\}, 1\}\underline{v}$.

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

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

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

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

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

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

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

5. Numerical results

Throughout this section, we use (1.8) for ψ in (1.7), 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,

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

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

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

(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.2) and (A.1)). 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 $n \geqslant 1$, $(U_\varepsilon^{n,0}, V_\varepsilon^{n,0}, W_\varepsilon^{n,0}, Z_\varepsilon^{n,0}) \equiv (U_\varepsilon^{n-1}, V_\varepsilon^{n-1}, W_\varepsilon^{n-1}, Z_\varepsilon^{n-1})$, 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

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

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 $L := \min\{L_1, L_2\}$. On using the adaptive finite element code Alberta 1.2 (see [35]), we implemented the same mesh refinement strategy as in [9]. In particular, to improve efficiency we use the approximation $(\widetilde{P}_\varepsilon^{h,\tau})$ (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 $|U_\varepsilon^{n-1}| = 1, |V_\varepsilon^{n-1}| < 2/\sqrt{3}$. Hence we use a refined mesh with mesh size $h_f = 2^{3/2}L/N_f$ in these interfacial regions, and a coarser mesh of mesh size $h_c = 2^{3/2}L/N_c$ away from the interfaces. Here N_f and N_c are parameters (see [9, §5]). We note that as long as the ratio N_f/N_c is kept fixed through a convergence experiment with $h \to 0$, the quasi-uniformity assumption in (A) will not be violated. Furthermore, we choose N_f such that there are always approximately 8 mesh points across the interface in each direction. In particular, for $\mu \geqslant 0$ we will always have $h_f \leqslant \frac{3\sqrt{2}}{32}\gamma\pi$, whereas for $\mu < 0$ we ensure that $h_f \leqslant \frac{3\sqrt{2}}{32}(1 - \mu)^{-1/2}\gamma\pi$.

For our first experiments we choose $\mu = 0$ in (1.8). That means that the function Ψ is symmetric with respect to the three vertices A, B and C of $\mathcal{K}$. In particular, the 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 [26] that the true solution for the void boundary consists of four symmetric branches, where one branch is given by

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

Here $a = (\frac{A}{2\theta - \sin(2\theta)})^{1/2}$ with $A = \pi R^2$ being the total area of the void and $2\theta = 2\pi/3$ being the contact angle between grains and void. We chose the following parameters for $(\widetilde{P}_\varepsilon^{h,\tau})$: $L_1 = L_2 = 0.5$, $\gamma = 1/24\pi$, $\alpha = 0$, $T = 10^{-2}$, $\tau_n = \tau = 5 \times 10^{-8}$. 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 $N_c = 2$.
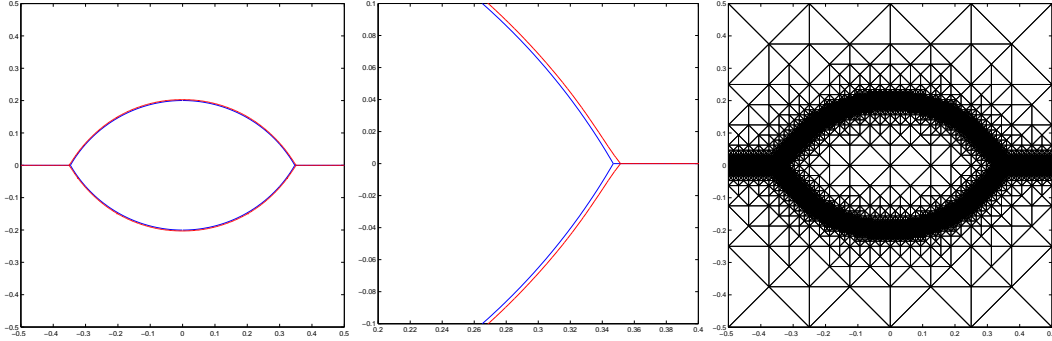
FIG. 3. ($\gamma = 1/24\pi$, $\alpha = 0$) Comparison between computed solution (red) and true solution (blue; see the pdf file for colour plots). The final triangulation is shown on the right.

The comparison between true solution and the numerically steady state can be seen in Figure 3, where we also include a detailed plot at a triple junction and 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 visualise 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.

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 $(\widetilde{P}_\varepsilon^{h,\tau})$: $L_1 = 1$, $L_2 = 0.5$, $\gamma = 1/24\pi$, $\alpha = 5\pi$, $T = 0.012$, $\tau_n = \tau = 5 \times 10^{-8}$. 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 4. 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.

