

PAPER

# Diffuse-interface approximation and weak–strong uniqueness of anisotropic mean curvature flow

Tim Laux<sup>1</sup>, Kerrek Stinson<sup>2</sup> and Clemens Ullrich<sup>3</sup>

<sup>1</sup>University of Regensburg, Regensburg, Germany

<sup>2</sup>University of Bonn, Bonn, Germany

<sup>3</sup>University of Erlangen-Nuremberg, Erlangen, Germany

**Corresponding author:** Kerrek Stinson; Email: [kerrek.stinson@hcm.uni-bonn.de](mailto:kerrek.stinson@hcm.uni-bonn.de)

**Received:** 11 April 2023; **Revised:** 01 April 2024; **Accepted:** 08 April 2024

**Keywords:** anisotropic mean curvature flow; distributional solution; sharp interface limit; weak–strong uniqueness

**2020 Mathematics Subject Classification:** 53E10 (Primary); 35A02, 35D30 (Secondary)

## Abstract

The purpose of this paper is to derive anisotropic mean curvature flow as the limit of the anisotropic Allen–Cahn equation. We rely on distributional solution concepts for both the diffuse and sharp interface models and prove convergence using relative entropy methods, which have recently proven to be a powerful tool in interface evolution problems. With the same relative entropy, we prove a weak–strong uniqueness result, which relies on the construction of gradient flow calibrations for our anisotropic energy functionals.

## 1. Introduction

We consider anisotropic mean curvature flow, a geometric evolution equation used to model microstructure in complex materials. The prototypical application is in multiphase grain growth for polycrystals. As noted in [28], isotropic models, such as mean curvature flow, fail to capture phenomenological features such as the dendritic growth of phases (see also [15]). In chemical kinetics, phase separation experiences anisotropy due to the underlying lattice orientation of the solid host material [5]. Similarly, many materials even display anisotropic surface tensions which are not smooth with respect to the interface orientation (see also [3, 53, 54]): here, one can even consider the household setting of salt (NaCl) and air. At the same time, the study of interface evolutions poses serious numerical and mathematical challenges and a large amount of insight has been gained by modeling such systems in terms of phase-field models, where one replaces interfaces by continuous order parameters (see, e.g., [15]).

In this paper, we prove convergence of solutions of the anisotropic scalar Allen–Cahn equation, a phase-field model, to anisotropic mean curvature flow using variational methods. This may be considered as a first step to proving convergence in the physically relevant vectorial setting. Our approach generally sheds light onto anisotropic mean curvature flow and further enables us to prove a weak–strong uniqueness result for the interface evolution.

Anisotropic mean curvature flow prescribes the evolution of an oriented hypersurface with the surface velocity determined by a weighted mean curvature. Fixing a surface tension  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$  and a mobility  $\mu : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$  (where one can think of extending from the sphere by one-homogeneity), we say that a time-parametrized collection of sets  $\{A(t)\}_{t \in [0, T]}$  evolves by anisotropic mean curvature flow if the boundary  $\Gamma(t) := \partial A(t)$  is smooth and satisfies

$$V = -\mu(v)H_\sigma \quad \text{on} \quad \bigcup_{t \in [0, T]} (\Gamma(t) \times \{t\}), \quad (1.1)$$



where  $\nu$  is the outer unit normal of  $A$ ,  $V$  is the surface normal velocity, and  $H_\sigma := \operatorname{div}_{\Gamma(t)}(D\sigma(\nu))$  is the anisotropic mean curvature with respect to  $\sigma$ . We note that  $D\sigma(\nu)$  is typically referred to as the Cahn–Hoffman vector field, a generalization of the surface normal. In the case that  $\sigma = \mu = |\cdot|$  are given by the Euclidean norm, one recovers the usual (isotropic) mean curvature flow  $V = -H$ .

Following the approach of Luckhaus and Sturzenhecker [47], one way to encode the motion (1.1) is through the characteristic functions  $\chi(t) := \chi_{A(t)}$ . Here  $A(t)$  are naturally given by sets of finite perimeter instead of smooth open sets, allowing for distributional (or  $BV$ ) weak solutions to (1.1). In the isotropic setting, solutions were derived via a minimizing movements scheme (see also [1]) for the perimeter functional. This was a natural approach as mean curvature flow can formally be viewed as the gradient flow of the perimeter functional with an appropriate metric [36, 51].

To carry this analogy to our setting, the curvature flow (1.1) seeks to minimize an anisotropic surface energy  $E : L^2_{\text{loc}}(\Omega) \rightarrow [0, \infty]$  defined by

$$E[u] := \begin{cases} c_0 \int_\Omega \sigma(\nu) |\nabla \chi| & \text{if } u = \chi \in BV_{\text{loc}}(\Omega; \{0, 1\}), \\ \infty & \text{otherwise.} \end{cases} \tag{1.2}$$

Here,  $\nu = -\frac{\nabla \chi}{|\nabla \chi|}$  is the measure-theoretic outer unit normal, and  $c_0$  is a positive constant quantifying surface energy. Formally speaking (see Subsection 2.2), anisotropic mean curvature flow is a gradient flow of the anisotropic perimeter  $E$  with respect to the weighted  $L^2$ -surface metric

$$(V, W)_\Gamma := c_0 \int_\Gamma \frac{1}{\mu(\nu)} V W d\mathcal{H}^{d-1}. \tag{1.3}$$

To construct solutions of interface evolutions, such as (1.1), one can approximate via diffuse-interface models. The idea is that the sharp interface  $\Gamma(t)$ , which captures a jump discontinuity of  $u = \chi$ , is replaced by a diffuse interface  $u_\varepsilon$  forming a continuous transition between values close to 1 and values close to 0. Diffuse-interface or phase-field models are often used in practice and especially for numerics, where tracking of the interface is reduced to a reaction-diffusion equation. Herein, we consider a phase-field approximation of the curvature flow (1.1) given by the anisotropic Allen–Cahn equation

$$\begin{aligned} 2g(-\nabla u_\varepsilon) \partial_t u_\varepsilon &= -\operatorname{div}(Df(-\nabla u_\varepsilon)) - \frac{1}{\varepsilon^2} W'(u_\varepsilon) && \text{in } \Omega \times (0, T), \\ u_\varepsilon(\cdot, 0) &= u_{\varepsilon,0} && \text{in } \Omega, \end{aligned} \tag{1.4}$$

where  $\Omega$  is the  $d$ -dimensional torus, the function  $W : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  is a double-well potential with its wells at 0 and 1, and the anisotropic surface tension and mobility are encoded by the functions

$$f : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}, \quad f(p) = \sigma^2(p) \tag{1.5}$$

and

$$g : \mathbb{R}^d \rightarrow \mathbb{R}_{>0}, \quad g(p) = \frac{\sigma(p) + 1}{\mu(p) + 1}. \tag{1.6}$$

For the isotropic case  $\sigma = \mu = |\cdot|$ , we simply obtain  $f = |\cdot|^2$  and  $g \equiv 1$ .

In contrast to the approach taken in [21], we have introduced a regularization (1.6) of the mobility  $\mu$  allowing us to take advantage of the gradient flow structure for the anisotropic Allen–Cahn equation. We define the anisotropic Cahn–Hilliard energy  $E_\varepsilon : L^2(\Omega) \rightarrow [0, \infty]$  by

$$E_\varepsilon[u] := \begin{cases} \frac{1}{2} \int_\Omega (\varepsilon f(-\nabla u) + \frac{1}{\varepsilon} W(u)) \, dx & \text{if } u \in H^1(\Omega) \text{ and } W \circ u \in L^1(\Omega), \\ \infty & \text{otherwise,} \end{cases} \tag{1.7}$$

and introduce a weighted  $L^2$ -metric given by

$$(v, w)_u := \int_\Omega \varepsilon g(-\nabla u) v w \, dx \tag{1.8}$$

for a given point  $u \in \text{dom}(E_\varepsilon)$ . Note that we omit the dependence on  $\varepsilon$  in the notation  $(\cdot, \cdot)_u$  for convenience. A formal calculation shows that the Allen–Cahn equation (1.4) is equivalent to

$$\partial_t u_\varepsilon(t) = -\nabla_{u_\varepsilon(t)} E_\varepsilon[u_\varepsilon(t)], \quad t \in [0, T],$$

where  $\nabla_{u_\varepsilon(t)}$  is the gradient on  $(L^2(\Omega), (\cdot, \cdot)_{u_\varepsilon(t)})$ . The above equation encapsulates (1.4) as a gradient flow, and this structure will be exploited for the construction of solutions to (1.4) in Section 3 as well as for the sharp-interface limit in Section 4.

A first indication that (1.4) approximates (1.1) is the  $\Gamma$ -convergence of the associated energies (see [10, 19]). With

$$c_0 := \int_0^1 \sqrt{W(s)} ds, \quad (1.9)$$

it was shown by Bouchitté [9] that  $E_\varepsilon \xrightarrow{\Gamma} E$  as  $\varepsilon \searrow 0$  with respect to the strong  $L^1$ -topology on the underlying space. Likewise in the spirit of Luckhaus and Modica [46], Cicalese et al. [18] verified an anisotropic Gibbs–Thomson relation for the energies connecting the first variation of (1.7) to the limiting minimal surface’s anisotropic curvature.

Early results on anisotropic mean curvature flow in the special case  $\mu = \sigma$  are due to Chen, Giga, and Goto [17], who proved the existence and uniqueness of viscosity solutions (up to fattening) for smooth surface tensions  $\sigma$ . Almgren, Taylor, and Wang [1] introduced a time discretization in the form of a minimizing movements scheme including crystalline surface tensions, yielding the so-called flat-flow solutions for anisotropic mean curvature flow. They also proved a short-time existence result for strong solutions if the surface tension  $\sigma$  is smooth. Bellettini and Paolini [7] provide a thorough introduction of the anisotropic mean curvature flow equation for a Finsler metric  $\sigma$ , i.e., with the surface tension possibly depending on the position in space, and argue formally that the time discretization, the level-set equation proposed in [17], and the anisotropic Allen–Cahn equation lead to solutions to anisotropic mean curvature flow.

Allowing for sufficiently regular convex surface tensions  $\sigma$  and arbitrary mobilities  $\mu$ , Elliott and Schätzle [21] proved that, in the sharp-interface limit, solutions to the anisotropic Allen–Cahn equation converge to anisotropic mean curvature flow in the sense of the viscosity formulation. Unlike in the present work, Elliott and Schätzle used a discontinuous, nonregularized version of  $g$  and, therefore, resorted to viscosity solutions of the phase-field equation. In [14], Chambolle and Novaga prove consistency of the minimizing movements scheme [1] and the MBO thresholding scheme for the energy (1.2) with anisotropic mean curvature flow using viscosity solutions. Similarly, using viscosity solutions and distributional solutions with an energy convergence hypothesis, Chambolle et al. [11] prove convergence of the minimizing movements scheme [1] for a translationally dependent energy to inhomogeneous anisotropic mean curvature flow.

In the case of non-smooth (or crystalline) surface tensions, the Cahn–Hoffman vector field is effectively described in terms of a differential inclusion  $v_\sigma \in \partial\sigma(v)$  and selection of the appropriate curvature can make the problem nonlocal. Recently, a variety of work has been invested in understanding crystalline curvature flow. Giga and Giga were among the first to develop a robust solution concept in the planar setting [29]. For crystalline surface tensions, existence and uniqueness was proven in dimension  $d = 3$  by Giga and Požár in [32], and the result was ultimately extended to arbitrary dimension in [33] by the same authors. Chambolle, Morini, and Ponsiglione [13] introduced a novel definition of supersolutions, subsolutions, and weak solutions to anisotropic mean curvature flow that is also based on level set techniques and is particularly useful for crystalline surface tensions. They presented an existence and uniqueness result up to fattening and a comparison principle. The same authors together with Novaga [12] extended this result from the special case  $\mu = \sigma$  to arbitrary mobilities  $\mu$ .

Many of the above results are qualitative, but in the spirit of Chen [16], given the power of viscosity methods, quantitative rates of convergence have been derived and we refer the interested reader to [6, 31] and references therein.

In contrast to the above approaches, we will apply relative entropy methods to identify the limit of (1.4). A related approach regarding a localized energy excess was introduced by the first author with Otto in [41] where they proved convergence of the MBO thresholding scheme to multi-phase mean curvature flow. A nontrivial modification of this idea – based on controlling the tilt excess of the sharp or diffuse interface with regard to a smooth approximation – has been used to prove convergence of the vectorial Allen–Cahn equation to multiphase mean curvature flow [42] and derive an associated rate of convergence [25] along with optimal quantitative convergence rates for the Allen–Cahn equation to mean curvature flow in the two-phase setting [24]. A key feature of viscosity type solutions is the associated comparison principle, which automatically provides uniqueness of solutions up to the issue of fattening. However, in the multiphase case, fundamentally different tools are needed to address uniqueness: The relative entropy method [23] has been used to prove weak–strong uniqueness results for a variety of geometric evolution equations including planar multiphase flows (see, e.g., [22, 35, 36, 40]).

In this paper, we prove convergence of the anisotropic Allen–Cahn equation (1.4) to anisotropic mean curvature flow (1.1) for arbitrary Lipschitz mobilities  $\mu$  and  $C^2$  surface tensions  $\sigma$  under an energy convergence hypothesis, as is often used in application [38, 41, 43, 47]. This result provides a complete proof for the result first announced in [39]. Further, we prove weak–strong uniqueness of the associated distributional solution concept for anisotropic mean curvature flow: If a smooth solution (which we will endow with the structure of a calibrated evolution) and a  $BV$  solution share the same initial data, it follows that both solutions coincide for all times in their common interval of existence. We summarize these results here and refer to Theorems 4.1 and 5.2 for precise details.

**Theorem 1.1.** *Let  $\mu$  be a Lipschitz mobility,  $\sigma$  a  $C^2$  uniformly convex surface tension, and  $\Omega$  the  $d$ -dimensional torus. Then the following holds:*

- *Any sequence of weak solutions  $u_\varepsilon$  of the anisotropic Allen–Cahn equation (1.4) with well-prepared initial conditions has a subsequence converging to some limit  $u = \chi$  with  $\chi : \Omega \times (0, T) \rightarrow \{0, 1\}$  as  $\varepsilon \rightarrow 0$ . Under an energy convergence hypothesis,  $u$  is a weak solution of anisotropic mean curvature flow.*
- *Let  $\sigma$  and  $\mu$  be smooth. If  $\{\mathcal{A}(t)\}_{t \in [0, T]}$  is a strong solution of anisotropic mean curvature flow (1.1) and  $\chi : \Omega \times (0, T) \rightarrow \{0, 1\}$  is a distributional solution of anisotropic mean curvature flow with the same initial condition, i.e.,  $\chi(\cdot, 0) = \chi_{\mathcal{A}(0)}$ , then  $\chi \equiv \chi_{\mathcal{A}}$  for all  $t \in [0, T]$ .*

We remark that convergence of the anisotropic Allen–Cahn equation is well-studied, but as far as the authors are aware, this has exclusively been done from the perspective of viscosity solutions. Our result considers this from the distributional setting and may ultimately be amenable to tackling the multiphase setting that is most relevant to physical applications. Further, our uniqueness result for distributional solutions shows that it may be possible to obtain quantitative convergence rates in the spirit of Fischer et al. [24] for the anisotropic Allen–Cahn equation.

The structure of the paper is as follows: In Section 2, we introduce the admissible class of anisotropies and, based on the closely connected notions of anisotropic surface energy and anisotropic mean curvature, derive the anisotropic mean curvature flow equation as a formal gradient flow. Here, we further introduce our notion of a distributional solution to anisotropic mean curvature flow. As a preparation for the convergence result, Section 3 discusses the anisotropic Allen–Cahn equation (1.4) and establishes the existence – via a time discretization – and regularity of weak solutions. Section 4 is devoted to the proof of the conditional convergence theorem. Finally, Section 5 covers the weak–strong uniqueness theorem.

## Notation

Throughout the paper, let  $d \geq 2$  be the ambient dimension. We consider the equations with periodic boundary conditions, i.e., for the domain we will always choose the flat torus  $\Omega := \mathbb{R}^d / \mathbb{Z}^d$ .

The  $i$ -th unit vector will be denoted by  $e_i$ , and the identity matrix in dimension  $d$  will be written as  $I_d \in \mathbb{R}^{d \times d}$ . For the scalar products of vectors  $a, b \in \mathbb{R}^d$  and of matrices  $A, B \in \mathbb{R}^{d \times d}$ , we write  $a \cdot b := \sum_{i=1}^d a_i b_i$  and  $A : B := \sum_{i,j=1}^d (A)_{ij} (B)_{ij}$ , respectively. The tensor product of  $a, b \in \mathbb{R}^d$  is a matrix  $a \otimes b \in \mathbb{R}^{d \times d}$ , which is given by  $(a \otimes b)_{ij} = a_i b_j$ .

The symbol  $\nabla$  is reserved for derivatives with respect to the space variable  $x \in \Omega$ . For a vector field  $X : \Omega \rightarrow \mathbb{R}^d$ , the derivative  $\nabla X$  is a  $(d \times d)$ -matrix defined by  $(\nabla X)_{ij} = \partial_j X_i$ . Derivatives with respect to variables  $p \in \mathbb{R}^d$  will be denoted by  $D = D_p$ . In contrast to  $\nabla$ ,  $D\sigma(v)$  will denote the column vector.

For a set  $A \subset \mathbb{R}^d$ ,  $\chi_A$  is the characteristic function taking the value 1 on  $A$  and 0 in the complement.

By a standard abuse of notation, we refer to the distributional derivative in space of a function  $\chi \in BV(\Omega; \{0, 1\})$  by  $\nabla \chi$ . In the case that  $\chi \in L^1((0, T); BV(\Omega))$ , we write  $\nabla \chi := \nabla \chi(\cdot, t) \otimes \mathcal{L}_{(0,T)}^1$  in the notation of Young measures.

The divergence of a vector field  $X : \Omega \rightarrow \mathbb{R}^d$  is  $\operatorname{div} X = \sum_{i=1}^d \partial_i X_i$ . The divergence of a matrix field  $A : \Omega \rightarrow \mathbb{R}^{d \times d}$  is vector-valued,  $(\operatorname{div} A)_i = \sum_{j=1}^d \partial_j (A)_{ij}$ .

For a vector  $a \in \mathbb{R}^d$  we define the differential operator  $a \cdot \nabla := \sum_{i=1}^d a_i \partial_i$ . In particular, for a vector field  $X$ , we have  $(a \cdot \nabla)X = (\nabla X)a$ .

## 2. Anisotropies and anisotropic mean curvature

In this section, we develop the necessary mathematical preliminaries for the rest of the paper. We introduce the notion of admissible anisotropies in Subsection 2.1 and provide a calculation clarifying the formal view of anisotropic mean curvature flow as gradient flow in Subsection 2.2. Finally, we show in the smooth setting how the anisotropic curvature can be reinterpreted via integration by parts, allowing us to introduce a distributional solution for anisotropic mean curvature flow in Subsection 2.3.

### 2.1 Admissible anisotropies

In this subsection, we introduce a suitable class of anisotropic surface tensions and mobilities and discuss some elementary facts about admissible anisotropic surface tensions.

**Definition 2.1.** An admissible pair of anisotropies  $(\sigma, \mu)$  is comprised of an admissible surface tension  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$  that satisfies

- (S1) positive 1-homogeneity, i.e.,  $\sigma(\lambda p) = \lambda \sigma(p)$  for all  $p \in \mathbb{R}^d, \lambda > 0$ ,
- (S2) positive definiteness, i.e.,  $\sigma(p) = 0$  if and only if  $p = 0$ ,
- (S3) smoothness, i.e.,  $\sigma \in C^2(\mathbb{R}^d \setminus \{0\})$ ,
- (S4) uniform convexity, i.e.,  $\sigma^2 : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$  is strongly convex in the sense that there exists a constant  $c > 0$  satisfying

$$\sigma^2((1-t)p + tq) \leq (1-t)\sigma^2(p) + t\sigma^2(q) - c \frac{t(1-t)}{2} |p - q|^2$$

for all  $p, q \in \mathbb{R}^d, t \in (0, 1)$ ,

and an admissible mobility  $\mu : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$  that satisfies

- (M1) positive 1-homogeneity, i.e.,  $\mu(\lambda p) = \lambda \mu(p)$  for all  $p \in \mathbb{R}^d, \lambda > 0$ ,
- (M2) positive definiteness, i.e.,  $\mu(p) = 0$  if and only if  $p = 0$ ,
- (M3) regularity, i.e.,  $\mu \in C^{0,1}(\mathbb{R}^d)$ .

The polar norm of  $\sigma$  is

$$\sigma^\circ : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}, \quad \sigma^\circ(q) = \sup_{p: \sigma(p) \leq 1} p \cdot q. \tag{2.1}$$

**Example 2.2.** A useful class of admissible surface tensions, which is suitable to approximate arbitrary norms, is introduced by Barrett, Garcke, and Nürnberg in [4, (1.12)] as follows: Let  $q \in [1, \infty)$ ,  $L \in \mathbb{N}$ , and let  $G_1, G_2, \dots, G_L \in \mathbb{R}^{d \times d}$  be symmetric and positive definite. We define

$$\sigma(p) := \left( \sum_{l=1}^L \sigma_l(p)^q \right)^{\frac{1}{q}}, \quad \sigma_l(p) := \sqrt{p \cdot G_l p}.$$

One can show that  $\sigma$  is an admissible surface tension in the above sense (see also [30, Remark 1.7.5] and [7, Example 4.6]).

For convenience, let us collect several useful properties of surface tensions.

**Lemma 2.3.** Let  $\sigma$  be an admissible surface tension as in Definition 2.1. Then

- (i)  $\sigma^\circ$  is an admissible surface tension and  $\sigma^{\circ\circ} = \sigma$ ,
- (ii)  $p \cdot q \leq \sigma(p)\sigma^\circ(q)$  for all  $p, q \in \mathbb{R}^d$ ,
- (iii) we have  $0 < \min_{S^{d-1}} \sigma \leq \max_{S^{d-1}} \sigma < \infty$ , and

$$(\min_{S^{d-1}} \sigma)|p| \leq \sigma(p) \leq (\max_{S^{d-1}} \sigma)|p| \quad \text{for all } p \in \mathbb{R}^d,$$

- (iv)  $D\sigma$  is positively 0-homogeneous and  $D^2\sigma$  is positively  $(-1)$ -homogeneous, i.e.,

$$D\sigma(\lambda p) = D\sigma(p) \quad \text{and} \quad D^2\sigma(\lambda p) = \frac{1}{\lambda} D^2\sigma(p) \quad \text{for all } \lambda > 0, p \in \mathbb{R}^d \setminus \{0\},$$

- (v)  $p \cdot D\sigma(p) = \sigma(p)$  for all  $p \in \mathbb{R}^d \setminus \{0\}$ ,
- (vi)  $\sigma^\circ(D\sigma(p)) = 1$  for all  $p \in \mathbb{R}^d \setminus \{0\}$ ,
- (vii)  $\sigma^\circ(q)D\sigma(D\sigma^\circ(q)) = q$  for all  $q \in \mathbb{R}^d \setminus \{0\}$ .

**Proof.** (ii) can be shown directly from the definition of  $\sigma^\circ$  and the positive 1-homogeneity of  $\sigma$ . (iii) and (iv) follow from the positive 1-homogeneity and continuity resp. smoothness of  $\sigma$ . (v) is a result of (iv) and the fundamental theorem of calculus, cf. [30, Section 1.7.2]:

$$\sigma(p) = \int_0^1 \frac{d}{d\lambda} \sigma(\lambda p) d\lambda = \int_0^1 D\sigma(\lambda p) \cdot p d\lambda = D\sigma(p) \cdot p.$$

Giga [30, Section 1.7.2] also provides the proofs for (vi), (vii), and (i) except the uniform convexity of  $\sigma^\circ$  and the duality statement  $\sigma^{\circ\circ} = \sigma$ . A proof for  $\sigma^{\circ\circ} = \sigma$  can be found in [50, Theorem 15.1 and Corollary 15.1.1]. The uniform convexity of  $\sigma^\circ$  can be proved in a similar way to [45, Proposition 1.e.2]. □

A useful trick with regard to the Euclidean metric  $|\cdot|$  is to control quadratic errors for unit vectors via the inequality

$$|p - p'|^2 = 1 + |p'|^2 - 2p \cdot p' \leq 2(1 - p \cdot p') \quad \text{whenever } |p| = 1 \text{ and } |p'| \leq 1,$$

with equality if and only if  $|p'| = 1$ . In order to introduce a tilt excess functional for anisotropic mean curvature flow, we are interested in an anisotropic counterpart to the above inequality.

The suitable anisotropic inequality bounds squared distances  $|p - p'|^2$  by a term of the form  $\sigma(p) - |p'|D\sigma(p') \cdot p$ . However, since the mapping  $p' \mapsto |p'|D\sigma(p')$  is, in general, not continuously differentiable at  $p' = 0$ , we will use a truncated version instead. To this end, we fix a cutoff function  $\psi \in C^\infty([0, \infty))$  satisfying

$$\psi(r) = 0 \quad \text{for all } r \leq \frac{1}{4}, \quad \psi(r) = 1 \quad \text{for all } r \geq \frac{1}{2}, \quad \text{and } \psi' \geq 0. \tag{2.2}$$

The following lemma contains two estimates featuring the truncated version.

**Lemma 2.4.** *Let  $\sigma$  be an admissible surface tension. There exist constants  $c_\sigma, C_\sigma > 0$  depending only on  $\sigma$  such that*

(i) *we have*

$$\sigma(p) - |p'| \psi(|p'|) D\sigma(p') \cdot p \geq c_\sigma |p - p'|^2 \tag{2.3}$$

*for all  $p, p' \in \mathbb{R}^d$  such that  $|p| = 1$  and  $|p'| \leq 1$ , and*

(ii) *we have*

$$\sigma(p) - |p'| \psi(|p'|) D\sigma(p') \cdot p \leq C_\sigma \left( |p - p'|^2 + (1 - |p'|) \right) \tag{2.4}$$

*for all  $p, p' \in \mathbb{R}^d$  such that  $|p| = 1$  and  $|p'| \leq 1$ .*

**Proof.** Variants of (i) and (ii) were provided by Dziuk [20, Proposition 2.2] and Laux [39, Lemma 3.3 and Section 4.2], respectively. The proof is included here for the reader’s convenience.

For inequality (i), we consider two cases with respect to  $p'$ : First, if  $D\sigma(p') \cdot p < \frac{\sigma(p)}{2}$  or  $p' = 0$ , we use the estimate  $|p - p'|^2 \leq 4$  to obtain

$$\begin{aligned} \sigma(p) - |p'| \psi(|p'|) D\sigma(p') \cdot p &\geq \sigma(p) \left( 1 - \frac{1}{2} |p'| \psi(|p'|) \right) \\ &\geq \frac{1}{2} \sigma(p) \\ &\geq \frac{1}{8} (\min_{S^{d-1}} \sigma) |p - p'|^2. \end{aligned} \tag{2.5}$$

Second, if  $p' \neq 0$  and  $D\sigma(p') \cdot p \geq \frac{\sigma(p)}{2}$ , we have  $\left| p + \frac{p'}{|p'|} \right| \geq \frac{\min_{S^{d-1}} \sigma}{\max_{S^{d-1}} |D\sigma|} > 0$  since

$$\begin{aligned} 0 \leq D\sigma(p') \cdot p \leq D\sigma(p') \cdot \left( -\frac{p'}{|p'|} \right) + (\max_{S^{d-1}} |D\sigma|) \left| p + \frac{p'}{|p'|} \right| \\ \leq -(\min_{S^{d-1}} \sigma) + (\max_{S^{d-1}} |D\sigma|) \left| p + \frac{p'}{|p'|} \right|. \end{aligned}$$

The uniform convexity assumption (S4) can equivalently be stated as follows (see [30, Remark 1.7.5]): There exists a constant  $\underline{\sigma} > 0$  such that

$$D^2\sigma(p^*) \geq \frac{\underline{\sigma}}{|p^*|} \quad \text{on } \{p^*\}^\perp \tag{2.6}$$

for all  $p^* \neq 0$ . We use a second-order Taylor expansion around  $\frac{p'}{|p'|}$  and write the remainder in terms of an intermediate point  $p^* = (1 - t)p + t \frac{p'}{|p'|}$  with  $t \in [0, 1]$ . Together with the convexity property (2.6), it follows that

$$\begin{aligned} \sigma(p) - D\sigma(p') \cdot p &= \sigma(p) - \sigma\left(\frac{p'}{|p'|}\right) - D\sigma\left(\frac{p'}{|p'|}\right) \cdot \left(p - \frac{p'}{|p'|}\right) \\ &= \frac{1}{2} \left(p - \frac{p'}{|p'|}\right) \cdot D^2\sigma(p^*) \left(p - \frac{p'}{|p'|}\right) \\ &\geq \frac{\underline{\sigma}}{2|p^*|} \left| \left(p - \frac{p'}{|p'|}\right) - \left(\left(p - \frac{p'}{|p'|}\right) \cdot \frac{p^*}{|p^*|}\right) \frac{p^*}{|p^*|} \right|^2 \end{aligned}$$

$$\begin{aligned} &\geq \frac{\sigma}{8} \left| p + \frac{p'}{|p'|} \right|^2 \left| p - \frac{p'}{|p'|} \right|^2 \\ &\geq \frac{\sigma}{8} \frac{(\min_{S^{d-1}} \sigma)^2}{(\max_{S^{d-1}} |D\sigma|)^2} \left| p - \frac{p'}{|p'|} \right|^2. \end{aligned} \tag{2.7}$$

Furthermore, the assumption  $D\sigma(p') \cdot p \geq \frac{\sigma(p)}{2}$  allows us to compute

$$(1 - |p'|\psi(|p'|)) D\sigma(p') \cdot p \geq (1 - |p'|) \frac{\sigma(p)}{2} \geq \frac{\min_{S^{d-1}} \sigma}{2} \left| \frac{p'}{|p'|} - p' \right|^2. \tag{2.8}$$

Finally, a combination of (2.7) and (2.8) together with an application of Young’s inequality yields

$$\begin{aligned} \sigma(p) - |p'|\psi(|p'|)D\sigma(p') \cdot p &= \sigma(p) - D\sigma(p') \cdot p + (1 - |p'|\psi(|p'|)) D\sigma(p') \cdot p \\ &\geq 2c_\sigma \left( \left| p - \frac{p'}{|p'|} \right|^2 + \left| \frac{p'}{|p'|} - p' \right|^2 \right) \\ &\geq c_\sigma |p - p'|^2, \end{aligned} \tag{2.9}$$

which completes the proof of (i) in the second case.

For inequality (ii), we distinguish two cases again. First, in the easier case  $|p - p'| \geq 1$ , we can estimate

$$\sigma(p) - |p'|\psi(|p'|)D\sigma(p') \cdot p \leq \max_{S^{d-1}} \sigma + \max_{S^{d-1}} |D\sigma| \leq \left( \max_{S^{d-1}} \sigma + \max_{S^{d-1}} |D\sigma| \right) |p - p'|^2. \tag{2.10}$$

Second, if  $|p - p'| < 1$ , then

$$\left| p - \frac{p'}{|p'|} \right|^2 = 2 - 2p \cdot \frac{p'}{|p'|} \leq 2 + \frac{1}{|p'|} (|p - p'|^2 - 1) < 2$$

and, therefore,

$$\left| (1 - t)p + t \frac{p'}{|p'|} \right|^2 = 1 - (t - t^2) \left| p - \frac{p'}{|p'|} \right|^2 \geq 1 - \frac{1}{4} \left| p - \frac{p'}{|p'|} \right|^2 > \frac{1}{2}$$

for all  $t \in [0, 1]$ . As in the proof of inequality (i) above, we introduce a second-order Taylor expansion around  $\frac{p'}{|p'|}$  with an intermediate point  $p^* = (1 - t)p + t \frac{p'}{|p'|}$ , where  $t \in [0, 1]$ . Using this expansion, Lemma 2.3(ii), (vi), and Young’s inequality, we obtain

$$\begin{aligned} \sigma(p) - |p'|\psi(|p'|)D\sigma(p') \cdot p &= \sigma(p) - \sigma\left(\frac{p'}{|p'|}\right) - D\sigma\left(\frac{p'}{|p'|}\right) \cdot \left(p - \frac{p'}{|p'|}\right) \\ &\quad + (1 - |p'|\psi(|p'|)) D\sigma(p') \cdot p \\ &\leq \frac{1}{2} \left(p - \frac{p'}{|p'|}\right) \cdot D^2\sigma(p^*) \left(p - \frac{p'}{|p'|}\right) + (1 - |p'|\psi(|p'|)) \sigma(p) \\ &\leq \frac{\sqrt{2}}{2} (\max_{S^{d-1}} |D^2\sigma|) \left| p - \frac{p'}{|p'|} \right|^2 + 2(\max_{S^{d-1}} \sigma) (1 - |p'|) \\ &\leq \sqrt{2} (\max_{S^{d-1}} |D^2\sigma|) |p - p'|^2 \\ &\quad + \left( \sqrt{2} (\max_{S^{d-1}} |D^2\sigma|) + 2(\max_{S^{d-1}} \sigma) \right) (1 - |p'|), \end{aligned} \tag{2.11}$$

from which (ii) follows in the case  $|p - p'| < 1$ . □

The following fact on the duality of  $\sigma$  and  $\sigma^\circ$  will be used at a later point.



**Lemma 2.5.** *Let  $B \in L^1(\Omega)^d$ , and let  $\sigma$  be an admissible surface tension. Then*

$$\int_{\Omega} \sigma(B) dx = \sup_{\eta} \int_{\Omega} B \cdot \eta \, dx,$$

where the supremum is taken over all  $\eta \in C^1(\Omega)^d$  such that  $\sigma^\circ(\eta) \leq 1$ .

**Proof.** For  $\eta \in C^1(\Omega)^d$  such that  $\sigma^\circ(\eta) \leq 1$ , we have that

$$\int_{\Omega} \sigma(B) \, dx \geq \int_{\Omega} \sigma(B)\sigma^\circ(\eta) \, dx \geq \int_{\Omega} B \cdot \eta \, dx$$

by Lemma 2.3. To see the other bound, whenever  $B \neq 0$  note that  $\sigma(B) = B \cdot D\sigma(B)$  and that  $\sigma^\circ(D\sigma(B)) = 1$  by Lemma 2.3. As  $\{\nu : \sigma^\circ(\nu) \leq 1\}$  is a convex set, if  $\eta$  is given by a smooth mollification of  $D\sigma(B)$ , the inequality  $\sigma^\circ(\eta) \leq 1$  still holds. Approximating  $D\sigma(B)$  by such  $\eta$  shows that  $\int_{\Omega} \sigma(B) dx \leq \sup_{\eta} \int_{\Omega} B \cdot \eta \, dx$ , concluding the proof.  $\square$

### 2.2 Anisotropic mean curvature and surface energy

In the remainder of this section, we assume that  $(\sigma, \mu)$  is an admissible pair of anisotropies according to Definition 2.1.

Similarly to [7], we introduce the  $\sigma$ -mean curvature of a hypersurface  $\Gamma$  as the surface divergence of the Cahn–Hoffman vector of the outer unit normal  $\nu$ , i.e.,  $H_\sigma = \operatorname{div}_\Gamma(D\sigma(\nu))$ . For the notion of the Cahn–Hoffman vector, see [30, Section 1.3]. The surface divergence is given by  $\operatorname{div}_\Gamma X := \operatorname{div} X - \nu \cdot \nabla X \nu$ . Observe that, for every  $C^1$ -extension of the normal  $\nu$  to the whole space and every  $x \in \Gamma$ , we have

$$\begin{aligned} H_\sigma(x) &= \operatorname{div}_\Gamma(D\sigma(\nu))(x) \\ &= \operatorname{div}(D\sigma(\nu))(x) - \nu(x) \cdot \nabla(D\sigma(\nu))(x)\nu(x) \\ &= \operatorname{div}(D\sigma(\nu))(x) - \nu(x) \cdot D^2\sigma(\nu(x))(\nabla\nu(x))\nu(x) \\ &= \operatorname{div}(D\sigma(\nu))(x), \end{aligned}$$

where the last step uses the fact that  $D^2\sigma(p)p = \frac{d}{ds} \Big|_{s=0} D\sigma(e^s p) = \frac{d}{ds} \Big|_{s=0} D\sigma(p) = 0$  for all  $p \in \mathbb{R}^d \setminus \{0\}$  by the positive 0-homogeneity of  $D\sigma$ , Lemma 2.3(iv).

Furthermore, we define the anisotropic surface energy  $E$  as in (1.2). The relation between the anisotropic mean curvature and anisotropic surface energy becomes clear from the following theorem, which deals with the first variation and direction of steepest descent of the functional  $E$ :

**Theorem 2.6.** *Let  $A \subset \mathbb{R}^d$  be a bounded open set with  $C^2$ -boundary, and let  $\nu$  denote the outer unit normal on  $\partial A$ . Given a compactly supported vector field  $B \in C_c^1(\mathbb{R}^d)^d$ , we define a one-parameter family of diffeomorphisms*

$$\Psi : \mathbb{R}^d \times (-\|\nabla B\|_{L^\infty}^{-1}, \|\nabla B\|_{L^\infty}^{-1}) \rightarrow \mathbb{R}^d, \quad \Psi_t(x) = x + tB(x).$$

For  $|t| < \|\nabla B\|_{L^\infty}^{-1}$ , let  $A(t) := \Psi_t(A)$  and  $u(x, t) := \chi_{\Omega(t)}(x)$ . Then

$$\frac{d}{dt} E[u(\cdot, t)] \Big|_{t=0} = c_0 \int_{\partial A} H_\sigma \nu \cdot B d\mathcal{H}^{d-1}. \tag{2.12}$$

Moreover, we define  $\|B\|_* := \left( c_0 \int_{\partial A} \sigma^\circ(-\operatorname{sgn}(H_\sigma)B)^2 \frac{\sigma(v)^2}{\mu(v)} d\mathcal{H}^{d-1} \right)^{\frac{1}{2}}$ , so that  $\|\cdot\|_*$  is bounded from below and above by the  $L^2$ -norm for vector-valued functions  $\partial A \rightarrow \mathbb{R}^d$ . If we denote the right-hand side of (2.12) by the dual pairing  $(dE, B)$ , then a solution of the minimization problem

$$\min \{ (dE, B) \mid \|B\|_* \leq 1 \} \tag{2.13}$$

is given by a positive multiple of  $-H_\sigma \frac{\mu(v)}{\sigma(v)} D\sigma(\nu)$ .