We also include an experiment that produces a travelling wave solution in the absence of electromigration, first mentioned in [30] (see also [27]). 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

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

with $y = (-0.3, -0.3)$, $R = 0.25$. The refinement parameters were $N_f = 256$ and $N_c = 2$. We used the scaling $\ell(\gamma) = \beta\gamma$ with $\beta \in \{1, 1/24\pi\}$ and the following parameters for $(\widetilde{P}_\varepsilon^{h,\tau})$: $L_1 = 1$, $L_2 = 0.5$, $\gamma = 1/24\pi$, $\alpha = 0$, as well as $T = 0.14$, $\tau_n = \tau = 5 \times 10^{-6}$ for $\beta = 1$ and $T = 2.6 \times 10^{-3}$, $\tau_n = \tau = 5 \times 10^{-8}$ for $\beta = 1/24\pi$. The evolution is shown in Figure 5. We note that the travelling wave solutions reported in [30, 27] are for a limiting motion that differs slightly from our sharp interface limit (A.13). In particular, there the material boundaries move by surface diffusion, whereas here, in the limit $\gamma \to 0$, we find that they evolve by a combination of surface diffusion and surface attachment limited kinetics (see (1.3)).
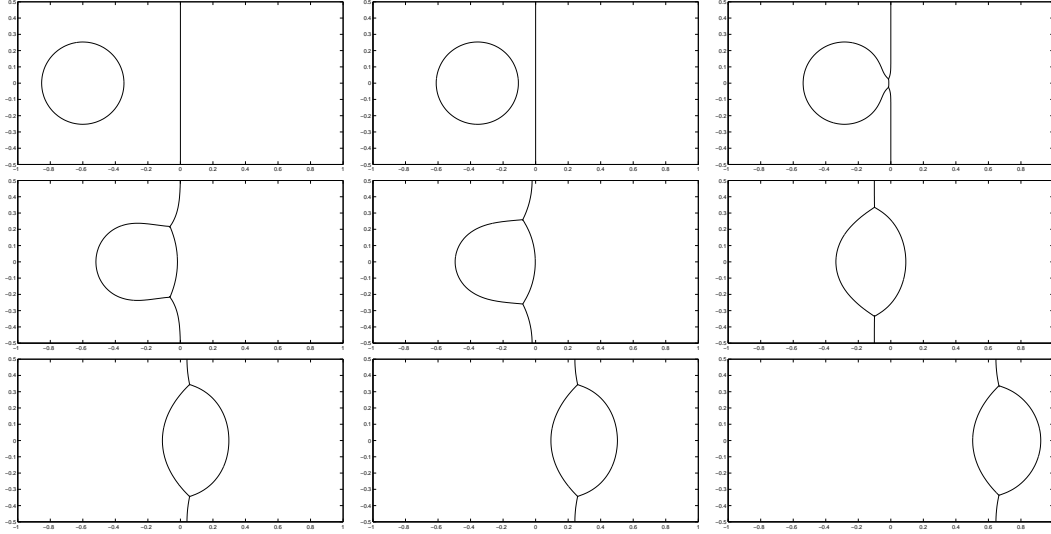
FIG. 4. ($\gamma = 1/24\pi$, $\alpha = 5\pi$) Solution (U_ε, V_ε) at times $t = 0$, 2×10^{-3}, 2.6×10^{-3}, 2.8×10^{-3}, 3×10^{-3}, 4×10^{-3}, 6×10^{-3}, 8×10^{-3}, 0.012.
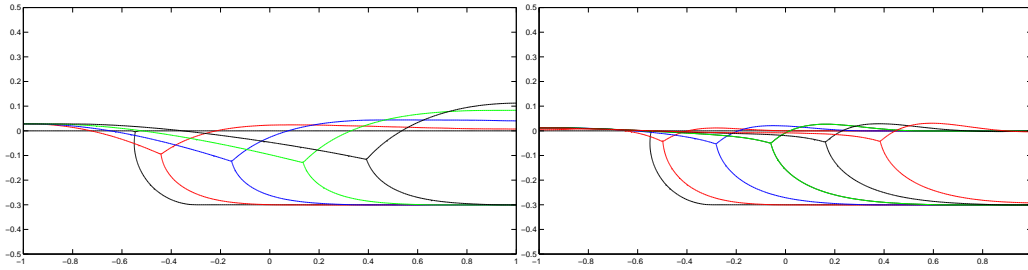

FIG. 5. ($\gamma = 1/24\pi$, $\alpha = 0$) Solution (U_ε, V_ε) at times $t = 0$, 0.02, 0.06, 0.1, 0.14 for $\beta = 1$ (left); and $t = 0$, 2×10^{-4}, 8×10^{-4}, 1.4×10^{-3}, 2×10^{-3}, 2.6×10^{-3} for $\beta = 1/24\pi$ (right).