**Proof.** The identity for the first variation (2.12) was stated by Almgren, Taylor, and Wang [1, Section 2.2] and proved by Bellettini and Paolini in the more general context of Finsler geometry [7, Theorem 5.1]. The proof of the steepest descent statement follows the idea of [7, Proposition 5.1] but incorporates the (arbitrary) mobility  $\mu$ :

Let  $B \in L^2(\partial A)^d$ , then by Lemma 2.3(ii) and the Cauchy–Schwarz inequality it follows that

$$\begin{aligned} (dE, B) &= c_0 \int_{\partial A} H_\sigma \nu \cdot B d\mathcal{H}^{d-1} \\ &= -c_0 \int_{\partial A} |H_\sigma| \nu \cdot (-\operatorname{sgn}(H_\sigma)B) d\mathcal{H}^{d-1} \\ &\geq -c_0 \int_{\partial A} |H_\sigma| \sigma(\nu) \sigma^\circ(-\operatorname{sgn}(H_\sigma)B) d\mathcal{H}^{d-1} \\ &\geq -c_0 \left( \int_{\partial A} |H_\sigma|^2 \mu(\nu) d\mathcal{H}^{d-1} \right)^{\frac{1}{2}} \left( \int_{\partial A} \sigma^\circ(-\operatorname{sgn}(H_\sigma)B)^2 \frac{\sigma(\nu)^2}{\mu(\nu)} d\mathcal{H}^{d-1} \right)^{\frac{1}{2}} \\ &= -\sqrt{c_0} \left( \int_{\partial A} |H_\sigma|^2 \mu(\nu) d\mathcal{H}^{d-1} \right)^{\frac{1}{2}} \|B\|_*, \end{aligned}$$

and Lemma 2.3(v), (vi) yield that the first inequality is an equality if  $B = -\lambda \operatorname{sgn}(H_\sigma) D\sigma(\nu)$  for some nonnegative function  $\lambda : \partial A \rightarrow \mathbb{R}_{\geq 0}$ . Therefore, if  $B$  is a constant positive multiple of  $-H_\sigma \frac{\mu(\nu)}{\sigma(\nu)} D\sigma(\nu)$ , then both inequalities are equalities. The statement follows by the positive 1-homogeneity of  $\|\cdot\|_*$ .  $\square$

This theorem helps us to justify the gradient flow structure for anisotropic mean curvature flow as introduced earlier via (1.2) and (1.3): Considering that tangential components of the velocity  $B$  do not contribute to the variation of  $A$ , let us restrict ourselves to velocities of the form  $B = -\lambda \operatorname{sgn}(H_\sigma) D\sigma(\nu)$ , where  $\lambda \in C(\partial A; \mathbb{R}_{\geq 0})$ . The normal component of such a vector field  $B$  is  $V = B \cdot \nu = -\lambda \operatorname{sgn}(H_\sigma) \sigma(\nu)$ , and the metric term becomes

$$\|B\|_*^2 = c_0 \int_{\partial A} \sigma^\circ(\lambda D\sigma(\nu))^2 \frac{\sigma(\nu)^2}{\mu(\nu)} d\mathcal{H}^{d-1} = c_0 \int_{\partial A} \lambda^2 \frac{\sigma(\nu)^2}{\mu(\nu)} d\mathcal{H}^{d-1} = c_0 \int_{\partial A} \frac{1}{\mu(\nu)} V^2 d\mathcal{H}^{d-1},$$

which is in accordance with (1.3). The velocity of steepest descent which was given in Theorem 2.6 satisfies

$$V = -H_\sigma \frac{\mu(\nu)}{\sigma(\nu)} D\sigma(\nu) \cdot \nu = -\mu(\nu) H_\sigma$$

by Lemma 2.3(v), which (formally) verifies the gradient flow structure of (1.1).

### 2.3 Distributional solutions to anisotropic mean curvature flow

We introduce a distributional formulation for (1.1) that was proposed in [39]. The idea behind this formulation is to encode the  $\sigma$ -mean curvature by a  $(d \times d)$ -matrix via an integration by parts as follows:

**Remark 2.7.** Let  $A \subset \mathbb{R}^d$  be a bounded open set with  $C^2$ -boundary. For every vector field  $B \in C^1(\mathbb{R}^d)^d$ , we have

$$\int_{\partial A} (\sigma(\nu) I_d - \nu \otimes D\sigma(\nu)) : \nabla B d\mathcal{H}^{d-1} = \int_{\partial A} H_\sigma B \cdot \nu d\mathcal{H}^{d-1}. \tag{2.14}$$

To see this, we extend the normal  $\nu$  to a vector field  $\nu \in C^1(\mathbb{R}^d)^d$ , e.g., as a truncation of the normalized gradient  $\frac{\nabla \operatorname{sdist}}{|\nabla \operatorname{sdist}|}$  of the signed distance function. There exists an open neighborhood  $U$  of  $\partial A$  where  $\nu \neq 0$ ,

and in this neighborhood we can decompose the vector field  $B$  as

$$B = \left( B - \frac{1}{\sigma(v)}(B \cdot v)D\sigma(v) \right) + \frac{1}{\sigma(v)}(B \cdot v)D\sigma(v) \\ =: B_1 + B_2.$$

The restriction of  $B_1$  to  $\partial A$  is tangential since, by Lemma 2.3(v), we have  $B_1 \cdot \nu = 0$ . By the divergence theorem on the closed surface  $\partial A$ , we obtain

$$0 = \int_{\partial A} \operatorname{div}_{\partial A} (\sigma(v)B_1) \, d\mathcal{H}^{d-1} \\ = \int_{\partial A} (\sigma(v)\operatorname{div}B_1 + D\sigma(v) \cdot \nabla \nu B_1 - \sigma(v)\nu \cdot \nabla B_1 \nu) \, d\mathcal{H}^{d-1} \\ = \int_{\partial A} (\sigma(v)\operatorname{div}B_1 - \nu \cdot \nabla B_1 D\sigma(v) + (D\sigma(v) - \sigma(v)\nu) \cdot \nabla \nu B_1 + \\ + \nu \cdot \nabla B_1 (D\sigma(v) - \sigma(v)\nu)) \, d\mathcal{H}^{d-1} \\ = \int_{\partial A} ((\sigma(v)I_d - \nu \otimes D\sigma(v)) : \nabla B_1 + (D\sigma(v) - \sigma(v)\nu) \cdot \nabla \nu B_1 + \\ - B_1 \cdot \nabla \nu (D\sigma(v) - \sigma(v)\nu)) \, d\mathcal{H}^{d-1} \\ = \int_{\partial A} (\sigma(v)I_d - \nu \otimes D\sigma(v)) : \nabla B_1 \, d\mathcal{H}^{d-1}. \tag{2.15}$$

In this computation, the third equality follows by adding zero and using that  $\nu \cdot \nabla \nu B_1 = B_1 \cdot \nabla (\frac{1}{2}|\nu|^2) = 0$ . For the fourth equality observe that  $(\nabla B_1)^T \nu + B_1)^T \nu + (\nabla \nu)^T B_1 = \nabla(B_1 \cdot \nu) = 0$  since  $B_1$  is tangential. holds true because the second fundamental form

$$\mathbb{I}_x : T_x \partial A \times T_x \partial A \rightarrow \mathbb{R}, \quad \mathbb{I}_x(v, w) := w \cdot (\nu \cdot \nabla) \nu(x)$$

is symmetric for all  $x \in \partial A$  (see [30, Section 1.3]).

On the other hand, we have

$$\nabla B_2 = -\frac{1}{\sigma^2(v)}(B \cdot v)D\sigma(v) \otimes (\nabla \nu)^T D\sigma(v) + \frac{1}{\sigma(v)}D\sigma(v) \otimes (\nabla \nu)^T B \\ + \frac{1}{\sigma(v)}D\sigma(v) \otimes (\nabla B)^T \nu + \frac{1}{\sigma(v)}(B \cdot v)D^2\sigma(v)\nabla \nu,$$

which yields

$$(\sigma(v)I_d - \nu \otimes D\sigma(v)) : \nabla B_2 = \operatorname{div}(D\sigma(v))B \cdot \nu = H_\sigma B \cdot \nu. \tag{2.16}$$

From (2.15) and (2.16) we obtain (2.14).

Using the integration by parts (2.14), we arrive at a distributional formulation for anisotropic mean curvature which resembles the isotropic version due to Luckhaus and Sturzenhecker [47]. Following [39], our definition for  $BV$  solutions to anisotropic mean curvature flow also includes an optimal energy dissipation inequality, which alludes to the gradient flow structure of the problem.

**Definition 2.8** (*Distributional solutions to anisotropic mean curvature flow*). Let  $\chi_0 \in BV(\Omega; \{0, 1\})$ . A distributional (or  $BV$ ) solution to  $V = -\mu(v)H_\sigma$  with initial condition  $\chi_0$  is a function  $\chi \in BV(\Omega \times (0, T); \{0, 1\}) \cap C^{0, \frac{1}{2}}([0, T]; L^1(\Omega)) \cap L^\infty(0, T; BV(\Omega))$  such that

(i) there exists a  $|\nabla\chi|$ -measurable function  $V : \Omega \times (0, T) \rightarrow \mathbb{R}$  such that

$$\int_0^T \int_{\Omega} V^2 |\nabla\chi| < \infty, \tag{2.17}$$

which is the normal velocity in the sense that for all  $\zeta \in C^1(\Omega \times [0, T])$  and  $T' \in (0, T)$ :

$$\begin{aligned} \int_{\Omega} \zeta(x, T') \chi(x, T') dx - \int_{\Omega} \zeta(x, 0) \chi_0(x) dx \\ = \int_0^{T'} \int_{\Omega} \partial_t \zeta(x, t) \chi(x, t) dx dt + \int_0^{T'} \int_{\Omega} \zeta(x, t) V(x, t) |\nabla\chi|, \end{aligned} \tag{2.18}$$

(ii) the relation  $V = -\mu(v)H_{\sigma}$  is satisfied in a distributional sense, i.e., for all  $B \in C^1(\Omega \times [0, T])^d$ , we have

$$-\int_0^T \int_{\Omega} \frac{1}{\mu(v)} VB \cdot v |\nabla\chi| = \int_0^T \int_{\Omega} \nabla B : (\sigma(v)I_d - v \otimes D\sigma(v)) |\nabla\chi|, \tag{2.19}$$

and

(iii) the function  $\chi$  satisfies the optimal energy dissipation inequality

$$\int_{\Omega} \sigma(v) |\nabla\chi(T')| + \int_0^{T'} \int_{\Omega} \frac{1}{\mu(v)} V^2 |\nabla\chi| \leq \int_{\Omega} \sigma(v) |\nabla\chi_0| \tag{2.20}$$

for every  $T' \in [0, T]$ .

A prime example of anisotropic mean curvature flow is the following self-similar solution, which generalizes the evolution of shrinking spheres by (isotropic) mean curvature flow. This example of motion by anisotropic mean curvature flow is derived in [30, Section 1.7.2].

**Example 2.9** (The Wulff shape). *Let  $\sigma$  be an admissible surface tension, and suppose that  $\mu = \sigma$ . The Wulff shape associated with  $\sigma$  is the set*

$$\mathcal{W}_{\sigma} := \left\{ x \in \mathbb{R}^d \mid \sigma^{\circ}(x) \leq 1 \right\}. \tag{2.21}$$

Then

- the  $\sigma$ -mean curvature of  $\partial\mathcal{W}_{\sigma}$  is  $H_{\sigma} \equiv d - 1$ , and
- the family  $\{\Gamma(t)\}_{t \in [0, \frac{1}{2(d-1)}}$  defined by  $\Gamma(t) = \sqrt{1 - 2(d-1)t}$   $\partial\mathcal{W}_{\sigma}$  evolves by anisotropic mean curvature flow  $V = -\sigma(v)H_{\sigma}$ .

In Theorem 4.1, we will prove convergence of the anisotropic Allen–Cahn equation (1.4) to a distributional solution of anisotropic mean curvature flow in the sense of Definition 2.8, but first we must introduce an appropriate notion of solution for (1.4).

### 3. The anisotropic Allen–Cahn equation

The goal of this chapter is to construct solutions to the anisotropic Allen–Cahn equation (1.4) and to establish spatial  $H^2$ -regularity for these solutions.

From now on, our assumptions on the double-well potential  $W$  are

- (W1)  $W(0) = W(1) = 0$  and  $W(s) > 0$  for all  $s \in \mathbb{R} \setminus \{0, 1\}$ ,
- (W2)  $W \in C^1(\mathbb{R})$ ,
- (W3)  $W$  is twice differentiable at 0 and 1, with  $W''(0), W''(1) > 0$ ,

- (W4) there exists  $\lambda > 0$  such that  $W$  is  $\lambda$ -convex (see Remark 3.1), and
- (W5)  $W$  is nonincreasing on  $(-\infty, 0]$  and nondecreasing on  $[1, \infty)$ .

While (W1) and (W2) are common assumptions on double-well potentials, (W3) is motivated by the desirable property of exponential convergence of solutions to (1.4) to the wells far away from the diffuse interface and is not necessary for the main results of existence and sharp-interface limit. This assumption was used in a similar way by Sternberg [52]. A reference which makes use of the exponential convergence is [24], where Fischer, Simon, and the first author derive a convergence rate for the sharp-interface limit in the isotropic case.

Assumption (W4) is needed in the proof of existence of solutions to guarantee the convergence of the approximate energies. Assumption (W5) is dispensable if one only considers solutions  $u_\varepsilon$  to (1.4) that satisfy  $0 \leq u_\varepsilon \leq 1$  almost everywhere.

Up to a linear scaling, a possible choice for  $W$  is the standard double-well potential  $W(s) = \frac{9}{16} (1 - s^2)^2$ . Here, the prefactor  $\frac{9}{16}$  is chosen such that  $c_0 := \int_{-1}^1 \sqrt{W(s)} ds = 1$ .

**Remark 3.1.** Let  $\mathcal{X}$  be a Hilbert space and  $\lambda > 0$ . A function  $F : \mathcal{X} \rightarrow [0, \infty]$  is called  $\lambda$ -convex if the following equivalent conditions are satisfied:

- (i) For all  $x_0 \in \mathcal{X}$ , the mapping  $x \mapsto F(x) + \frac{\lambda}{2} \|x - x_0\|_{\mathcal{X}}^2$  is convex,
- (ii) for all  $x, y \in \mathcal{X}$  and  $\mu \in (0, 1)$ , we have

$$F((1 - \mu)x + \mu y) \leq (1 - \mu)F(x) + \mu F(y) + \frac{\lambda}{2} \mu(1 - \mu) \|y - x\|_{\mathcal{X}}^2.$$

Furthermore, we will always assume that  $(\sigma, \mu)$  is an admissible pair of anisotropies in accordance with Definition 2.1. The information on the surface tension  $\sigma$  and mobility  $\mu$  is contained in the functions  $f, g$  as defined in (1.5) and (1.6), respectively. We remark that there is some freedom with regard to the choice of  $g$ : Since  $\frac{\sigma(p)}{\mu(p)}$  is not defined for  $p = 0$ , adding  $+1$  in the numerator as well as the denominator (cf. (1.6)) is a means of avoiding a singularity at 0. Our existence and conditional convergence statements rely merely on the following properties of  $g$ :

**Lemma 3.2.** *Letting  $g$  be as in (1.6), the following holds:*

- (i) The function  $g : \mathbb{R}^d \rightarrow \mathbb{R}_{>0}$  is continuous,
- (ii) we have  $c_g := \inf_{\mathbb{R}^d} g > 0$ , and
- (iii)

$$\lim_{|p| \rightarrow \infty} \left| g(p) - \frac{\sigma(p)}{\mu(p)} \right| = 0. \tag{3.1}$$

The following weak solution concept for the anisotropic Allen–Cahn equation combines an integration by parts in space (ii) with an optimal energy dissipation identity (iii), which hints at the gradient flow structure of (1.4). Prescribing the initial data as in (i) is possible because  $u_\varepsilon(0)$  is well-defined via the embedding  $H^1(\Omega \times (0, T)) \subset C^{0, \frac{1}{2}}([0, T]; L^2(\Omega))$ .

**Definition 3.3** (Solutions to the anisotropic Allen–Cahn equation). Let  $\varepsilon > 0$  and  $u_{\varepsilon,0} \in \text{dom}(E_\varepsilon)$ . A function

$$u_\varepsilon \in H^1(\Omega \times (0, T)) \cap L^\infty(\Omega \times (0, T))$$

is a solution to the anisotropic Allen–Cahn equation with initial data  $u_{\varepsilon,0}$  if

- (i)

$$u_\varepsilon(0) = u_{\varepsilon,0}, \tag{3.2}$$

(ii)

$$\int_0^T \int_{\Omega} g(-\nabla u_{\varepsilon}) \partial_t u_{\varepsilon} \varphi \, dx dt = \int_0^T \int_{\Omega} \frac{1}{2} \left( Df(-\nabla u_{\varepsilon}) \cdot \nabla \varphi - \frac{1}{\varepsilon^2} W'(u_{\varepsilon}) \varphi \right) \, dx dt \tag{3.3}$$

for all  $\varphi \in C^1(\Omega \times [0, T])$ , and

(iii)

$$E_{\varepsilon}[u_{\varepsilon}(T')] + \int_0^{T'} \int_{\Omega} \varepsilon g(-\nabla u_{\varepsilon}) (\partial_t u_{\varepsilon})^2 \, dx dt = E_{\varepsilon}[u_{\varepsilon,0}] \tag{3.4}$$

for all  $T' \in (0, T]$ .

We comment on the typical behavior of the anisotropic Allen–Cahn equation (1.4).

**Remark 3.4.** The reaction term  $-\frac{1}{\varepsilon^2} W'(u_{\varepsilon})$  forces the solution  $u_{\varepsilon}$  towards the wells of  $W$ , i.e., towards the set  $\{0, 1\}$ . The interplay of the reaction term with the anisotropic diffusion term  $-\operatorname{div}(Df(-\nabla u_{\varepsilon}))$  results in a transition layer of width  $\mathcal{O}(\varepsilon)$  whose shape also depends on the direction of the approximate outer normal  $-\frac{\nabla u_{\varepsilon}}{|\nabla u_{\varepsilon}|}$ . More precisely, the typical width of the transition layer is  $\varepsilon \sigma\left(-\frac{\nabla u_{\varepsilon}}{|\nabla u_{\varepsilon}|}\right)$ .

To see this, one can choose the 1-dimensional stationary ansatz

$$u_{\varepsilon} : \mathbb{R}^d \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}, \quad u_{\varepsilon}(x, t) = U_{\varepsilon}(x \cdot \nu), \tag{3.5}$$

with  $\nu \in S^{d-1}$  being a fixed unit-length vector. Together with a monotonicity assumption on the transition, this admits the solution

$$u_{\varepsilon}(x, t) = \Theta\left(-\frac{x \cdot \nu}{\varepsilon \sigma(\nu)}\right), \tag{3.6}$$

where  $\Theta$  is the unique solution to

$$\Theta' = \sqrt{W(\Theta)}, \quad \Theta(0) = 1/2. \tag{3.7}$$

With this in mind, we note that the weighted  $L^2$ -metric (1.8) depends on  $\sigma$  and  $\mu$ , whereas anisotropic mean curvature is a gradient flow with respect to the metric (1.3), which depends only on  $\mu$ . This can be explained by the idea that, in the anisotropic Allen–Cahn equation, the metric (1.8) has to compensate for the typical width of the transition layer varying with the orientation of the normal.