5.1   *Different scalings*

In this subsection, we provide numerical simulations that highlight the difference between the two scalings $\ell(\gamma) = \gamma$ and $\ell(\gamma) = \gamma^2$ and their respective sharp interface limits, as discussed in the Appendix. 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. $\gamma = 1/12\pi$, $1/24\pi$, $1/48\pi$. In particular, we set $L_1 = L_2 = 0.5$, $T = 4 \times 10^{-3}$, $\tau_n = \tau = 288(\gamma\pi)^2 \times 10^{-7}$, $\varepsilon = 48\gamma\pi \times 10^{-5}$ and used the appropriate refinement parameters $N_f = (32/3)(1/\gamma\pi)$ and $N_c = N_f/8$. Considerations using formal asymptotic expansions (see (A.12)) show 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 6, where we plot the results for $\gamma = 1/12\pi$, $\gamma = 1/24\pi$ and $\gamma = 1/48\pi$.
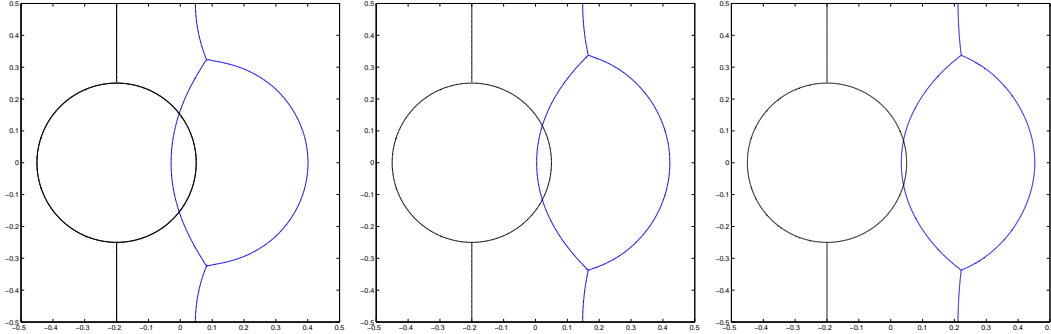
FIG. 6. ($\alpha = 5\pi$) Solution (U_ε, V_ε) at times $t = 0$, $T = 4 \times 10^{-3}$ for $\gamma = 1/12\pi$, $\gamma = 1/24\pi$ and $\gamma = 1/48\pi$.
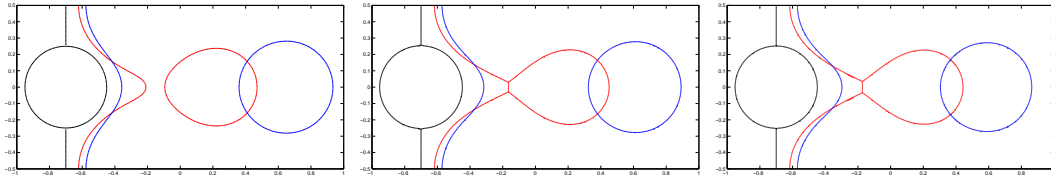

FIG. 7. ($\alpha = 5\pi$) Solution (U_ε, V_ε) at times $t = 0$, 0.04, $T = 0.056$ for $\gamma = 1/12\pi$, $\gamma = 1/24\pi$ and $\gamma = 1/48\pi$.

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.13)). 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 $N_f = (32/3)(1/\gamma\pi)$, $N_c = N_f/8$. In Figure 7 one can see that the void detaches from the grain boundary. Note also the very good agreement between the results as γ is decreased.

5.2 Different contact angles

In this subsection, we report on contact angles for the triple junction that are different from the symmetric case $2\pi/3$. 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 (1.7)) such that the grain and material boundaries have different surface energies. To this end, we use (1.8) with $\mu \neq 0$.