### 3.1 Existence of weak solutions

**Theorem 3.5** (Existence). *Let  $\varepsilon > 0$  and  $u_0 \in \operatorname{dom}(E_{\varepsilon}) \cap L^{\infty}(\Omega)$ . For any finite time horizon  $T \in (0, \infty)$ , there exists a solution  $u$  to the anisotropic Allen–Cahn equation in the sense of Definition 3.3 with initial data  $u_0$ . Furthermore,  $u$  satisfies*

$$\|u - 1/2\|_{L^{\infty}(\Omega \times (0, T))} \leq \max\{\|u_0 - 1/2\|_{L^{\infty}(\Omega)}, 1/2\}. \tag{3.8}$$

To prove this existence theorem, we exploit the gradient-flow structure of the equation and construct solutions via a minimizing movements scheme. To this end, we consider an approximation of the PDE (1.4) that replaces the time derivative by difference quotients: Given a time-step size  $h > 0$ , let us look for functions  $\{u_h^n\}_{n \in \mathbb{N}_0}$  that solve

$$2g(-\nabla u_h^{n-1}) \frac{u_h^n - u_h^{n-1}}{h} = -\operatorname{div}(Df(-\nabla u_h^n)) - \frac{1}{\varepsilon^2} W'(u_h^n), \quad n \in \mathbb{N}, \tag{3.9}$$

$$u_h^0 = u_0. \tag{3.10}$$

In fact, equation (3.9) is the strong Euler–Lagrange equation of the variational problem

$$u_h^n \in \arg \min_{u \in L^2(\Omega)} \left\{ E_{\varepsilon}[u] + \frac{1}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2 \right\}, \quad n \in \mathbb{N}, \tag{3.11}$$

where we recall the energy (1.7) and the inner-product (1.8) for which we use the notation

$$\|v\|_u := (v, v)_u^{\frac{1}{2}},$$

in which the dependence on  $\varepsilon$  is suppressed again.

Equation (3.9) does not precisely resemble an implicit Euler scheme as it features the explicit term  $g(-\nabla u_h^{n-1})$  rather than the implicit term  $g(-\nabla u_h^n)$ . Instead (3.9) can be viewed an implicit-explicit splitting discretization scheme of (1.4). This is mirrored in the minimization problem (3.11) by the occurrence of the metric  $(\cdot, \cdot)_{u_h^{n-1}}$ , which is taken with respect to the constant point  $u_h^{n-1}$  and therefore does not depend on  $u$ . Thanks to this choice, one can immediately prove that (3.11) admits a unique solution as soon as  $h$  is small enough, using the direct method in the calculus of variations:

**Lemma 3.6.** *Let  $E_\varepsilon$  be defined as in (1.7) and assumptions (W1)–(W5) hold. It follows that:*

- (i) *The energy  $E_\varepsilon$  is  $\frac{\lambda}{2\varepsilon}$ -convex on  $L^2(\Omega)$ ,*
- (ii) *if  $h < \frac{2\varepsilon^2 c_g}{\lambda}$ , then  $u \mapsto E_\varepsilon[u] + \frac{1}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2$  is strongly convex,*
- (iii) *for all  $h > 0$ , the mapping  $u \mapsto E_\varepsilon[u] + \frac{1}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2$  is lower semicontinuous on  $L^2(\Omega)$ ,*
- (iii) *for  $h < \frac{2\varepsilon^2 c_g}{\lambda}$  and  $u_h^{n-1} \in \text{dom}(E_\varepsilon)$ , there exists a unique minimizer  $u_h^n \in \text{dom}(E_\varepsilon)$  for (3.11), and  $\|u_h^n\|_{L^\infty} \leq \max\{\|u_h^{n-1}\|_{L^\infty}, 1\}$ .*

**Proof.**

- (i) Let  $u, v \in L^2(\Omega)$  and  $\mu \in [0, 1]$ . We will use the equivalent formulation (ii) in Remark 3.1, so that we have to show that

$$E_\varepsilon[(1 - \mu)u + \mu v] \leq (1 - \mu)E_\varepsilon[u] + \mu E_\varepsilon[v] + \frac{\lambda}{4\varepsilon} \mu(1 - \mu) \|v - u\|_{L^2}^2$$

for all  $u, v \in L^2(\Omega)$  and  $\mu \in (0, 1)$ . We can assume without loss of generality that  $u, v \in \text{dom}(E_\varepsilon)$ . Exploiting the convexity of  $f$  as well as the  $\lambda$ -convexity of  $W$  (W4), we obtain

$$\begin{aligned} E_\varepsilon[(1 - \mu)u + \mu v] &= \frac{1}{2} \int_\Omega \left( \varepsilon f(-(1 - \mu)\nabla u - \mu\nabla v) + \frac{1}{\varepsilon} W((1 - \mu)u + \mu v) \right) dx \\ &\leq \frac{1}{2} \int_\Omega (\varepsilon(1 - \mu)f(-\nabla u) + \varepsilon\mu f(-\nabla v)) dx \\ &\quad + \frac{1}{2} \int_\Omega \left( \frac{1 - \mu}{\varepsilon} W(u) + \frac{\mu}{\varepsilon} W(v) + \frac{\lambda}{2\varepsilon} \mu(1 - \mu) |v - u|^2 \right) dx \\ &\leq (1 - \mu)E_\varepsilon[u] + \mu E_\varepsilon[v] + \frac{\lambda}{4\varepsilon} \mu(1 - \mu) \|v - u\|_{L^2}^2. \end{aligned} \tag{3.12}$$

- (ii) This follows from (i) and quantifying the strong convexity of the metric term: Given  $u, v \in L^2(\Omega)$  and  $\mu \in (0, 1)$ , we compute

$$\begin{aligned} &\frac{1}{2h} \|((1 - \mu)u + \mu v) - u_h^{n-1}\|_{u_h^{n-1}}^2 \\ &= \frac{1 - \mu}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2 + \frac{\mu}{2h} \|v - u_h^{n-1}\|_{u_h^{n-1}}^2 \\ &\quad - \frac{\mu(1 - \mu)}{2h} \int_\Omega \varepsilon g(-\nabla u_h^{n-1}) |v - u|^2 dx \end{aligned}$$

$$\begin{aligned} &\leq \frac{1-\mu}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2 + \frac{\mu}{2h} \|v - u_h^{n-1}\|_{u_h^{n-1}}^2 \\ &\quad - \frac{\varepsilon c_g}{2h} \mu(1-\mu) \|v - u\|_{L^2}^2. \end{aligned} \tag{3.13}$$

Combining (3.12) with (3.13) yields strong convexity for  $u \mapsto E_\varepsilon[u] + \frac{1}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2$  if  $\frac{\lambda}{4\varepsilon} < \frac{\varepsilon c_g}{2h}$ , or, equivalently, if  $h < \frac{2\varepsilon^2 c_g}{\lambda}$ .

- (iii) This will follow from the lower semi-continuity of the individual terms. To see that the anisotropic Dirichlet energy is lower semi-continuous, we can assume without restriction that  $v_k \in H^1(\Omega)$  for all  $k \in \mathbb{N}$ , so that it remains to prove that

$$\frac{\varepsilon}{2} \int_{\Omega} f(-\nabla v) dx \leq \liminf_{k \rightarrow \infty} \frac{\varepsilon}{2} \int_{\Omega} f(-\nabla v_k) dx.$$

Without loss of generality,  $v_k$  converge to  $v$  weakly in  $H^1(\Omega)$ , and lower semi-continuity follows from convexity of  $f$ . The lower semicontinuity of the potential term  $\int_{\Omega} W(u) dx$  and the metric term  $u \mapsto \frac{1}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2$  follow from Fatou's lemma.

- (iv) The functional to be minimized is proper. Let  $\{v_k\}_{k \in \mathbb{N}}$  be a minimizing sequence, i.e.,

$$\lim_{k \rightarrow \infty} \left( E_\varepsilon[v_k] + \frac{1}{2h} \|v_k - u_h^{n-1}\|_{u_h^{n-1}}^2 \right) = \inf_{u \in L^2(\Omega)} \left\{ E_\varepsilon[u] + \frac{1}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2 \right\}.$$

Every minimizing sequence is bounded in  $H^1(\Omega)$  since

$$\begin{aligned} \|v_k\|_{H^1}^2 &\leq 2 \|v_k - u_h^{n-1}\|_{L^2}^2 + 2 \|u_h^{n-1}\|_{L^2}^2 + \|\nabla v_k\|_{L^2}^2 \\ &\leq \frac{2}{\varepsilon c_g} \|v_k - u_h^{n-1}\|_{u_h^{n-1}}^2 + 2 \|u_h^{n-1}\|_{L^2}^2 + \frac{1}{(\min_{S^{d-1}} \sigma)^2} \int_{\Omega} f(-\nabla v_k) dx \\ &\leq 2 \|u_h^{n-1}\|_{L^2}^2 + \left( \frac{4h}{\varepsilon c_g} + \frac{2}{\varepsilon(\min_{S^{d-1}} \sigma)^2} \right) \left( E_\varepsilon[v_k] + \frac{1}{2h} \|v_k - u_h^{n-1}\|_{u_h^{n-1}}^2 \right). \end{aligned} \tag{3.14}$$

By Rellich's theorem, there exists a subsequence converging in  $L^2(\Omega)$  to a limit function  $v \in L^2(\Omega)$ . The lower semicontinuity (iii) then shows that

$$v \in \arg \min_{u \in L^2(\Omega)} \left\{ E_\varepsilon[u] + \frac{1}{2h} \|u - u_h^{n-1}\|_{u_h^{n-1}}^2 \right\}.$$

The minimizer  $v$  is unique due to the strong convexity (ii). For the  $L^\infty$ -bound, we define a truncated version of  $u_h^n$  by

$$v(x) := \begin{cases} -\alpha^{n-1} + \frac{1}{2} & \text{if } u_h^n(x) < -\alpha^{n-1} + \frac{1}{2}, \\ u_h^n(x) & \text{if } |u_h^n(x) - \frac{1}{2}| \leq \alpha^{n-1}, \\ \alpha^{n-1} + \frac{1}{2} & \text{if } u_h^n(x) > \alpha^{n-1} + \frac{1}{2}, \end{cases}$$

where  $\alpha^{n-1} := \max\{\|u_h^{n-1} - \frac{1}{2}\|_{L^\infty}, \frac{1}{2}\}$ . Letting  $A := \{x \in \Omega \mid v(x) \neq u_h^n(x)\}$ , it follows from  $u_h^n \in \text{dom}(E_\varepsilon) \subseteq H^1(\Omega)$  and a general fact about Sobolev spaces that  $v \in H^1(\Omega)$  and

$$\nabla v(x) = \begin{cases} \nabla u_h^n(x) & \text{for almost every } x \in \Omega \setminus A, \\ 0 & \text{for almost every } x \in A. \end{cases} \tag{3.15}$$



Then, using (3.15), the positive definiteness of  $\sigma$ , the monotonicity assumption (W5), and the pointwise inequality  $|v - u_h^{n-1}| \leq |u_h^n - u_h^{n-1}|$ , one obtains

$$\begin{aligned} E_\varepsilon[v] + \frac{1}{2h} \|v - u_h^{n-1}\|_{u_h^{n-1}}^2 &\leq \frac{\varepsilon}{2} \int_\Omega f(-\nabla u_h^n) dx + \frac{1}{2\varepsilon} \int_\Omega W(u_h^n) dx \\ &\quad + \frac{\varepsilon}{2h} \int_\Omega g(-\nabla u_h^{n-1}) |u_h^n - u_h^{n-1}|^2 dx \\ &= E_\varepsilon[u_h^n] + \frac{1}{2h} \|u_h^n - u_h^{n-1}\|_{u_h^{n-1}}^2. \end{aligned}$$

By the uniqueness of minimizers, this implies that  $v = u_h^n$  and, therefore,

$$\|u_h^n - \frac{1}{2}\|_{L^\infty} = \|v - \frac{1}{2}\|_{L^\infty} \leq \alpha^{n-1} = \max \left\{ \|u_h^{n-1} - \frac{1}{2}\|_{L^\infty}, \frac{1}{2} \right\}.$$

□

In order to pass to the limit  $h \searrow 0$ , we define one affine and two piecewise constant interpolations: For  $n \in \mathbb{N}$  such that  $t \in [(n-1)h, nh)$  let

$$\begin{aligned} u_h(t) &:= \frac{t - (n-1)h}{h} u_h^n + \frac{nh - t}{h} u_h^{n-1}, \\ \bar{u}_h(t) &:= u_h^n, \\ \underline{u}_h(t) &:= u_h^{n-1}, \end{aligned} \tag{3.16}$$

where  $u_h^0 \in \text{dom}(E)$  and  $u_h^n$  are defined inductively via (3.11) for  $n \in \mathbb{N}$ .

Before we turn to the limiting function  $u$ , let us prove the following inequality, which is a consequence of the  $\frac{\lambda}{2\varepsilon}$ -convexity of  $E_\varepsilon$  and the minimizing property of the time steps  $u_h^n, u_h^{n-1}$  (3.11):

**Lemma 3.7.** *Let  $h < \frac{2\varepsilon^2 c_g}{\lambda}$  and  $w \in L^2(\Omega \times (0, T))$ . Then*

$$E_\varepsilon[w(t)] \geq E_\varepsilon[\bar{u}_h(t)] - \int_\Omega \varepsilon g(-\nabla \underline{u}_h(t)) \partial_t u_h(t) (w(t) - \bar{u}_h(t)) dx - \frac{\lambda}{4\varepsilon} \int_\Omega |w(t) - \bar{u}_h(t)|^2 dx \tag{3.17}$$

for almost every  $t \in (0, T)$ .

**Proof.** It suffices to prove the corresponding inequality involving the time steps  $u_h^n, n \in \mathbb{N}$ : Let  $h < \frac{\varepsilon^2 c_g}{2\lambda}$  and  $n \in \mathbb{N}$ . Then for all  $v \in L^2(\Omega)$ , we have

$$E_\varepsilon[v] \geq E_\varepsilon[u_h^n] - \int_\Omega \varepsilon g(-\nabla u_h^{n-1}) \frac{u_h^n - u_h^{n-1}}{h} (v - u_h^n) dx - \frac{\lambda}{4\varepsilon} \int_\Omega |v - u_h^n|^2 dx. \tag{3.18}$$

Clearly, one can write the interpolations  $u_h, \bar{u}_h$ , and  $\underline{u}_h$  on the left-hand side of (3.17) in terms of the time steps, choosing  $n = n(t) \in \mathbb{N}$  such that  $t \in [(n(t)-1)h, n(t)h)$ . We also observe that the piecewise affine interpolation  $u_h \in C([0, T]; L^2(\Omega))$  has a weak time derivative  $\partial_t u_h$  given by

$$\partial_t u_h(t) = \frac{\bar{u}_h(t) - \underline{u}_h(t)}{h} = \frac{u_h^{n(t)} - u_h^{n(t)-1}}{h}$$

for almost every  $t \in (0, T)$ . Therefore, (3.17) follows from (3.18) by reformulating the interpolation in terms of the time steps and using  $v = w(t)$  for fixed  $t \in (0, T)$ .

To prove (3.18), let  $v \in L^2(\Omega)$  and  $\delta \in (0, 1)$ . Again, we use the equivalent characterization of  $\frac{\lambda}{2\varepsilon}$ -convexity in Remark 3.1(ii). For computational ease, first note, that

$$|\delta v + (1 - \delta)u_h^n - u_h^{n-1}|^2 = \delta^2 |v - u_h^n|^2 + 2\delta(v - u_h^n)(u_h^n - u_h^{n-1}) + |u_h^n - u_h^{n-1}|^2.$$

Subtracting  $(1 - \delta)|u^n - u^{n-1}|^2$  from both sides of the above equality, we may directly compute

$$\begin{aligned}
 E_\varepsilon[v] &\geq \frac{1}{\delta} \left( E_\varepsilon[\delta v + (1 - \delta)u_h^n] - (1 - \delta)E_\varepsilon[u_h^n] \right) - (1 - \delta) \frac{\lambda}{4\varepsilon} \int_\Omega |v - u_h^n|^2 dx \\
 &= \frac{1}{\delta} \left( E_\varepsilon[\delta v + (1 - \delta)u_h^n] + \frac{1}{2h} \|\delta v + (1 - \delta)u_h^n - u_h^{n-1}\|_{u_h^{n-1}}^2 \right. \\
 &\quad \left. - (1 - \delta)E_\varepsilon[u_h^n] - (1 - \delta) \frac{1}{2h} \|u_h^n - u_h^{n-1}\|_{u_h^{n-1}}^2 \right) \\
 &\quad - \frac{1}{2h} \int_\Omega \varepsilon g(-\nabla u_h^{n-1}) (2(u_h^n - u_h^{n-1})(v - u_h^n) + (u_h^n - u_h^{n-1})^2 + \delta(u_h^n - v)^2) dx \\
 &\quad - (1 - \delta) \frac{\lambda}{4\varepsilon} \int_\Omega |v - u_h^n|^2 dx \\
 &\geq \frac{1}{\delta} \left( E_\varepsilon[u_h^n] + \frac{1}{2h} \|u_h^n - u_h^{n-1}\|_{u_h^{n-1}}^2 \right. \\
 &\quad \left. - (1 - \delta)E_\varepsilon[u_h^n] - (1 - \delta) \frac{1}{2h} \|u_h^n - u_h^{n-1}\|_{u_h^{n-1}}^2 \right) \\
 &\quad - \frac{1}{2h} \int_\Omega \varepsilon g(-\nabla u_h^{n-1}) (2(u_h^n - u_h^{n-1})(v - u_h^n) + (u_h^n - u_h^{n-1})^2 + \delta(u_h^n - v)^2) dx \\
 &\quad - (1 - \delta) \frac{\lambda}{4\varepsilon} \int_\Omega |v - u_h^n|^2 dx \\
 &= E_\varepsilon[u_h^n] - \int_\Omega \varepsilon g(-\nabla u_h^{n-1}) \frac{u_h^n - u_h^{n-1}}{h} (v - u_h^n) dx - \frac{\delta}{2h} \int_\Omega \varepsilon g(-\nabla u_h^{n-1}) (u_h^n - v)^2 dx \\
 &\quad - (1 - \delta) \frac{\lambda}{4\varepsilon} \int_\Omega |v - u_h^n|^2 dx, \tag{3.19}
 \end{aligned}$$

where the third step is an application of (3.11), with  $\delta v + (1 - \delta)u_h^n$  acting as a contender for the minimization problem. Taking the limit  $\delta \searrow 0$  finally yields (3.18).  $\square$

An application of the Arzelà–Ascoli theorem, as performed in the lemma below, shows that a subsequence of  $\{u_h\}_h$  converges strongly, and the derivatives in time and space will turn out to converge weakly in  $L^2$ . However, due to the nonlinear term  $g(-\nabla u)$  in the anisotropic Allen–Cahn equation (1.4), weak convergence of the gradients is not sufficient to show that the limit function solves the equation. We will therefore use a convexity argument to also prove strong convergence of the gradients.

**Lemma 3.8.** *Let  $\{u_h^0\}_{h \in (0, \frac{2\varepsilon^2 c_E}{\lambda})} \subset \text{dom}(E)$  such that  $\limsup_{h \searrow 0} E_\varepsilon[u_h^0] < \infty$  and  $\limsup_{h \searrow 0} \|u_h^0\|_{L^\infty(\Omega)} < \infty$ . Then there exist a sequence  $h \searrow 0$ , denoted without relabeling, and a limit function  $u \in H^1(\Omega \times (0, T))$  such that*

(i)

$$u_h(t), \bar{u}_h(t), \underline{u}_h(t) \longrightarrow u(t) \quad \text{in } L^2(\Omega) \text{ uniformly in } t \in [0, T],$$

(ii)

$$\partial_t u_h \rightharpoonup \partial_t u \quad \text{in } L^2(\Omega \times (0, T)),$$

(iii)

$$\nabla u_h, \nabla \bar{u}_h, \nabla \underline{u}_h \longrightarrow \nabla u \quad \text{in } L^2(\Omega \times (0, T))^d$$

as  $h \searrow 0$ , where  $u_h, \bar{u}_h$ , and  $\underline{u}_h$  are constructed as above. Furthermore, the mapping  $u : [0, T] \rightarrow L^2(\Omega)$  is  $\frac{1}{2}$ -Hölder continuous and satisfies the  $L^\infty$ -bound  $\|u - \frac{1}{2}\|_{L^\infty(\Omega \times (0, T))} \leq \max \left\{ \limsup_{h \searrow 0} \|u_h^0 - \frac{1}{2}\|_{L^\infty(\Omega)}, \frac{1}{2} \right\}$ .

**Proof.**

(i) The first step is to prove a  $H^1$ -bound for the minimizers  $u_h^n, n \in \mathbb{N}_0$ :

$$\begin{aligned} \|u_h^n\|_{H^1(\Omega)}^2 &= \|u_h^n\|_{L^2(\Omega)}^2 + \|\nabla u_h^n\|_{L^2(\Omega)}^2 \\ &\leq \|u_h^n\|_{L^\infty(\Omega)}^2 + \frac{1}{\min_{S^{d-1}} \sigma^2} \int_{\Omega} f(-\nabla u_h^n) dx \\ &\leq \|u_h^n\|_{L^\infty(\Omega)}^2 + \frac{2}{\varepsilon \min_{S^{d-1}} \sigma^2} E_\varepsilon[u_h^n] \\ &\leq \max \left\{ \|u_h^0\|_{L^\infty(\Omega)}^2, 1 \right\} + \frac{2}{\varepsilon \min_{S^{d-1}} \sigma^2} E_\varepsilon[u_h^0], \end{aligned} \tag{3.20}$$

where the last inequality uses Lemma 3.6(iv) and (3.11) repeatedly.

For any  $t \in [0, T]$ , the function  $u_h(t)$  is a convex combination of two minimizers  $u_h^{n-1}, u_h^{n-1}, u_h^n$  of (3.11). Applying the triangle inequality yields

$$\begin{aligned} \sup_{t \in [0, T]} \limsup_{h \searrow 0} \|u_h(t)\|_{H^1(\Omega)} &\leq \left( \max \left\{ \limsup_{h \searrow 0} \|u_h^n\|_{L^\infty(\Omega)}, 1 \right\} \right. \\ &\quad \left. + \frac{2}{\varepsilon \min_{S^{d-1}} \sigma^2} \limsup_{h \searrow 0} E_\varepsilon[u_h^0] \right)^{\frac{1}{2}} \\ &< \infty. \end{aligned} \tag{3.21}$$

It follows from Rellich’s compact embedding theorem that any sequence  $\{u_{h_k}(t)\}_{k \in \mathbb{N}} \subset \mathbb{N}$  with  $t \in [0, T]$  and  $\lim_{k \rightarrow \infty} h_k = 0$  has a subsequence converging in  $L^2(\Omega)$ . To prove equicontinuity in time, which is the second requirement for the Arzelà–Ascoli theorem, we will – as is standard with minimizing movement approximations – show the stronger statement that  $u_h \in C([0, T]; L^2(\Omega))$  are  $\frac{1}{2}$ -Hölder continuous and the Hölder constants are uniformly bounded as  $h \searrow 0$ . Indeed, for  $n \in \mathbb{N}$ , we use the definition of the metric  $(\cdot, \cdot)_{u_h^{n-1}}$  and the minimization property (3.11) to show that

$$\begin{aligned} \|u_h^n - u_h^{n-1}\|_{L^2(\Omega)}^2 &\leq \frac{1}{\varepsilon C_g} \|u_h^n - u_h^{n-1}\|_{u_h^{n-1}}^2 \\ &= \frac{2h}{\varepsilon C_g} \left( E_\varepsilon[u_h^n] + \frac{1}{2h} \|u_h^n - u_h^{n-1}\|_{u_h^{n-1}}^2 - E_\varepsilon[u_h^n] \right) \\ &\leq \frac{2h}{\varepsilon C_g} (E_\varepsilon[u_h^{n-1}] - E_\varepsilon[u_h^n]). \end{aligned} \tag{3.22}$$

Inequality (3.22) allows us to bound the  $L^2$ -norms of the weak time derivatives  $\partial_t u_h$  as

$$\begin{aligned} \|\partial_t u_h\|_{L^2(0, T; L^2(\Omega))}^2 &\leq \sum_{n=1}^{\lceil \frac{T}{h} \rceil} \int_{(n-1)h}^{nh} \|\partial_t u_h(t)\|_{L^2(\Omega)}^2 dt \\ &= \sum_{n=1}^{\lceil \frac{T}{h} \rceil} \int_{(n-1)h}^{nh} \frac{\|u_h^n - u_h^{n-1}\|_{L^2(\Omega)}^2}{h^2} dt \end{aligned}$$

$$\begin{aligned} &\leq \sum_{n=1}^{\lceil \frac{T}{h} \rceil} \frac{2}{\varepsilon C_g} (E_\varepsilon[u_h^{n-1}] - E_\varepsilon[u_h^n]) \\ &\leq \frac{2}{\varepsilon C_g} E_\varepsilon[u_h^0]. \end{aligned} \tag{3.23}$$

The fundamental theorem of calculus for vector-valued functions together with the Cauchy–Schwarz inequality now yields

$$\begin{aligned} \|u_h(t) - u_h(s)\|_{L^2(\Omega)} &\leq \int_{(s,t)} \|\partial_r u_h(r)\|_{L^2(\Omega)} dr \\ &\leq \|\partial_r u_h\|_{L^2(s,t;L^2(\Omega))} \sqrt{|t-s|} \\ &\leq \sqrt{\frac{2E_\varepsilon[u_h^0]}{\varepsilon C_g}} \sqrt{|t-s|} \end{aligned} \tag{3.24}$$

for all  $s, t \in [0, T]$ . This is the one-dimensional case of Morrey’s inequality.

By (3.24) and the assumption on the initial data, the Hölder constants are bounded as  $h \searrow 0$ . Families of functions with bounded Hölder constants are equicontinuous. These equicontinuity and precompactness statements allow us to apply the Arzelà–Ascoli theorem, which states that there exists a subsequence  $h \searrow 0$  and a function  $u \in C([0, T]; L^2(\Omega))$  such that  $u_h \rightarrow u$  in  $C([0, T]; L^2(\Omega))$ . Furthermore, the Hölder continuity carries over to the limit function, i.e.,

$$\|u(t) - u(s)\|_{L^2(\Omega)} \leq \sqrt{\frac{2 \limsup_{h \searrow 0} E_\varepsilon[u_h^0]}{\varepsilon 0}} E_\varepsilon[u_h^0] \varepsilon C_g \sqrt{|t-s|} \quad \text{for all } s, t \in [0, T]. \tag{3.25}$$

One can extract a further subsequence  $\{u_{h_k}\}_{k \in \mathbb{N}}$  such that  $h_k \searrow 0$  and  $u_{h_k}(x, t) \rightarrow u(x, t)$  almost everywhere in  $\Omega \times (0, T)$ . This pointwise convergence together with the triangle inequality yields

$$\begin{aligned} \|u - \frac{1}{2}\|_{L^\infty(\Omega \times (0,T))} &\leq \limsup_{k \rightarrow \infty} \|u_{h_k} - \frac{1}{2}\|_{L^\infty(\Omega \times (0,T))} \\ &\leq \limsup_{h \searrow 0} \|u_h - \frac{1}{2}\|_{L^\infty(\Omega \times (0,T))} \\ &\leq \limsup_{h \searrow 0} \sup_{n \in \mathbb{N}} \|u_h^n - \frac{1}{2}\|_{L^\infty(\Omega)} \\ &\leq \max \left\{ \limsup_{h \searrow 0} \|u_h^0 - \frac{1}{2}\|_{L^\infty(\Omega)}, \frac{1}{2} \right\}. \end{aligned}$$

In a last step, we argue briefly that the piecewise constant interpolations  $\bar{u}_h$  and  $\underline{u}_h$  converge uniformly.

Given  $t \in [0, T]$ , we find  $n \in \mathbb{N}$  such that  $t \in [(n-1)h, nh)$ . Then

$$\|\bar{u}_h(t) - u_h(t)\|_{L^2(\Omega)} = \|u_h(nh) - u_h(t)\|_{L^2(\Omega)} \leq \sqrt{\frac{2E_\varepsilon[u_h^0]}{\varepsilon C_g}} \sqrt{nh-t} \leq \sqrt{\frac{2E_\varepsilon[u_h^0]}{\varepsilon C_g}} \sqrt{h},$$

and a similar argument shows that

$$\|\underline{u}_h(t) - u_h(t)\|_{L^2(\Omega)} \leq \sqrt{\frac{2E_\varepsilon[u_h^0]}{\varepsilon C_g}} \sqrt{h}.$$

Combining this bound with the uniform convergence  $u_h \rightarrow u$  in  $C([0, T]; L^2(\Omega))$ , we obtain

$$\begin{aligned} & \limsup_{h \searrow 0} \sup_{t \in [0, T]} \max \{ \|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)}, \|\underline{u}_h(t) - u(t)\|_{L^2(\Omega)} \} \\ & \leq \limsup_{h \searrow 0} \sup_{t \in [0, T]} \|u_h(t) - u(t)\|_{L^2(\Omega)} + \lim_{h \searrow 0} \sqrt{\frac{2E_\varepsilon[u_h^0]}{\varepsilon C_g}} \sqrt{h} \\ & = 0. \end{aligned}$$

- (ii) It follows from (i) that  $u_h \rightarrow u$  as  $h \searrow 0$  in  $L^2(0, T; L^2(\Omega)) \cong L^2(\Omega \times (0, T))$ . Furthermore, by (3.23), the time derivatives  $\{\partial_t u_h\}$  are bounded in  $L^2(\Omega \times (0, T))$  as  $h \searrow 0$ . By, e.g., [2, Prop. 2.5(b)], one concludes that  $u$  has a weak time derivative  $\partial_t u \in L^2(\Omega \times (0, T))$ , and that  $\partial_t u_h \rightarrow \partial_t u$  in  $L^2(\Omega \times (0, T))$  as  $h \searrow 0$ .
- (iii) We will prove the desired convergence result for the piecewise constant interpolation  $\bar{u}_h$  first. Integrating (3.20) from 0 to  $T$  shows that the gradients  $\{\nabla \bar{u}_h\}$  are bounded in  $L^2(\Omega \times (0, T))^d$  as  $h \searrow 0$ . Just like in (ii), it then follows from the strong convergence  $\bar{u}_h \rightarrow u$  in  $L^2(\Omega \times (0, T))$  that the limit function  $u$  has a weak gradient  $\nabla u \in L^2(\Omega \times (0, T))^d$ , and that  $\nabla \bar{u}_h \rightharpoonup \nabla u$  as  $h \searrow 0$  in  $L^2(\Omega \times (0, T))^d$ .

The next step is to upgrade this weak convergence statement for the gradients to strong convergence in  $L^2$ . The key ingredient for the argument will be the energy convergence

$$E_\varepsilon[\bar{u}_h(\cdot)] \rightarrow E_\varepsilon[u(\cdot)] \quad \text{in } L^1(0, T). \tag{3.26}$$

Let us first argue how (3.26) implies strong convergence of the gradients. Since the Cahn–Hilliard energy  $E_\varepsilon$  is made up of a Dirichlet energy and a nonconvex term involving  $W$ , we will use the convergence result (3.26) for the Cahn–Hilliard energies and a liminf inequality for the nonconvex part to derive a limsup inequality for the Dirichlet energy, which will allow us to prove the strong convergence statement.

Similarly to the proof of Lemma 3.6(iii), an application of Fatou’s lemma yields

$$\liminf_{h \searrow 0} \int_0^T \int_\Omega W(\bar{u}_h) dx dt \geq \int_0^T \int_\Omega W(u) dx dt. \tag{3.27}$$

One can now test (3.26) with a constant test function and combine this convergence statement with (3.27), which yields

$$\begin{aligned} \limsup_{h \searrow 0} \int_0^T \int_\Omega f(-\nabla \bar{u}_h) dx dt &= \frac{2}{\varepsilon} \lim_{h \searrow 0} \int_0^T E_\varepsilon[\bar{u}_h(t)] dt - \frac{1}{\varepsilon^2} \liminf_{h \searrow 0} \int_0^T \int_\Omega W(\bar{u}_h) dx dt \\ &\leq \frac{2}{\varepsilon} \int_0^T E_\varepsilon[u(t)] dt - \frac{1}{\varepsilon^2} \int_0^T \int_\Omega W(u) dx dt \\ &= \int_0^T \int_\Omega f(-\nabla u) dx dt. \end{aligned} \tag{3.28}$$

The function  $f$  is strongly convex, i.e., there exists a constant  $c > 0$  such that  $f(p') \geq f(p) + Df(p) \cdot (p' - p) + \frac{c}{2} |p' - p|^2$  for all  $p, p' \in \mathbb{R}^d$ . Thus, if (3.28) holds true, then

$$\begin{aligned} & \limsup_{h \searrow 0} \frac{c}{2} \int_0^T \int_\Omega |\nabla \bar{u}_h - \nabla u|^2 dx dt \\ & \leq \limsup_{h \searrow 0} \left( \int_0^T \int_\Omega f(-\nabla \bar{u}_h) dx dt - \int_0^T \int_\Omega f(-\nabla u) dx dt \right. \\ & \quad \left. - \int_0^T \int_\Omega Df(-\nabla u) \cdot (\nabla u - \nabla \bar{u}_h) dx dt \right) \end{aligned}$$

$$\begin{aligned} &\leq \limsup_{h \searrow 0} \int_0^T \int_{\Omega} f(-\nabla \bar{u}_h) dx dt - \int_0^T \int_{\Omega} f(-\nabla u) dx dt \\ &\quad - \liminf_{h \searrow 0} \int_0^T \int_{\Omega} Df(-\nabla u) \cdot (\nabla u - \nabla \bar{u}_h) dx dt \\ &\leq 0, \end{aligned}$$

where the last step also uses the fact that  $\nabla u_h \rightarrow \nabla u$  in  $L^2(\Omega \times (0, T))^d$ . This computation implies that  $\nabla \bar{u}_h \rightarrow \nabla u$  strongly in  $L^2(\Omega \times (0, T))$ .

In order to prove (3.26), it suffices to consider nonnegative test functions  $\zeta \in L^\infty(0, T; \mathbb{R}_{\geq 0})$ . We will prove a liminf inequality and a limsup inequality separately.

On the one hand, by Fatou’s lemma, the lower semicontinuity of  $E_\varepsilon$ , and (i), we obtain

$$\liminf_{h \searrow 0} \int_0^T \zeta(t) E_\varepsilon[\bar{u}_h(t)] dt \geq \int_0^T \zeta(t) \liminf_{h \searrow 0} E_\varepsilon[\bar{u}_h(t)] dt \geq \int_0^T \zeta(t) E_\varepsilon[u(t)] dt. \tag{3.29}$$

On the other hand, choosing  $w(x, t) := u(x, t)$  in (3.17) and integrating against  $\zeta$  yields

$$\begin{aligned} \int_0^T \zeta(t) E_\varepsilon[u(t)] dt &\geq \int_0^T \zeta(t) E_\varepsilon[\bar{u}_h(t)] dt \\ &\quad - \int_0^T \zeta(t) \int_{\Omega} \left( \varepsilon g(-\nabla \underline{u}_h) \partial_t u_h (u - \bar{u}_h) + \frac{\lambda}{4\varepsilon} |u - \bar{u}_h|^2 \right) dx dt. \end{aligned}$$

In the limit  $h \searrow 0$ , the second integral on the right-hand side vanishes since  $\bar{u}_h \rightarrow u$  in  $L^2(\Omega \times (0, T))$  and  $g$  is bounded in  $L^\infty$  by Lemma 3.2. Thus,

$$\int_0^T \zeta(t) E_\varepsilon[u(t)] dt \geq \limsup_{h \searrow 0} \int_0^T \zeta(t) E_\varepsilon[\bar{u}_h(t)] dt. \tag{3.30}$$

Combining (3.29) and (3.30) yields (3.26).

We have shown that  $\nabla \bar{u}_h \rightarrow \nabla u$  in  $L^2(\Omega \times (0, T))$  as  $h \searrow 0$ . As for  $\nabla \underline{u}_h$ , one computes

$$\begin{aligned} &\|\nabla \underline{u}_h - \nabla u\|_{L^2(\Omega \times (0, T))}^2 \\ &= \int_0^h \int_{\Omega} |\nabla \underline{u}_h - \nabla u|^2 dx dt + \int_h^T \int_{\Omega} |\nabla \underline{u}_h - \nabla u|^2 dx dt \\ &= \int_0^h \int_{\Omega} |\nabla u_h^0 - \nabla u|^2 dx dt + \int_h^T \int_{\Omega} |\nabla \bar{u}_h(x, t-h) - \nabla u(x, t)|^2 dx dt \\ &\leq 2h \int_{\Omega} |\nabla u_h^0|^2 dx + 2 \int_0^h \int_{\Omega} |\nabla u|^2 dx dt \\ &\quad + 2 \int_0^{T-h} \int_{\Omega} |\nabla \bar{u}_h(x, t) - \nabla u(x, t)|^2 dx dt \\ &\quad + 2 \int_0^{T-h} \int_{\Omega} |\nabla u(x, t) - \nabla u(x, t+h)|^2 dx dt \\ &\longrightarrow 0 \quad \text{as } h \searrow 0, \end{aligned}$$

where the last integral vanishes as  $h \searrow 0$  due to the continuity of translation.

Lastly,  $\nabla u_h(x, t)$  is a convex combination of  $\nabla \bar{u}_h(x, t)$  and  $\nabla \underline{u}_h(x, t)$  for almost all  $(x, t) \in \Omega \times (0, T)$ . Hence, it follows from  $\nabla \bar{u}_h \rightarrow \nabla u$  and  $\nabla \underline{u}_h \rightarrow \nabla u$  that, as  $h \searrow 0$ , we have  $\nabla u_h \rightarrow \nabla u$  in  $L^2(\Omega \times (0, T))$  as well. □

We are now ready to pass to the limit  $h \searrow 0$  in (3.17), which will eventually allow us to prove Theorem 3.5. This is the content of the following lemma:

**Lemma 3.9.** *Let  $w \in L^2(\Omega \times (0, T))$  and  $\zeta \in L^\infty(0, T)$  such that  $\zeta \geq 0$  almost everywhere. Then*

$$\int_0^T \zeta(t) \left( E_\varepsilon[w(t)] - E_\varepsilon[u(t)] + \int_\Omega \varepsilon g(-\nabla u) \partial_t u (w - u) dx + \frac{\lambda}{4\varepsilon} \int_\Omega |w - u|^2 dx \right) dt \geq 0. \tag{3.31}$$

In other words, for every  $w \in L^2(\Omega \times (0, T))$ , the inequality  $E_\varepsilon[w(t)] \geq E_\varepsilon[u(t)] - \int_\Omega \varepsilon g(-\nabla u) \partial_t u (w - u) dx - \frac{\lambda}{4\varepsilon} \int_\Omega |w - u|^2 dx$  holds true for almost every  $t \in (0, T)$ .