Assume we are given the ratio of the surface energies for the grain and material boundaries, σ^A/σ^B, where we have adopted the notation of the Appendix (see (A.9)). Then this angle law, where $\sigma^B = \sigma^C$ is the surface energy of the material boundary and σ^A is the surface energy of the grain boundary, yields

$$\theta^A = 2\arccos\left(\frac{1}{2}\frac{\sigma_{\text{grain}}}{\sigma_{\text{mat}}}\right).$$

Using (A.4) we compute for $\mu \in (-2, 4/7)$ that

$$\frac{\sigma_{\text{grain}}}{\sigma_{\text{mat}}} = \frac{\frac{2}{3}\pi(1-\mu)^{1/2}}{\frac{2}{3}\pi(1-\mu/4)^{1/2}} = 2\left(\frac{1-\mu}{4-\mu}\right)^{1/2}. \tag{5.4}$$

In the derivation of (A.4) it is assumed that the first order solution to the variational inequality (1.6c) leads after a suitable rescaling to a minimizer 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 $(\widetilde{P}_\varepsilon^{h,\tau})$ 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 $\mu = 1/2$, we computed the different surface energies for the grain and material boundaries in this way and obtained a ratio $\sigma_{\text{grain}}/\sigma_{\text{mat}} \approx 0.758$, i.e. almost exactly the value $2/\sqrt{7}$ 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).
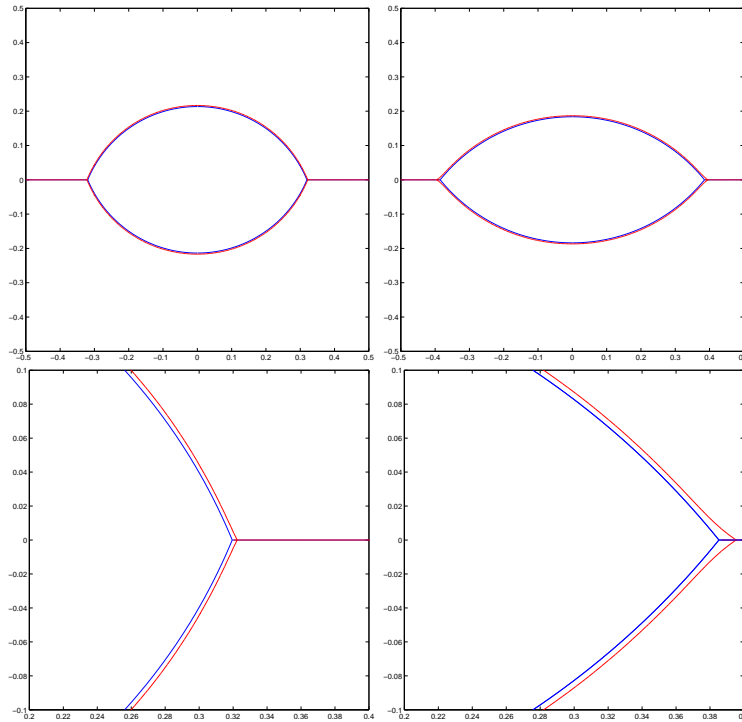

FIG. 8. ($\gamma = 1/24\pi$, $\mu = 1/2$ (left) and $\gamma = \sqrt{2}/24\pi$, $\mu = -1$ (right)) Comparison between computed solution (red) and true solution (blue) for $\theta^A = 135°$ and $\theta^A = 102°$, respectively.

Next, we computed the different surface energies for the grain and material boundaries numerically for the case $\mu = -1$ and obtained a ratio $\sigma_{\text{grain}}/\sigma_{\text{mat}} \approx 1.26$, i.e. almost exactly the

value $(8/5)^{1/2}$ 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 $(\widetilde{P}_{\varepsilon}^{h,\tau})$ as before, except $\gamma = \sqrt{2}/24\pi$.

A. Formal asymptotic expansions

In this appendix we give a short discussion of the sharp interface limit of (1.6a–f) when $\gamma \to 0$. A more detailed analysis of the formal asymptotic expansions can be found in [7]. To identify the limit, three different types of expansions have to be used. In regions where either a grain or a 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. We refer to [11, 24, 32, 23] for details on the method of formally matched asymptotic expansions for systems with triple junctions, and we state here only the results if one applies the method to our system. However, it should be noted that the above references only consider smooth potentials. Therefore we have to be particularly careful with the asymptotics at triple junctions with our obstacle potential, as we have to deal with a variational inequality as opposed to an equation.

The equations for the outer expansion imply that the vector (u, v) attains to leading order 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 find that it solves Laplace's equation in the regions where (u, v) is either B or C.

Now the inner expansion has to be used to determine the governing equations on the interface. There are three interfaces (curves in two dimensions) for which we seek these laws. Let $\Gamma^{ij} = (\Gamma^{ij}(t))_{t \geqslant 0}$ with either $(i, j) = (A, B)$, (B, C) or (C, A) be a smooth evolving curve, describing an interface between regions occupied by i and j.

As usual in the theory of formally matched asymptotic expansions for phase field systems, one introduces new coordinates (ρ, s). Here s is an arclength parameter along the interface and $\rho = \gamma^{-1}d$ is a rescaled signed distance where $d(x, t)$ is the signed distance of a point x to $\Gamma^{ij}(t)$, which is positive if x belongs to phase j. In the following we will suppress the indices ij if no confusion arises.

Considering (1.6c), we find that the leading order solution, which we denote by (u_0, v_0), has the following structure. At a grain boundary, Γ^{BC}, with

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

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

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

is a solution. A simple computation shows that we have to require $\mu \in (-2, 4/7)$ in order to make sure that (u_0, v_0) solves the variational inequality. Similarly, at a material boundary, Γ^{AB}, with

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

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

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

is to leading order a solution of the variational inequality (1.6c). The solution of the material boundary CA is then given, through symmetry, as $(u_0, v_0)(\rho) = (\bar{u}, (1 + \bar{u})/\sqrt{3})(-\rho)$.

For later use we compute the interfacial energy

$$
\sigma = \int_{-\infty}^{\infty} [\tfrac{1}{2}((\partial_\rho u_0)^2 + (\partial_\rho v_0)^2) + \Psi(u_0, v_0)] \, \mathrm{d}\rho = \int_{-\infty}^{\infty} [(\partial_\rho u_0)^2 + (\partial_\rho v_0)^2] \, \mathrm{d}\rho
$$
$$
= 2 \int_{-\infty}^{\infty} \sqrt{(\partial_\rho u_0)^2 + (\partial_\rho v_0)^2} \sqrt{\tfrac{1}{2} \Psi(u_0, v_0)} \, \mathrm{d}\rho \tag{A.3}
$$

of the solutions (u_0, v_0) above. The formula (A.3) coincides with σ^{ij} in (1.1) if (u_0, v_0), upon rescaling, is not only a stationary point but in fact the minimum in (1.1) (see [36]). 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 $(u_0, v_0) = (\bar{u}, \pm(1 + \bar{u})/\sqrt{3})$ at the material boundary, we obtain

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

respectively.

To derive equations for the grain boundary and the void boundaries, a solvability condition for the first order equation in (1.6c) has to be used (see [12, 24] for similar approaches). If we employ this approach in our context, we infer, depending on the scaling in (1.6b), for the grain boundary that

$$
\kappa = 0 \quad \text{if } \ell(\gamma) := \beta\gamma^2 \quad \text{and} \quad \beta\omega\mathcal{V} = \sigma\kappa \quad \text{if } \ell(\gamma) := \beta\gamma, \tag{A.5}
$$

where $\omega := \int_{-\infty}^{\infty} (\partial_\rho v_0)^2 \, \mathrm{d}\rho = \sigma_{\text{grain}} = \tfrac{2}{3}\pi(1 - \mu)^{1/2}$, on recalling (A.3)–(A.4). Obviously, the factors ω and σ cancel in (A.5). 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 (1.6b) degenerates on grain boundaries, i.e. we obtain $\partial u/\partial t = 0$, and (1.6e) has no interfacial structure on grain boundaries since $c(u_0)$ is constant. In order to derive the governing equation for the void boundary we have to use an approach introduced in [13]. There the authors showed that the diffusion equation (1.6a) gives

$$
-\mathcal{V}[u_0]_i^j = M\partial_{ss}(w_0 + \alpha\phi_0), \tag{A.6}
$$

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 $M := \int_{-\infty}^{\infty} b(u_0(\rho)) \, \mathrm{d}\rho = \rho_m = \pi(4 - \mu)^{-1/2}$. Exploiting (1.6b) and the variational inequality gives to the order $\mathcal{O}(1)$, on recalling (A.3),

$$
\sigma\kappa = [u_0]_i^j w_0 \quad \text{if } \ell(\gamma) := \beta\gamma^2 \quad \text{and} \quad \sigma\kappa = [u_0]_i^j w_0 + \beta\omega\mathcal{V} \quad \text{if } \ell(\gamma) := \beta\gamma, \tag{A.7}
$$

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

For the material interfaces $(i, j) = (A, B)$, (C, A), and the grain interface (B, C), we derive from (A.5), (A.6) and (A.7) for the scaling $\ell(\gamma) := \beta\gamma$ that

$$-2\mathcal{V}^{AB} = M^{AB}\partial_{ss}(w_0^{AB} + \alpha\phi_0^{AB}), \qquad 2w_0^{AB} + \beta\omega^{AB}\mathcal{V}^{AB} = \sigma^{AB}\kappa^{AB}, \qquad \text{(A.8a)}$$

$$2\mathcal{V}^{CA} = M^{CA}\partial_{ss}(w_0^{CA} + \alpha\phi_0^{CA}), \qquad -2w_0^{CA} + \beta\omega^{CA}\mathcal{V}^{CA} = \sigma^{CA}\kappa^{CA}, \qquad \text{(A.8b)}$$

$$\beta\omega^{BC}\mathcal{V}^{BC} = \sigma^{BC}\kappa^{BC}, \qquad \text{(A.8c)}$$

where, on recalling (A.4), we have

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

The evolution laws (A.8a,b) for the material interfaces combine surface diffusion and surface attachment limited kinetics (SALK), which was discussed in [37]; see also [21].