**Proof.** First, by the weak convergence (3.26), it follows that

$$\lim_{h \searrow 0} \int_0^T \zeta(t) E_\varepsilon[\bar{u}_h(t)] dt = \int_0^T \zeta(t) E_\varepsilon[u(t)] dt \tag{3.32}$$

for any function  $\zeta \in L^\infty(0, T)$  such that  $\zeta \geq 0$ .

Second, for the term involving the time derivative, we know from Lemma 3.8(ii) that  $\partial_t u_h \rightharpoonup \partial_t u$  in  $L^2(\Omega \times (0, T))$ . Thus, to derive the convergence of the integrals, it suffices to show that  $g(-\nabla \underline{u}_h)(w - \bar{u}_h) \rightarrow g(-\nabla u)(w - u)$  (strongly) in  $L^2(\Omega \times (0, T))$ .

By Lemma 3.8(iii), we know that  $\nabla \bar{u}_h, \nabla \underline{u}_h \rightarrow \nabla u$  in  $L^2(\Omega \times (0, T))$ . Since  $g$  is a continuous function, every subsequence of  $g(-\nabla \underline{u}_h)$  as  $h \searrow 0$  has a further subsequence that converges pointwise almost everywhere to  $g(-\nabla u)$ . By the uniform boundedness of  $g$  and the dominated convergence theorem, it follows that

$$g(-\nabla \underline{u}_h(t))(w(t) - \bar{u}_h(t)) \rightarrow g(-\nabla u(t))(w(t) - u(t)) \quad \text{in } L^2(\Omega) \text{ for almost all } t \in (0, T).$$

It remains to show the  $L^2$ -convergence on the product space  $\Omega \times (0, T)$ . Using the generalized dominated convergence theorem with

$$\begin{aligned} \|g(-\nabla \underline{u}_h(t))(w(t) - \bar{u}_h(t))\|_{L^2(\Omega)} &\leq (\sup_{\mathbb{R}^d} g) \left( \int_\Omega (w(t) - \bar{u}_h(t))^2 dx \right)^{\frac{1}{2}} \\ &\rightarrow (\sup_{\mathbb{R}^d} g) \left( \int_\Omega (w(t) - u(t))^2 dx \right)^{\frac{1}{2}} \quad \text{in } L^2(0, T), \end{aligned}$$

it follows that

$$\int_0^T \|g(-\nabla \underline{u}_h(t))(w(t) - \bar{u}_h(t)) - g(-\nabla u(t))(w(t) - u(t))\|_{L^2(\Omega)}^2 dt \rightarrow 0,$$

i.e., by Fubini's theorem,  $g(-\nabla \underline{u}_h)(w - \bar{u}_h) \rightarrow g(-\nabla u)(w - u)$  in  $L^2(\Omega \times (0, T))$ . In total, we obtain

$$\begin{aligned} \lim_{h \searrow 0} \int_0^T \zeta(t) \int_\Omega \varepsilon g(-\nabla \underline{u}_h(t)) \partial_t u_h(t) (w(t) - \bar{u}_h(t)) dx dt \\ = \int_0^T \zeta(t) \int_\Omega \varepsilon g(-\nabla u(t)) \partial_t u(t) (w(t) - u(t)) dx dt. \end{aligned} \tag{3.33}$$

Third, it follows from Lemma 3.8(i) that

$$\lim_{h \searrow 0} \frac{\lambda}{4\varepsilon} \int_0^T \zeta(t) \int_\Omega |w(t) - \bar{u}_h(t)|^2 dx dt = \frac{\lambda}{4\varepsilon} \int_0^T \zeta(t) \int_\Omega |w(t) - u(t)|^2 dx dt. \tag{3.34}$$

Integrating (3.17) against  $\zeta$ , taking the limit  $h \searrow 0$ , and plugging in (3.32)–(3.34) proves the desired inequality.  $\square$

With the help of Lemma 3.9, it can be seen that  $-\partial_t u(t)$  lies in the  $u(t)$ -subdifferential  $\partial_{u(t)} E_\varepsilon[u(t)]$  for almost every  $t \in (0, T)$ . Thus, it seems plausible that the limiting trajectory  $u(t)$  is a gradient flow for  $E_\varepsilon$ .

**Proof of Theorem 3.5.** Choosing  $u_h^0 = u_0$ , one can construct the functions  $u_h, \bar{u}_h$ , and  $\underline{u}_h$  by Lemma 3.6 and (3.16). Invoking Lemma 3.8 yields a subsequence  $h \searrow 0$  as well as a function  $u \in H^1(\Omega \times (0, T)) \cap$

$L^\infty(\Omega \times (0, T))$  such that  $\|u - \frac{1}{2}\|_{L^\infty(\Omega \times (0, T))} \leq \max\{\|u_0 - \frac{1}{2}\|_{L^\infty(\Omega)}, \frac{1}{2}\}$ , and such that the three convergence results in Lemma 3.8(i)–(iii) hold true. It remains to be shown that  $u$  is a weak solution to the anisotropic Allen–Cahn equation with initial data  $u_0$  in the sense of Definition 3.3.

As for the initial condition, it follows from the uniform convergence in Lemma 3.8(i) and the choice  $u_h^0 = u_0$  that

$$u(0) = \lim_{h \searrow 0} u_h(0) = \lim_{h \searrow 0} u_h^0 = u_0,$$

where all limits are in  $L^2(\Omega)$ .

For the optimal energy dissipation relation (3.4) let  $T' \in (0, T]$  be a fixed time horizon. We define an extension of  $u : \Omega \times [0, T'] \rightarrow \mathbb{R}$  to  $\Omega \times \mathbb{R}$  by

$$\tilde{u}(x, t) = \begin{cases} u(x, t) & \text{if } 0 \leq t \leq T', \\ u(x, 0) & \text{if } t < 0, \\ u(x, T') & \text{if } t > T'. \end{cases}$$

Then  $\tilde{u} \in H^1_{\text{loc}}(\Omega \times \mathbb{R})$ , with the weak time derivative being given by  $\partial_t \tilde{u}(x, t) = \chi_{\Omega \times (0, T')} \partial_t u(x, t)$ . Taking  $a > 0$ ,  $\zeta := \frac{1}{a} \chi_{(0, T')}$ , and  $w(x, t) := \tilde{u}(x, t + a)$  in (3.31) yields

$$\begin{aligned} & \int_0^{T'} \int_\Omega \varepsilon g(-\nabla u(t)) \partial_t u(t) \frac{\tilde{u}(t+a) - \tilde{u}(t)}{a} dx dt + a \int_0^{T'} \int_\Omega \frac{\lambda}{4\varepsilon} \left( \frac{\tilde{u}(t+a) - \tilde{u}(t)}{a} \right)^2 dx dt \\ & \geq - \int_0^{T'} \frac{E_\varepsilon[\tilde{u}(t+a)] - E_\varepsilon[\tilde{u}(t)]}{a} dt \\ & = \frac{1}{a} \int_0^a E_\varepsilon[\tilde{u}(t)] dt - \frac{1}{a} \int_{T'}^{T'+a} E_\varepsilon[\tilde{u}(t)] dt \\ & = \frac{1}{a} \int_0^a E_\varepsilon[u(t)] dt - E_\varepsilon[u(T')], \end{aligned} \tag{3.35}$$

where we have replaced  $u$  by  $\tilde{u}$  on  $\Omega \times (0, T')$  several times and used the definition of  $\tilde{u}$  in the last line.

Let us now take  $a \searrow 0$  in the inequality above. As a consequence of the lower semicontinuity of  $E_\varepsilon$  and the  $L^2$ -continuity of the map  $t \mapsto u(t)$ , we have

$$\liminf_{a \searrow 0} \frac{1}{a} \int_0^a E_\varepsilon[u(t)] dt \geq E_\varepsilon[u(0)] = E_\varepsilon[u_0].$$

Furthermore, it follows from a general fact about Sobolev functions that

$$\frac{\tilde{u}(\cdot + a) - \tilde{u}(\cdot)}{a} \rightharpoonup \partial_t \tilde{u} (= \partial_t u) \quad \text{in } L^2(\Omega \times (0, T)) \text{ as } a \searrow 0$$

(see [34, Lemma 7.23 and proof of Lemma 7.24]). In particular, the integral  $\int_0^{T'} \int_\Omega \frac{\lambda}{2\varepsilon} \left( \frac{\tilde{u}(t+a) - \tilde{u}(t)}{a} \right)^2 dx dt$  is bounded as  $a \searrow 0$ .

Therefore, the limiting inequality of (3.35) as  $a \searrow 0$  reads

$$\int_0^{T'} \int_\Omega \varepsilon g(-\nabla u(t)) (\partial_t u(t))^2 dx dt \geq E_\varepsilon[u_0] - E_\varepsilon[u(T')], \tag{3.36}$$

which is one inequality in the optimal energy dissipation relation (3.4).

For the converse inequality, we proceed analogously by taking  $w(x, t) := \tilde{u}(x, t - a)$  and  $a \searrow 0$ . Using first the fact that  $\frac{\tilde{u}(\cdot - a) - \tilde{u}(\cdot)}{a} \rightharpoonup -\partial_t u$  in  $L^2(\Omega \times (0, T'))$ , and second the inequality  $\liminf_{a \searrow 0} \frac{1}{a} \int_{T'-a}^{T'} E_\varepsilon[u(t)] dt \geq E_\varepsilon[u(T')]$ , we obtain

$$\int_0^{T'} \int_\Omega \varepsilon g(-\nabla u(t)) (\partial_t u(t))^2 dx dt \leq E_\varepsilon[u_0] - E_\varepsilon[u(T')]. \tag{3.37}$$



The combination of (3.36) and (3.37) is precisely the optimal energy dissipation identity (3.4).

For the distributional formulation of the PDE (3.3) let  $\varphi \in C^1(\Omega \times [0, T])$  and  $s > 0$ . By plugging in  $\zeta \equiv 1$  and  $w = u + s\varphi$  into (3.31), we obtain

$$\int_0^T \left( E_\varepsilon[u(t) + s\varphi(t)] - E_\varepsilon[u(t)] + \int_\Omega \varepsilon g(-\nabla u(t)) \partial_t u(t) s \varphi(t) dx + \frac{\lambda}{4\varepsilon} \int_\Omega s^2 \varphi(t)^2 dx \right) dt \geq 0.$$

Dividing by  $s$  and taking  $s \searrow 0$  leads to

$$\begin{aligned} \int_0^T \int_\Omega \varepsilon g(-\nabla u) \partial_t u \varphi \, dx dt &\geq - \liminf_{s \searrow 0} \int_0^T \frac{E_\varepsilon[u(t) + s\varphi(t)] - E_\varepsilon[u(t)]}{s} dt \\ &= -\frac{1}{2} \liminf_{s \searrow 0} \left( \varepsilon \int_0^T \int_\Omega \frac{f(-\nabla u - s\nabla \varphi) - f(-\nabla u)}{s} dx dt + \frac{1}{\varepsilon} \int_0^T \int_\Omega \frac{W(u + s\varphi) - W(u)}{s} dx dt \right) \\ &= \frac{\varepsilon}{2} \int_0^T \int_\Omega Df(-\nabla u) \cdot \nabla \varphi \, dx dt - \frac{1}{2\varepsilon} \int_0^T \int_\Omega W'(u) \varphi \, dx dt, \end{aligned}$$

where the limit and the integrals can be interchanged in the last step because  $u$  is essentially bounded.

We divide by  $\varepsilon$  to obtain one inequality in (3.3). The converse inequality follows by taking  $s < 0$ ,  $s \nearrow 0$ . □

### 3.2 Regularity of weak solutions

The remainder of this section is devoted to proving a regularity result which states that solutions to the anisotropic Allen–Cahn equation have weak second derivatives in space. We apply a difference quotient method for elliptic regularity. The idea for this proof is taken from [26]. The fact that  $f$  is, in general, not twice continuously differentiable on  $\mathbb{R}^d$  due to a singularity at 0 will not pose a problem when proving the existence of second derivatives, but it prevents us from deriving an additional PDE for the second derivatives.

**Theorem 3.10** (Spatial regularity of distributional solutions). *Let  $\varepsilon > 0$ . If  $u$  is a solution to (1.4) in the sense of Definition 3.3, then  $u \in L^2(0, T; H^2(\Omega))$ .*

It suffices to prove the following

**Claim 1.** *Let  $u \in H^1(\Omega)$ , and suppose that  $u$  is a weak solution to  $-\operatorname{div}(Df(-\nabla u)) = h$  for some function  $h \in L^2(\Omega)$ . More precisely, we assume that*

$$\int_\Omega Df(-\nabla u) \cdot \nabla w \, dx = \int_\Omega h w \, dx$$

for all  $w \in H^1(\Omega)$ . Then  $u$  has a weak second derivative in space  $\nabla^2 u \in L^2(\Omega)^{d \times d}$  and  $\|\nabla^2 u\|_{L^2(\Omega)} \leq C \|h\|_{L^2(\Omega)}$ .

Theorem 3.10 follows from this claim by choosing  $h = 2g(-\nabla u) \partial_t u + \frac{1}{\varepsilon^2} W'(u)$  and slicing in time. Then we have  $h(\cdot, t) \in L^2(\Omega)$  for almost every  $t \in (0, T)$  since  $\partial_t u \in L^2(\Omega \times (0, T))$ ,  $u \in L^\infty(\Omega \times (0, T))$ ,

and  $W \in C^1(\mathbb{R})$ . The fact that (3.3) holds true for test functions  $w \in L^2(0, T; H^1(\Omega))$  instead of just  $C^1(\Omega \times [0, T])$  can be shown with a density argument.

**Proof of the claim.** Given a function  $w$  as in the claim, we define the difference quotient

$$D_l^s w(x) := \frac{w(x + se_l) - w(x)}{s}$$

for  $s \neq 0$  and  $l = 1, \dots, d$ . Then  $D_l^s w \in L^2(0, T; H^1(\Omega))$ . The weak formulation of the PDE and the substitution  $y = x + se_l$  yield

$$\begin{aligned} \int_{\Omega} h D_l^s w \, dx &= \int_{\Omega} Df(-\nabla u) \cdot \nabla D_l^s w \, dx \\ &= \int_{\Omega} Df(-\nabla u(x, t)) \cdot \frac{\nabla w(x + se_l) - \nabla w(x)}{s} \, dx \\ &= \int_{\Omega} \frac{Df(-\nabla u(x - se_l)) - Df(-\nabla u(x))}{s} \cdot \nabla w(x) \, dx. \end{aligned} \tag{3.38}$$

In order to deal with the right-hand side, we will use the following identity for  $i = 1, 2, \dots, d$  and for almost every  $x \in \Omega$ :

$$\begin{aligned} \partial_{\xi_i} f(-\nabla u(x - se_i)) - \partial_{\xi_i} f(-\nabla u(x)) &= - \sum_{j=1}^d \int_0^1 \partial_{\xi_i \xi_j}^2 f(-r \nabla u(x - se_i) - (1-r) \nabla u(x)) \, dr \\ &\quad \cdot (\partial_j u(x - se_i) - \partial_j u(x)) \\ &= \sum_{j=1}^d A_{ij}^s(x) (\partial_j u(x) - \partial_j u(x - se_i)), \end{aligned} \tag{3.39}$$

where the shorthand notation in the last line is to be read as

$$A_{ij}^s(x) := \int_0^1 \partial_{\xi_i \xi_j}^2 f(-r \nabla u(x - se_i) - (1-r) \nabla u(x)) \, dr.$$

If the line segment between  $\nabla u(x - se_i)$  and  $\nabla u(x)$  does not contain the origin, then  $f$  is twice continuously differentiable in a neighborhood of the line segment, so that (3.39) follows immediately from the fundamental theorem of calculus. On the other hand, if  $0 \in \text{conv}(\{\nabla u(x - se_i), \nabla u(x, t)\})$ , then (3.39) can be deduced by decomposing the line segment into two parts and applying the fundamental theorem of calculus on each part.

By the strong convexity of  $f$ , there exists a constant  $c > 0$  such that

$$\xi \cdot D^2 f(p) \xi \geq c |\xi|^2$$

for all  $\xi \in \mathbb{R}^d, p \in \mathbb{R}^d \setminus \{0\}$ . Using the definition of  $A_{ij}^s$ , we obtain

$$\nabla D_l^{-s} u \cdot A^s(x) \nabla D_l^{-s} u \geq c |\nabla D_l^{-s} u|^2 \tag{3.40}$$

for almost every  $x \in \Omega$ .

One can now choose  $w = D_l^{-s} u$  in (3.38) and use (3.39) with (3.40) as is standard (see, e.g., [34, Lemma 7.23]) to conclude

$$\int_{\Omega} |\nabla D_l^{-s} u|^2 \, dx \leq C \int_{\Omega} h^2 \, dx < \infty.$$

As  $\nabla D_l^{-s} u \rightharpoonup \partial_l \nabla u$  as  $s \rightarrow 0$ , this concludes the claim. □

#### 4. Convergence of the anisotropic Allen–Cahn equation to anisotropic mean curvature flow

The goal of this section is to prove a conditional convergence result for the sharp-interface limit of the anisotropic Allen–Cahn equation. In order to conclude that the limit is a *BV* solution to (1.1) in the sense of Definition 2.8, a critical assumption is the convergence of the time-integrated energies. A statement of the result and a sketch of the proof are included in the notes [39]. Here, we complete the proof and correct a mistake with respect to the phase-field equation (1.4): In [39], the function  $g$  was defined as  $g(p) = \frac{|p|}{\mu(p)}$ , thereby missing the effect of the anisotropic surface tension  $\sigma$  on  $g$ .

**Theorem 4.1.** *Let  $(\sigma, \mu)$  be an admissible pair of anisotropies, and let  $f, g$  be given by equations (1.5), (1.6), respectively. Suppose that  $W$  is a double-well potential satisfying (W1)–(W5). Let  $\{u_\varepsilon\}_{\varepsilon>0}$  be solutions to the anisotropic Allen–Cahn equation in the sense of Definition 3.3 with initial data  $u_{\varepsilon,0} \in \text{dom}(E_\varepsilon)$  and such that  $\|u_\varepsilon - \frac{1}{2}\|_{L^\infty(\Omega \times (0,T))} \leq \max\{\|u_{\varepsilon,0} - \frac{1}{2}\|_{L^\infty(\Omega)}, \frac{1}{2}\}$ . Assume that*

- the initial conditions are well-prepared,

$$u_{\varepsilon,0} \longrightarrow u_0 \quad \text{in } L^2(\Omega) \quad \text{and} \quad E_\varepsilon[u_{\varepsilon,0}] \longrightarrow E[u_0] =: E_0 < \infty \tag{4.1}$$

as  $\varepsilon \searrow 0$ , and that

- there is a uniform  $L^\infty$ -bound on the initial conditions,

$$\limsup_{\varepsilon \searrow 0} \|u_{\varepsilon,0}\|_{L^\infty(\Omega)} < \infty. \tag{4.2}$$

Then there exists a subsequence  $\varepsilon \searrow 0$  as well as a function

$$u = \chi \in BV(\Omega \times (0, T); \{-1, 1\}) \cap C^{0, \frac{1}{2}}([0, T]; L^1(\Omega)) \cap L^\infty(0, T; BV(\Omega))$$

such that  $u_\varepsilon \rightarrow u$  in  $L^2(\Omega \times (0, T))$  as  $\varepsilon \searrow 0$ .