If we choose the scaling $\ell(\gamma) := \beta\gamma^2$ instead of $\ell(\gamma) := \beta\gamma$ in the evolution equation (1.6b) we derive from (A.6), (A.5) and (A.7) that

$$\mathcal{V}^{AB} = -\tfrac{1}{2}M^{AB}\partial_{ss}(\tfrac{1}{2}\sigma^{AB}\kappa^{AB} + \alpha\phi_0^{AB}), \qquad \mathcal{V}^{CA} = -\tfrac{1}{2}M^{CA}\partial_{ss}(\tfrac{1}{2}\sigma^{CA}\kappa^{CA} - \alpha\phi_0^{CA}),$$
$$\sigma^{BC}\kappa^{BC} = 0.$$

Therefore under this scaling the evolution of the void interface is given by surface diffusion (see [13]), 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. Using the approach in [11], where an asymptotic expansion close to the triple junction was considered, we obtain

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

where τ_Γ^1, τ_Γ^2 and τ_Γ^3 are the tangents to the interfaces at the triple junction all pointing away from the triple junction. This is the force balance at the triple junction and a simple computation shows that the above identity is equivalent to Young's law,

$$\frac{\sin\theta^A}{\sigma^A} = \frac{\sin\theta^B}{\sigma^B} = \frac{\sin\theta^C}{\sigma^C}, \qquad \text{(A.9)}$$

where θ^A, θ^B and θ^C are the angles that the regions A, B and C form at the triple junction. Using the ideas in [32, 24] we can also derive a flux balance condition

$$M^C\partial_s(w_0^C + \alpha\phi_0^C) + M^B\partial_s(w_0^B + \alpha\phi_0^B) = 0 \qquad \text{(A.10)}$$

at the triple junction.

It remains to determine an additional condition at the triple junction, which is related to the fact that the chemical potential is continuous. In fact, similarly to [32, 24], we derive that at the triple junction

$$w_0^C = w_0^B. \qquad \text{(A.11)}$$

We remark that the choice of scaling $\ell(\gamma) := \beta\gamma$ or $\beta\gamma^2$ does not affect the conditions (A.9), (A.10) and (A.11) at the triple junction, as the equation (1.6b) was not used to derive them. Of course, under the scaling $\ell(\gamma) := \beta\gamma^2$ we deduce from (A.7) and (A.11) that at the triple junction

$$\sigma^C \kappa^C = -\sigma^B \kappa^B.$$

Finally, when an interface meets the external boundary, further boundary conditions have to hold which can be derived as in [24, 32]. We include these conditions in the summary below.

To summarize, we obtain, depending on the scaling in (1.6b), the following two sharp interface problems. In both cases we find that at a triple junction the identities (A.9), (A.10) and (A.11) have to hold for w and ϕ. When an interface meets an external boundary a 90° angle condition has to hold. In addition, at points where the material boundary intersects $\partial\Omega$, we have $\partial_s(w^i + \alpha\phi^i) = 0$ for $i \in \{B, C\}$. Firstly, the scaling $\ell(\gamma) := \beta\gamma^2$ leads to

$$\begin{aligned}
\mathcal{V}^C &= -\tfrac{1}{2}M^C \partial_{ss}(\tfrac{1}{2}\sigma^C\kappa^C + \alpha\phi^C) && \text{on } \Gamma^C, \\
\mathcal{V}^B &= -\tfrac{1}{2}M^B \partial_{ss}(\tfrac{1}{2}\sigma^B\kappa^B - \alpha\phi^B) && \text{on } \Gamma^B, \\
0 &= \kappa^A && \text{on } \Gamma^A.
\end{aligned} \tag{A.12}$$

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

$$\begin{aligned}
\pm 2\mathcal{V}^i &= M^i \partial_{ss}(w^i + \alpha\phi^i) && \text{on } \Gamma^i \text{ for } i \in \{B, C\}, \\
\beta\omega^i \mathcal{V}^i &= \sigma^i\kappa^i \pm 2w^i && \text{on } \Gamma^i \text{ for } i \in \{B, C\}, \\
\beta\omega^A \mathcal{V}^A &= \sigma^A\kappa^A && \text{on } \Gamma^A,
\end{aligned} \tag{A.13}$$