If, furthermore,

$$\lim_{\varepsilon \searrow 0} \int_0^T E_\varepsilon[u_\varepsilon(t)] dt = \int_0^T E[u(t)] dt, \tag{4.3}$$

then  $u$  is a distributional solution to anisotropic mean curvature flow with initial condition  $u(0) = u_0$  in the sense of Definition 2.8.

Similarly to [47], where a minimizing movements construction for isotropic mean curvature is performed, the intuitive meaning of the assumption of energy convergence (4.3) is that no surface area is lost in the limit.

We will first argue by compactness that a limiting function  $u$  exists. Assumption (4.3) then allows us to prove an equipartition of energy between the anisotropic Dirichlet energy and the nonconvex potential energy as  $\varepsilon \searrow 0$ . In addition, this assumption guarantees the existence of a normal velocity  $V$  in the sense of (2.18). We introduce a relative entropy functional which serves as a tilt excess. Using this tilt excess to bound the occurring error terms, we will derive the weak formulation (2.19) from the distributional formulation (3.3) and the optimal energy dissipation inequality (2.20) from (3.4).

##### 4.1 Compactness

To prove the compactness statement for the solutions  $\{u_\varepsilon\}_{\varepsilon>0}$  to the anisotropic Allen–Cahn equation from Theorem 3.5, we will first show a  $W^{1,1}$ -bound for the compositions of a suitable continuous function with  $u_\varepsilon$ . This argument was performed, for example, by Fonseca and Tartar [27] in the isotropic case

(see also [44]). However, some simplifications are possible since we assume an  $L^\infty$ -bound on the functions  $\{u_\varepsilon\}$  instead of prescribing growth conditions on the double-well potential  $W$ .

**Lemma 4.2.** *Let  $u_\varepsilon$  be solutions of the anisotropic Allen–Cahn equation as in Theorem 3.5. Under the assumptions of Theorem 4.1 except (4.3), there exist a subsequence  $\varepsilon \searrow 0$  and a function  $u = \chi \in BV(\Omega \times (0, T); \{0, 1\}) \cap C^{0, \frac{1}{2}}([0, T]; L^1(\Omega)) \cap L^\infty(0, T; BV(\Omega))$  such that*

- $u_\varepsilon \rightarrow u$  in  $L^2(\Omega \times (0, T))$ ,
- $\phi \circ u_\varepsilon \xrightarrow{*} \phi \circ u = c_0 \chi$  in  $BV(\Omega \times (0, T))$ , and
- $\phi \circ u_\varepsilon(t) \xrightarrow{*} c_0 \chi(t)$  in  $BV(\Omega)$  for all  $t \in [0, T]$

as  $\varepsilon \searrow 0$ , where the function  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  is defined via

$$\phi(z) := \int_0^z \sqrt{W(u)} du. \tag{4.4}$$

The first step of the proof is to show that

$$\limsup_{\varepsilon \searrow 0} \|\phi \circ u_\varepsilon\|_{W^{1,1}(\Omega \times (0, T))} < \infty. \tag{4.5}$$

For this step, we observe that, by assumption (4.2), there exists a finite constant  $R > 0$  such that  $u_\varepsilon(x, t) \in [-R, R]$  for almost every  $(x, t) \in \Omega \times (0, T)$  and for all sufficiently small  $\varepsilon > 0$ . Using the continuity of  $\phi$ , we first estimate

$$\int_0^T \int_\Omega |\phi(u_\varepsilon)| \, dxdt \leq T \sup_{z \in [-R, R]} |\phi(z)| < \infty \tag{4.6}$$

for  $\varepsilon > 0$  small enough.

Furthermore, the chain rule for Sobolev functions is applicable to the compositions  $\phi \circ u_\varepsilon$  even though  $\phi$  need not be globally Lipschitz continuous: By virtue of its definition as a primitive function,  $\phi$  is continuously differentiable. In particular,  $\phi'|_{[-R, R]}$  is bounded, and we can assume without restriction that  $\phi'$  is bounded. Thus, one can estimate

$$\begin{aligned} \int_0^T \int_\Omega |\nabla(\phi \circ u_\varepsilon)| \, dxdt &= \int_0^T \int_\Omega \phi'(u_\varepsilon) |\nabla u_\varepsilon| \, dxdt \\ &\leq \frac{1}{\min_{S^{d-1}} \sigma} \int_0^T \int_\Omega \phi'(u_\varepsilon) \sigma(-\nabla u_\varepsilon) \, dxdt \\ &\leq \frac{1}{2 \min_{S^{d-1}} \sigma} \int_0^T \int_\Omega \left( \varepsilon \sigma(-\nabla u_\varepsilon)^2 + \frac{1}{\varepsilon} \phi'(u_\varepsilon)^2 \right) \, dxdt \\ &= \frac{1}{2 \min_{S^{d-1}} \sigma} \int_0^T \int_\Omega \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \, dxdt \\ &= \frac{1}{\min_{S^{d-1}} \sigma} \int_0^T E_\varepsilon[u_\varepsilon(t)] \, dt \end{aligned} \tag{4.7}$$

for  $\varepsilon > 0$  sufficiently small.

Lastly, a similar argument for small  $\varepsilon$  yields

$$\begin{aligned} \int_0^T \int_{\Omega} |\partial_t(\phi \circ u_\varepsilon)| \, dxdt &= \int_0^T \int_{\Omega} \phi'(u_\varepsilon) |\partial_t u_\varepsilon| \, dxdt \\ &\leq \left( \int_0^T \int_{\Omega} \frac{1}{\varepsilon} W(u_\varepsilon) \, dxdt \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} \varepsilon |\partial_t u_\varepsilon|^2 \, dxdt \right)^{\frac{1}{2}} \\ &\leq \left( 2 \int_0^T E_\varepsilon[u_\varepsilon(t)] \, dt \right)^{\frac{1}{2}} \left( \frac{1}{c_g} \int_0^T \int_{\Omega} \varepsilon g(-\nabla u_\varepsilon) |\partial_t u_\varepsilon|^2 \, dxdt \right)^{\frac{1}{2}} \\ &\leq \sqrt{\frac{2}{c_g} E_\varepsilon[u_{\varepsilon,0}]} \sqrt{T}, \end{aligned} \tag{4.8}$$

where the fourth inequality uses the optimal energy dissipation identity (3.4) in the first and in the second factor.

It is known from assumption (4.1) that

$$\limsup_{\varepsilon \searrow 0} E_\varepsilon[u_{\varepsilon,0}] = E_0 < \infty \quad \text{and} \quad \limsup_{\varepsilon \searrow 0} \int_0^T E_\varepsilon[u_\varepsilon(t)] \, dt \leq \lim_{\varepsilon \searrow 0} TE_\varepsilon[u_{\varepsilon,0}] = TE_0 < \infty,$$

where the second estimate makes use of the fact that  $t \mapsto E_\varepsilon[u_\varepsilon(t)]$  is nonincreasing by (3.4). Thus, the estimates (4.6)–(4.8) suffice to prove (4.5).

In addition to (4.5), we want to derive a uniform  $W^{1,1}$ -estimate for fixed times  $t \in (0, T)$ , namely

$$\limsup_{\varepsilon \searrow 0} \operatorname{ess\,sup}_{t \in (0,T)} \|\phi \circ u_\varepsilon(t)\|_{W^{1,1}(\Omega)} < \infty. \tag{4.9}$$

Indeed, it follows by Fubini’s theorem that  $\|u_\varepsilon(t)\|_{L^\infty} \leq R$  for almost every  $t \in (0, T)$  if  $\varepsilon > 0$  is small enough, and we can argue similarly as in (4.6) and (4.7) to show that

$$\int_{\Omega} |\phi \circ u_\varepsilon(x, t)| \, dx \leq \sup_{z \in [-R,R]} |\phi(z)| < \infty$$

and

$$\int_{\Omega} |\nabla(\phi \circ u_\varepsilon)(x, t)| \, dx \leq \frac{1}{\min_{\sigma} \sigma} E_\varepsilon[u_\varepsilon(t)] \leq \frac{1}{\min_{\sigma} \sigma} E_\varepsilon[u_{\varepsilon,0}] < \infty$$

for a.e.  $t \in (0, T)$  and all  $\varepsilon > 0$  sufficiently small. These two observations prove (4.9).

Bounded sequences in  $BV(\Omega \times (0, T))$  admit a weakly- $*$  convergent subsequence by a version of [2, Theorem 3.23]. Thus, by (4.5), we can find a subsequence  $\varepsilon \searrow 0$  such that  $\phi \circ u_\varepsilon$  converge weakly- $*$  in  $BV(\Omega \times (0, T))$  as  $\varepsilon \searrow 0$ . Let us denote the limiting function by  $v \in BV(\Omega \times (0, T))$ . Then, in particular,  $\phi \circ u_\varepsilon \rightarrow v$  in  $L^1(\Omega \times (0, T))$ , and we can assume without restriction that

$$\phi \circ u_\varepsilon(x, t) \longrightarrow v(x, t) \quad \text{for a.e. } (x, t) \in \Omega \times (0, T) \tag{4.10}$$

and

$$\phi \circ u_\varepsilon(t) \longrightarrow v(t) \quad \text{in } L^1(\Omega) \quad \text{for a.e. } t \in (0, T), \tag{4.11}$$

which can be accomplished by taking a further subsequence. It follows from (4.9), (4.11), and the characterization [2, Proposition 3.13] of weak- $*$  convergence in  $BV$  that  $\phi \circ u_\varepsilon(t) \overset{*}{\rightharpoonup} v(t)$  in  $BV(\Omega)$  for a.e.  $t \in (0, T)$ .

From now on, the limit  $\varepsilon \searrow 0$  is to be understood as the limit along a subsequence  $\{\varepsilon_j\}_{j \in \mathbb{N}}$  such that  $\varepsilon_j \searrow 0$  and  $\phi \circ u_{\varepsilon_j} \overset{*}{\rightharpoonup} v$  in  $BV(\Omega \times (0, T))$  as  $j \rightarrow \infty$ , and such that the pointwise convergence properties as in (4.10) and (4.11) hold true.

The following argument for the Hölder continuity  $v \in C^{0, \frac{1}{2}}([0, T]; L^1(\Omega))$  is adapted from [35, Lemma 2], where Hensel and the first author deal with the Hölder continuity in the isotropic case:

Let  $0 \leq s \leq t \leq T$ . If we proceed as in (4.8), but only integrate from  $s$  to  $t$ , we find

$$\begin{aligned} \int_{\Omega} |\phi(u_{\varepsilon}(x, t)) - \phi(u_{\varepsilon}(x, s))| dx &\leq \int_s^t \int_{\Omega} |\partial_i(\phi \circ u_{\varepsilon})| dx dt \\ &\leq \sqrt{\frac{2}{c_g}} E_{\varepsilon}[u_{\varepsilon,0}] \sqrt{t-s}. \end{aligned} \tag{4.12}$$

There exists a null set  $N \subset (0, T)$  such that, for all  $s, t \in (0, T) \setminus N$ , we can pass to the limit  $\varepsilon \searrow 0$  to obtain

$$\int_{\Omega} |v(x, t) - v(x, s)| dx \leq \sqrt{\frac{2}{c_g}} E_0 \sqrt{t-s}.$$

This allows us to redefine  $v$  on the null set  $\Omega \times N$  so that  $v \in C^{0, \frac{1}{2}}([0, T]; L^1(\Omega))$ .

We identify  $\phi \circ u_{\varepsilon}$  with their Hölder continuous representatives due to (4.12). If a sequence of uniformly Hölder continuous functions converges pointwise almost everywhere, it follows that the sequence converges pointwise. In particular, we can upgrade the  $L^1(\Omega)$ -convergence for a.e.  $t \in (0, T)$  in (4.11) to  $L^1(\Omega)$ -convergence for all  $t \in [0, T]$ .

Using the continuity  $\phi \circ u_{\varepsilon} \in C^{0, \frac{1}{2}}([0, T]; L^1(\Omega))$  and the lower semicontinuity of the variation [2, Remark 3.5], we conclude from (4.9) that

$$\limsup_{\varepsilon \searrow 0} \sup_{t \in [0, T]} \|\phi \circ u_{\varepsilon}(t)\|_{BV(\Omega)} < \infty,$$

i.e., we can also upgrade the essential boundedness of the  $BV$ -norm to uniform boundedness in time. From this, one finds that  $\phi \circ u_{\varepsilon}(t) \overset{*}{\rightharpoonup} v(t)$  in  $BV(\Omega)$  for all  $t \in [0, T]$  and that  $v \in L^{\infty}(0, T; BV(\Omega))$ . (This statement includes the measurability of the measure-valued map  $t \mapsto \nabla v(t)$  in the sense of [2, Definition 2.25], which follows from the fact that we have  $\sup_{t \in [0, T]} \int_{\Omega} |\nabla v(t)| < \infty$  and  $v(t) \rightarrow v(t_0)$  in  $L^1(\Omega)$  as  $t \rightarrow t_0$ , and therefore  $\nabla v(t) \overset{*}{\rightharpoonup} \nabla v(t_0)$  in  $\mathcal{M}(\Omega)$  as  $t \rightarrow t_0$ .)

Let us turn to the convergence result for the solutions  $u_{\varepsilon}$ . As the function  $\sqrt{W}$  is strictly positive except at two isolated points, the primitive function  $\phi$  is strictly increasing. Therefore, there exists a continuous inverse  $\phi^{-1}$ , and we obtain from (4.10) that

$$u_{\varepsilon}(x, t) \longrightarrow u(x, t) := \phi^{-1}(v(x, t)) \quad \text{as } \varepsilon \searrow 0 \quad \text{for almost all } (x, t) \in \Omega \times (0, T).$$

By (4.2) and the dominated convergence theorem, we may conclude  $u_{\varepsilon} \rightarrow u$  in  $L^2(\Omega \times (0, T))$  as  $\varepsilon \searrow 0$ .

To see that the limiting function  $u$  takes values in  $\{0, 1\}$  it suffices to use Fatou’s lemma:

$$\int_0^T \int_{\Omega} W(u) dx dt \leq \liminf_{\varepsilon \rightarrow 0} \int_0^T \int_{\Omega} W(u_{\varepsilon}) dx dt \leq \liminf_{\varepsilon} \int_0^T E_{\varepsilon}[u_{\varepsilon}(\cdot, t)] dt = 0.$$

Consequently, we have  $W(u) = 0$ , i.e.,  $u \in \{0, 1\}$  almost everywhere in  $\Omega \times (0, T)$ .

The representation  $v = \phi \circ u = c_0 \chi$  now follows from  $\phi(0) = 0$  and  $\phi(1) = c_0$ . Since  $v \in BV(\Omega \times (0, T); \mathbb{R}) \cap C^{0, \frac{1}{2}}([0, T]; L^1(\Omega)) \cap L^{\infty}(0, T; BV(\Omega))$ , it follows that the same holds true for  $u = \frac{1}{c_0} v$ .

### 4.2 Equipartition of energy

The following theorem states that, asymptotically as  $\varepsilon \searrow 0$ , the anisotropic Dirichlet energy  $\frac{1}{2} \int_0^T \int_{\Omega} \varepsilon f(-\nabla u_{\varepsilon}) dx dt$  and the nonconvex term  $\frac{1}{2} \int_0^T \int_{\Omega} \frac{1}{\varepsilon} W(u_{\varepsilon}) dx dt$  contribute equally to the Cahn-Hilliard energy. The equipartition of energy also holds true in a localized form. The argument relies crucially on the energy convergence assumption (4.3) and uses a trick introduced by Modica–Mortola for Gamma-convergence of the energies [49] (see also [8]). Equipartition results of this kind have been used to prove conditional convergence in the static or dynamic case since [46, 48].

Statements (i) and (ii) can be viewed primarily as preparatory results for the equipartition statements (iii)–(v).

**Theorem 4.3.** *Under the assumptions of Theorem 4.1, including the energy convergence (4.3), the following convergence statements hold true in the limit  $\varepsilon \searrow 0$ :*

(i)

$$\frac{1}{2} \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \xrightarrow{*} c_0 \sigma(v) |\nabla \chi| \quad \text{in } \mathcal{M}(\Omega \times [0, T]), \tag{4.13}$$

(ii)

$$\sigma(-\nabla(\phi \circ u_\varepsilon)) \xrightarrow{*} c_0 \sigma(v) |\nabla \chi| \quad \text{in } \mathcal{M}(\Omega \times [0, T]), \tag{4.14}$$

(iii)

$$\sqrt{\varepsilon f(-\nabla u_\varepsilon)} - \sqrt{\frac{1}{\varepsilon} W(u_\varepsilon)} \longrightarrow 0 \quad \text{in } L^2(\Omega \times (0, T)), \tag{4.15}$$

(iv)

$$\frac{1}{2} \left( \varepsilon f(-\nabla u_\varepsilon) - \frac{1}{\varepsilon} W(u_\varepsilon) \right) \longrightarrow 0 \quad \text{in } L^1(\Omega \times (0, T)), \tag{4.16}$$

(v)

$$\varepsilon f(-\nabla u_\varepsilon) \xrightarrow{*} c_0 \chi \quad \text{and} \quad \frac{1}{\varepsilon} W(u_\varepsilon) \xrightarrow{*} c_0 \sigma(v) |\nabla \chi| \quad \text{in } \mathcal{M}(\Omega \times [0, T]). \tag{4.17}$$

**Proof.**

(i) It suffices to show that

$$\frac{1}{2} \int_0^T \int_\Omega \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \zeta \, dx dt \rightarrow c_0 \int_0^T \int_\Omega \sigma(v) \zeta |\nabla \chi|$$

for all  $\zeta \in C^1(\Omega \times [0, T])$  with  $0 \leq \zeta \leq 1$ . By a density and linearity argument, the same convergence then holds true for all  $\zeta \in C(\Omega \times [0, T])$ .

If  $\zeta \in C^1(\Omega \times [0, T])$  and  $0 \leq \zeta \leq 1$ , we can use Young’s inequality and the chain rule for Sobolev functions to estimate

$$\begin{aligned} \frac{1}{2} \int_0^T \int_\Omega \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \zeta \, dx dt &\geq \int_0^T \int_\Omega \sqrt{f(-\nabla u_\varepsilon) W(u_\varepsilon)} \zeta \, dx dt \\ &= \int_0^T \int_\Omega \sigma(-\phi'(u_\varepsilon) \nabla u_\varepsilon) \zeta \, dx dt \\ &= \int_0^T \int_\Omega \sigma(-\nabla(\phi \circ u_\varepsilon)) \zeta \, dx dt. \end{aligned}$$

By Lemma 2.5, we have that

$$\int_0^T \int_\Omega \sigma(B) \, dx dt = \sup_{\substack{\eta \in C^1(\Omega \times [0, T]) \\ \sigma^2(\eta) \leq 1}} \int_0^T \int_\Omega B \cdot \eta \, dx dt \tag{4.18}$$

for all  $B \in L^1(\Omega \times (0, T))^d$ .