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

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

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

Acknowledgements

REFERENCES

1. AVERBUCH, A., ISRAELI, M., & RAVVE, I. Electromigration of intergranular voids in metal films for microelectronic interconnects. *J. Comput. Phys.* **186** (2003), 481–502. Zbl 1036.78514

2. BALDO, S. Minimal interface criterion for phase transitions in mixtures of Cahn–Hilliard fluids. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **7** (1990), 67–90. Zbl 0702.49009 MR 1051228

3. BARRETT, J. W., & BLOWEY, J. F. Finite element approximation of a degenerate Allen–Cahn/Cahn–Hilliard system. *SIAM J. Numer. Anal.* **39** (2001), 1598–1624. Zbl 1012.35004 MR 1885708

4. BARRETT, J. W., BLOWEY, J. F., & GARCKE, H. Finite element approximation of the Cahn–Hilliard equation with degenerate mobility. *SIAM J. Numer. Anal.* **37** (1999), 286–318. Zbl 0947.65109 MR 1742748

5. BARRETT, J. W., BLOWEY, J. F., & GARCKE, H. On fully practical finite element approximations of degenerate Cahn–Hilliard systems. *M2AN Math. Model. Numer. Anal.* **35** (2001), 713–748. Zbl 0987.35071 MR 1863277

6. BARRETT, J. W., GARCKE, H., & NÜRNBERG, R. Finite element approximation of a phase field model for surface diffusion of voids in a stressed solid. *Math. Comp.* **75** (2006), 7–41. Zbl 1078.74050 MR 2176388

7. BARRETT, J. W., GARCKE, H., & NÜRNBERG, R. On sharp interface limits of Allen–Cahn/Cahn–Hilliard variational inequalities. *Discrete Contin. Dyn. Syst.* (2007), to appear.

8. BARRETT, J. W., & NÜRNBERG, R. 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** (2004), 323–363. Zbl pre02109225 MR 2046180

9. BARRETT, J. W., NÜRNBERG, R., & STYLES, V. Finite element approximation of a phase field model for void electromigration. *SIAM J. Numer. Anal.* **42** (2004), 738–772. Zbl 1076.78012 MR 2084234

10. BHATE, D. N., KUMAR, A., & BOWER, A. F. Diffuse interface model for electromigration and stress voiding. *J. Appl. Phys.* **87** (2000), 1712–1721.

11. BRONSARD, L., & REITICH, F. On three-phase boundary motion and the singular limit of a vector-valued Ginzburg–Landau equation. *Arch. Ration. Mech. Anal.* **124** (1993), 355–379. Zbl 0785.76085 MR 1240580

12. CAGINALP, G., & FIFE, P. C. Dynamics of layered interfaces arising from phase boundaries. *SIAM J. Appl. Math.* **48** (1988), 506–518. MR 0941098

13. CAHN, J. W., ELLIOTT, C. M., & NOVICK-COHEN, A. The Cahn–Hilliard equation with a concentration dependent mobility: motion by minus the Laplacian of the mean curvature. *European J. Appl. Math.* **7** (1996), 287–301. Zbl 0861.35039 MR 1401172

14. CAHN, J. W., & NOVICK-COHEN, A. Evolution equations for phase separation and ordering in binary alloys. *J. Statist. Phys.* **76** (1994), 877–909. Zbl 0840.35110

15. CAHN, J. W., & NOVICK-COHEN, A. Limiting motion for an Allen–Cahn/Cahn–Hilliard system. *Free Boundary Problems, Theory and Applications* (Zakopane, 1995), Pitman Res. Notes Math. Ser. 363, Longman, Harlow (1996), 89–97. Zbl 0865.35145 MR 1462972

16. CAHN, J. W., & NOVICK-COHEN, A. Motion by curvature and impurity drag: Resolution of a mobility paradox. *Acta Mater.* **48** (2000), 3425–3440.

17. CIARLET, P. G. *The Finite Element Method for Elliptic Problems*. Stud. Math. Appl. 4, North-Holland, Amsterdam (1978). Zbl 0511.65078 MR 0520174

18. CIARLET, P. G. *Introduction to Numerical Linear Algebra and Optimisation*. Cambridge Univ. Press, Cambridge (1988). Zbl 0672.65001 MR 1015713

19. CUMMINGS, L. J., RICHARDSON, G., & BEN AMAR, M. Models of void electromigration. *European J. Appl. Math.* **12** (2001), 97–134. Zbl 0991.78003 MR 1931301

20. ELLIOTT, C. M. Approximation of curvature dependent interface motion. *The State of the Art in Numerical Analysis*, I. S. Duff and G. A. Watson (eds.), Oxford Univ. Press, New York (1997), 407–440. Zbl 0881.65131 MR 1628355

21. ELLIOTT, C. M., & GARCKE, H. Existence results for diffusive surface motion laws. *Adv. Math. Sci. Appl.* **7** (1997), 465–488. Zbl 0876.35050 MR 1454678

22. GARCKE, H., & NESTLER, B. A mathematical model for grain growth in thin metallic films. *Math. Models Methods Appl. Sci.* **10** (2000), 895–921. Zbl pre01882736 MR 1771511

23. GARCKE, H., NESTLER, B., & STOTH, B. On anisotropic order parameter models for multi-phase systems and their sharp interface limits. *Phys. D* **115** (1998), 87–108. Zbl 0936.82010 MR 1616772

24. GARCKE, H., & NOVICK-COHEN, A. A singular limit for a system of degenerate Cahn–Hilliard equations. *Adv. Differential Equations* **5** (2000), 401–434. Zbl 0988.35019 MR 1750107

25. GRÜN, G., & RUMPF, M. Nonnegativity preserving numerical schemes for the thin film equation. *Numer. Math.* **87** (2000), 113–152. Zbl 0988.76056 MR 1800156

26. ITO, K., & KOHSAKA, Y. Three-phase boundary motion by surface diffusion: stability of a mirror symmetric stationary solution. *Interfaces Free Bound.* **3** (2001), 45–80. Zbl 0972.35165 MR 1805078

27. KANEL, J., NOVICK-COHEN, A., & VILENKIN, A. Coupled surface and grain boundary motion: nonclassical traveling wave solutions. *Adv. Differential Equations* **9** (2004), 299–327. Zbl pre05054476 MR 2100630

28. KINDERLEHRER, D., & STAMPACCHIA, G. *An Introduction to Variational Inequalities and Their Applications*. Academic Press (1980). Zbl 0457.35001 MR 1786735

29. MAHADEVAN, M., & BRADLEY, R. M. Phase field model of surface electromigration in single crystal metal thin films. *Phys. D* **126** (1999), 201–213.

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

31. NOVICK-COHEN, A. The Cahn–Hilliard equation: mathematical and modelling perspectives. *Adv. Math. Sci. Appl.* **8** (1998), 965–985. Zbl 0917.35044 MR 1657208

32. NOVICK-COHEN, A. Triple-junction motion for an Allen–Cahn/Cahn–Hilliard system. *Phys. D* **137** (2000), 1–24. Zbl 0954.35081 MR 1738763

33. NOVICK-COHEN, A., & PERES HARI, L. Geometric motion for a degenerate Allen–Cahn/Cahn–Hilliard system: The partial wetting case. *Phys. D* **209** (2005), 205–235. Zbl pre02229549 MR 2167453

34. RUBINSTEIN, J., & STERNBERG, P. Nonlocal reaction-diffusion equations and nucleation. *IMA J. Appl. Math.* **48** (1992), 249–264. Zbl 0763.35051 MR 1167735

35. SCHMIDT, A., & SIEBERT, K. G. *Design of Adaptive Finite Element Software: The Finite Element Toolbox ALBERTA*. Lecture Notes in Comput. Sci. Engrg. 42, Springer, Berlin (2005). Zbl 1068.65138 MR 2127659

36. STERNBERG, P. Vector-valued local minimizers of nonconvex variational problems. *Current Directions in Nonlinear Partial Differential Equations* (Provo, UT, 1987), *Rocky Mountain J. Math.* **21** (1991), 799–807. MR 1121542 Zbl 0737.49009

37. TAYLOR, J. E., & CAHN, J. W. Linking anisotropic sharp and diffuse surface motion laws via gradient flows. *J. Statist. Phys.* **77** (1994), 183–197. Zbl 0844.35044 MR 1300532

38. XIA, L., BOWER, A. F., SUO, Z., & SHIH, C. A finite element analysis of the motion and evolution of voids due to strain and electromigration induced surface diffusion. *J. Mech. Phys. Solids* **45** (1997), 1473–1493. Zbl 0974.74566

39. ZHORNITSKAYA, L., & BERTOZZI, A. L. Positivity preservinng numerical schemes for lubrication-type equations. *SIAM J. Numer. Anal.* **37** (2000), 523–555. Zbl 0961.76060 MR 1740768