We apply (4.18) and the  $L^1$ -convergence  $\phi \circ u_\varepsilon \rightarrow c_0\chi$  as  $\varepsilon \searrow 0$  to the above estimate, which, recalling Lemma 2.3, yields

$$\begin{aligned}
 & \liminf_{\varepsilon \searrow 0} \frac{1}{2} \int_0^T \int_\Omega \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \zeta \, dxdt \\
 & \geq \liminf_{\varepsilon \searrow 0} \int_0^T \int_\Omega \sigma(-\nabla(\phi \circ u_\varepsilon)) \zeta \, dxdt \\
 & = \liminf_{\varepsilon \searrow 0} \sup_\eta \int_0^T \int_\Omega \eta \cdot (-\nabla(\phi \circ u_\varepsilon)) \zeta \, dxdt \\
 & = \liminf_{\varepsilon \searrow 0} \sup_\eta \int_0^T \int_\Omega (\phi \circ u_\varepsilon) \operatorname{div}(\zeta \eta) \, dxdt \\
 & \geq c_0 \sup_\eta \int_0^T \int_\Omega \chi \operatorname{div}(\zeta \eta) \, dxdt \\
 & = -c_0 \sup_\eta \int_0^T \int_\Omega \zeta \eta \cdot \frac{\nabla \chi}{|\nabla \chi|} |\nabla \chi| \\
 & = c_0 \sup_\eta \int_0^T \int_\Omega \zeta \eta \cdot \nu |\nabla \chi| \\
 & = c_0 \int_0^T \int_\Omega \sigma(\nu) \zeta |\nabla \chi|, \tag{4.19}
 \end{aligned}$$

where the supremum is taken over all  $\eta \in C^1(\Omega \times [0, T])^d$  such that  $\sigma^\circ(\eta) \leq 1$ . To prove the estimate from above, we observe that (4.19) also applies to the function  $1 - \zeta$  instead of  $\zeta$ , and use the energy convergence assumption (4.3). Indeed,

$$\begin{aligned}
 & \limsup_{\varepsilon \searrow 0} \frac{1}{2} \int_0^T \int_\Omega \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \zeta \, dxdt \\
 & = \limsup_{\varepsilon \searrow 0} \left( \frac{1}{2} \int_0^T \int_\Omega \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \, dxdt \right. \\
 & \quad \left. - \frac{1}{2} \int_0^T \int_\Omega \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) (1 - \zeta) \, dxdt \right) \\
 & = \lim_{\varepsilon \searrow 0} \int_0^T E_\varepsilon[u_\varepsilon(t)] \, dt \\
 & \quad - \liminf_{\varepsilon \searrow 0} \frac{1}{2} \int_0^T \int_\Omega \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) (1 - \zeta) \, dxdt \\
 & \leq \int_0^T E[u(t)] \, dt - c_0 \int_0^T \int_\Omega \sigma(\nu) (1 - \zeta) |\nabla \chi| \\
 & = c_0 \int_0^T \int_\Omega \sigma(\nu) \zeta |\nabla \chi|. \tag{4.20}
 \end{aligned}$$

The combination of the two inequalities (4.19) and (4.20) yields the first claim.



(ii) In the same manner as in (i), it suffices to show that

$$\int_0^T \int_{\Omega} \sigma(-\nabla(\phi \circ u_\varepsilon)) \zeta \, dxdt \rightarrow c_0 \int_0^T \int_{\Omega} \sigma(v) \zeta |\nabla \chi|$$

whenever  $\zeta \in C^1(\Omega \times [0, T])$  and  $0 \leq \zeta \leq 1$ .

The lim inf inequality follows from the chain of inequalities in (4.19).

For the lim sup inequality, we can use Young’s inequality and (i) to compute

$$\begin{aligned} \limsup_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \sigma(-\nabla(\phi \circ u_\varepsilon)) \zeta \, dxdt &\leq \lim_{\varepsilon \searrow 0} \frac{1}{2} \int_0^T \int_{\Omega} \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \zeta \, dxdt \\ &= c_0 \int_0^T \int_{\Omega} \sigma(v) \zeta |\nabla \chi|. \end{aligned}$$

(iii) By taking  $\zeta \equiv 1$  as test functions for the weak\* limits in (i) and (ii), we see that

$$\begin{aligned} \lim_{\varepsilon \searrow 0} \frac{1}{2} \int_0^T \int_{\Omega} \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \, dxdt &= \lim_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \sigma(-\nabla(\phi \circ u_\varepsilon)) \, dxdt \\ &= c_0 \int_0^T \int_{\Omega} \sigma(v) |\nabla \chi|. \end{aligned} \tag{4.21}$$

This observation allows us to compute

$$\begin{aligned} &\int_0^T \int_{\Omega} \left( \sqrt{\varepsilon f(-\nabla u_\varepsilon)} - \sqrt{\frac{1}{\varepsilon} W(u_\varepsilon)} \right)^2 \, dxdt \\ &= \int_0^T \int_{\Omega} \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \, dxdt - 2 \int_0^T \int_{\Omega} \sqrt{f(-\nabla u_\varepsilon) W(u_\varepsilon)} \, dxdt \\ &= \int_0^T \int_{\Omega} \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \, dxdt - 2 \int_0^T \int_{\Omega} \sigma(-\nabla(\phi \circ u_\varepsilon)) \, dxdt \\ &\rightarrow 0 \quad \text{as } \varepsilon \searrow 0. \end{aligned}$$

(iv) Young’s inequality yields

$$\begin{aligned} \left| \varepsilon f(-\nabla u_\varepsilon) - \frac{1}{\varepsilon} W(u_\varepsilon) \right| &\leq \delta \left( \sqrt{\varepsilon f(-\nabla u_\varepsilon)} + \sqrt{\frac{1}{\varepsilon} W(u_\varepsilon)} \right)^2 \\ &\quad + \frac{1}{4\delta} \left( \sqrt{\varepsilon f(-\nabla u_\varepsilon)} - \sqrt{\frac{1}{\varepsilon} W(u_\varepsilon)} \right)^2 \end{aligned}$$

for all  $\delta > 0$ . Taking the limit  $\varepsilon \searrow 0$  and using (iii) as well as (4.21), we obtain

$$\begin{aligned} \limsup_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \left| \varepsilon f(-\nabla u_\varepsilon) - \frac{1}{\varepsilon} W(u_\varepsilon) \right| \, dxdt &\leq \delta \limsup_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \left( \sqrt{\varepsilon f(-\nabla u_\varepsilon)} + \sqrt{\frac{1}{\varepsilon} W(u_\varepsilon)} \right)^2 \, dxdt \\ &= \delta \limsup_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) + 2\sigma(-\nabla(\phi \circ u_\varepsilon)) \right) \, dxdt \\ &= 4\delta c_0 \int_0^T \int_{\Omega} \sigma(v) |\nabla \chi|. \end{aligned}$$

Since  $\delta > 0$  is arbitrary, it follows that

$$\lim_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \left| \varepsilon f(-\nabla u_\varepsilon) - \frac{1}{\varepsilon} W(u_\varepsilon) \right| dxdt = 0.$$

(v) The two convergence results follow from adding and subtracting (i) and (iv), respectively. □

### 4.3 Construction of the normal velocity

This construction follows the argument in the first step of [42, Proposition 2.10], where it was carried out by Simon and the first author for the (multiphase) isotropic case. The idea is to introduce the velocity as a Radon–Nikodým density  $V := \frac{\partial_t \chi}{|\nabla \chi|}$ .

For simplicity of notation, we denote the energy density of the anisotropic Cahn–Hilliard energy by

$$e_\varepsilon(u, x) := \frac{\varepsilon}{2} f(-\nabla u(x)) + \frac{1}{2\varepsilon} W(u(x)).$$

**Lemma 4.4.** *Under the assumptions of Theorem 4.1, including the energy convergence (4.3), there exists a  $|\nabla \chi|$ -measurable normal velocity  $V : \Omega \times (0, T) \rightarrow \mathbb{R}$  satisfying (2.17) and (2.18).*

The first step of the proof is to show that the Radon–Nikodým theorem is applicable: Given a smooth test function  $\zeta \in C_c^\infty(\Omega \times (0, T))$ , the definition of distributional derivatives and an application of the Cauchy–Schwarz inequality yield

$$\begin{aligned} \left| \int_0^T \int_{\Omega} (\phi \circ u_\varepsilon) \partial_t \zeta dxdt \right| &= \left| \int_0^T \int_{\Omega} \partial_t (\phi \circ u_\varepsilon) \zeta dxdt \right| \\ &= \left| \int_0^T \int_{\Omega} \phi'(u_\varepsilon) \partial_t u_\varepsilon \zeta dxdt \right| \\ &\leq \left( \int_0^T \int_{\Omega} \varepsilon |\partial_t u_\varepsilon|^2 dxdt \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} \frac{1}{\varepsilon} W(u_\varepsilon) \zeta^2 dxdt \right)^{\frac{1}{2}} \\ &\leq \sqrt{2} \left( \int_0^T \int_{\Omega} \varepsilon |\partial_t u_\varepsilon|^2 dxdt \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} e_\varepsilon(u_\varepsilon(t), x) \zeta^2 dxdt \right)^{\frac{1}{2}}. \end{aligned} \tag{4.22}$$

It follows from (4.13) that

$$\lim_{\varepsilon \rightarrow 0} \int_0^T \int_{\Omega} e_\varepsilon(u_\varepsilon(t), x) \zeta^2 dxdt = c_0 \int_0^T \int_{\Omega} \zeta^2 \sigma(v) |\nabla \chi|. \tag{4.23}$$

Thus, by taking the limit inferior on both sides of (4.22) and recalling that  $\phi \circ u_\varepsilon \rightarrow c_0 \chi$  in  $L^1(\Omega \times (0, T))$  as  $\varepsilon \searrow 0$ , we find

$$\begin{aligned} c_0 \left| \int_0^T \int_{\Omega} \chi \partial_t \zeta dxdt \right| &\leq \sqrt{2} \liminf_{\varepsilon \rightarrow 0} \left( \int_0^T \int_{\Omega} \varepsilon |\partial_t u_\varepsilon|^2 dxdt \right)^{\frac{1}{2}} \left( c_0 \int_0^T \int_{\Omega} \zeta^2 \sigma(v) |\nabla \chi| \right)^{\frac{1}{2}} \\ &\leq \left( \frac{2E_0}{c_g} \right)^{\frac{1}{2}} \left( c_0 \int_0^T \int_{\Omega} \zeta^2 \sigma(v) |\nabla \chi| \right)^{\frac{1}{2}}, \end{aligned} \tag{4.24}$$

where the second inequality follows from the optimal energy dissipation inequality (3.4).

It is desirable to reformulate (4.24) in terms of open sets instead of test functions: Given an open set  $A \subseteq \Omega \times (0, T)$ , let us maximize the left-hand side of (4.24) over all  $\zeta \in C_c^\infty(A)$  with  $|\zeta| \leq 1$ .

This provides us with the inequality

$$\begin{aligned}
 |\partial_t \chi|(A) &\leq \left(\frac{2E_0}{c_0 c_g}\right)^{\frac{1}{2}} \sup_{\substack{\zeta \in C_c^\infty(A) \\ |\zeta| \leq 1}} \left(\int_0^T \int_\Omega \zeta^2 \sigma(v) |\nabla \chi|\right)^{\frac{1}{2}} \\
 &\leq \left(\frac{2E_0 \max_{S^{d-1}} \sigma}{c_0 c_g}\right)^{\frac{1}{2}} \sqrt{|\nabla \chi|(A)}.
 \end{aligned}
 \tag{4.25}$$

Making use of the outer regularity of Radon measures, we see that

$$|\partial_t \chi|(E) = \inf_{\substack{A \supseteq E \\ A \text{ open}}} |\partial_t \chi|(A) \leq \left(\frac{2E_0 \max_{S^{d-1}} \sigma}{c_0 c_g}\right)^{\frac{1}{2}} \inf_{\substack{A \supseteq E \\ A \text{ open}}} \sqrt{|\nabla \chi|(A)} = 0$$

for all Borel sets  $E \subseteq \Omega \times (0, T)$  such that  $|\nabla \chi|(E) = 0$ , i.e.,  $\partial_t \chi \ll |\nabla \chi|$ . By the Radon–Nikodým theorem, there exists a  $|\nabla \chi|$ -measurable function  $V$  such that  $\partial_t \chi = V|\nabla \chi|$  on the open set  $\Omega \times (0, T)$ .

In order to prove the square integrability (2.17) of  $V$ , we go back to (4.24). Let us first fix a finite number  $M > 0$  and find a sequence  $\{\zeta_k\}_{k \in \mathbb{N}}$  of smooth test functions such that  $\zeta_k \rightarrow V\chi_{\{|V| \leq M\}}$  in  $L^2(|\nabla \chi|)$  and  $|\zeta_k| \leq M$  for all  $k \in \mathbb{N}$ . Then it follows by dominated convergence that  $V\zeta_k \rightarrow V^2\chi_{\{|V| \leq M\}}$  in  $L^1(|\nabla \chi|)$  and, therefore,

$$\int_0^T \int_\Omega \chi \partial_t \zeta_k dxdt = - \int_0^T \int_\Omega \zeta_k \partial_t \chi = - \int_0^T \int_\Omega V \zeta_k |\nabla \chi| \longrightarrow - \int_0^T \int_\Omega V^2 \chi_{\{|V| \leq M\}} |\nabla \chi|,$$

whereas the  $L^2$ -convergence gives

$$\int_0^T \int_\Omega \zeta_k^2 \sigma(v) |\nabla \chi| \longrightarrow \int_0^T \int_\Omega V^2 \chi_{\{|V| \leq M\}} \sigma(v) |\nabla \chi|.$$

Plugging in  $\zeta_k$  and taking the limit  $k \rightarrow \infty$  in (4.24), we now obtain

$$\begin{aligned}
 c_0 \int_0^T \int_\Omega V^2 \chi_{\{|V| \leq M\}} |\nabla \chi| &\leq \left(\frac{2E_0}{c_g}\right)^{\frac{1}{2}} \left(c_0 \int_0^T \int_\Omega V^2 \chi_{\{|V| \leq M\}} \sigma(v) |\nabla \chi|\right)^{\frac{1}{2}} \\
 &\leq \left(\frac{2E_0 \max_{S^{d-1}} \sigma}{c_g}\right)^{\frac{1}{2}} \left(c_0 \int_0^T \int_\Omega V^2 \chi_{\{|V| \leq M\}} |\nabla \chi|\right)^{\frac{1}{2}}.
 \end{aligned}$$

By rearranging this inequality and taking  $M \rightarrow \infty$  with the help of the monotone convergence theorem, we find the desired integrability statement, namely

$$\int_0^T \int_\Omega V^2 |\nabla \chi| \leq \frac{2E_0 \max_{S^{d-1}} \sigma}{c_0 c_g} < \infty.
 \tag{4.26}$$

The final step is to prove the identity (2.18) (see also [35, Lemma 7]). Given a test function  $\zeta \in C^1(\Omega \times [0, T])$  and an intermediate time  $T' \in (0, T]$ , we introduce a cutoff  $\eta_\alpha$  in time such that

$$\eta_\alpha \in C^1(\mathbb{R}; [0, 1]), \quad \eta_\alpha \equiv 0 \quad \text{on } (-\infty, \frac{\alpha}{2}] \cup [T' - \frac{\alpha}{2}, \infty), \quad \eta_\alpha \equiv 1 \quad \text{on } [\alpha, T' - \alpha],$$

and set  $\zeta_\alpha(x, t) := \zeta(x, t)\eta_\alpha(t)$ . Then, in particular, we have  $\zeta_\alpha \in C_c^1(\Omega \times (0, T))$  for sufficiently small  $\alpha > 0$ . The definition of  $V$  as a Radon–Nikodým density yields

$$\begin{aligned}
 \int_0^{T'} \int_\Omega \zeta_\alpha(x, t) V(x, t) |\nabla \chi| &= - \int_0^{T'} \int_\Omega \chi(x, t) \partial_t \zeta_\alpha(x, t) dxdt \\
 &= - \int_0^{T'} \int_\Omega \chi(x, t) \eta_\alpha(t) \partial_t \zeta(x, t) dxdt \\
 &\quad - \int_0^{T'} \int_\Omega \chi(x, t) \zeta(x, t) dx \eta'_\alpha(t) dt.
 \end{aligned}
 \tag{4.27}$$

In the limit  $\alpha \searrow 0$ , the left-hand side integral and the first right-hand side integral converge due to the dominated convergence theorem. For the last integral, we observe that  $\eta'_\alpha \xrightarrow{*} \delta_0 - \delta_{T'}$  in  $\mathcal{M}(\mathbb{R})$  as  $\alpha \searrow 0$ . Furthermore, we have  $\chi = u \in C^{0, \frac{1}{2}}([0, T]; L^1(\Omega))$  by Lemma 4.2, and  $\zeta$  is uniformly continuous, so that the map  $t \mapsto \int_\Omega \chi(x, t)\zeta(x, t)dx$  is continuous on  $[0, T]$ . We can therefore compute the limits of all three integrals and obtain

$$\int_0^{T'} \int_\Omega \zeta(x, t)V(x, t)|\nabla\chi| = - \int_0^{T'} \int_\Omega \chi(x, t)\partial_t\zeta(x, t)dxdt - \int_\Omega \chi_0(x)\zeta(x, 0) dx + \int_\Omega \chi(x, T')\zeta(x, T') dx,$$

which proves (2.18).

### 4.4 Relative entropies

A key ingredient to derive the sharp interface limit is the following notions of relative entropies.

#### Definition 4.5.

- (i) Let  $u = \chi \in BV(\Omega)$  take values in  $\{0, 1\}$  almost everywhere. As in the definition of the anisotropic surface energy, let  $\nu := -\frac{\nabla\chi}{|\nabla\chi|}$  be the measure-theoretic outer unit normal. The relative entropy of  $u$  with respect to a vector field  $\xi \in C(\Omega)^d$  is

$$\mathcal{E}[u|\xi] := c_0 \int_\Omega (\sigma(\nu) - |\xi|\psi(|\xi|)D\sigma(\xi) \cdot \nu) |\nabla\chi|, \tag{4.28}$$

where  $\psi \in C^\infty([0, \infty))$  is a cutoff function such that the three properties in (2.2) hold true.

- (ii) Let  $\varepsilon > 0$  and  $u_\varepsilon \in H^1(\Omega) \cap L^\infty(\Omega)$ . The  $\varepsilon$ -relative entropy of  $u_\varepsilon$  with respect to a vector field  $\xi \in C(\Omega)^d$  is

$$\mathcal{E}_\varepsilon [u_\varepsilon|\xi] := \int_\Omega (\sigma(-\nabla u_\varepsilon) + |\xi|\psi(|\xi|)D\sigma(\xi) \cdot \nabla u_\varepsilon) \sqrt{W(u_\varepsilon)}dx. \tag{4.29}$$

#### Remark 4.6.

- (i) For  $u = \chi$  as in Definition 4.5(i), we can define the set of finite perimeter  $A := \{x \in \Omega \mid u(x) = 1\}$  (up to a null set). By De Giorgi’s structure theorem (see [2, Theorem 3.59]), one can replace the total variation measure  $|\nabla\chi|$  with the  $(d - 1)$ -dimensional Hausdorff measure and write

$$\mathcal{E}[u|\xi] = c_0 \int_{\partial^*A} (\sigma(\nu) - |\xi|\psi(|\xi|)D\sigma(\xi) \cdot \nu) d\mathcal{H}^{d-1},$$

with  $\partial^*$  denoting the reduced boundary of a set of finite perimeter.

- (ii) One can also show that  $|\nabla\chi(t)| \otimes \mathcal{L}^1 \llcorner (0, T) = |\nabla\chi|$  as Radon measures on  $\Omega \times (0, T)$ . Thus, integrating the relative entropy over time yields

$$\begin{aligned} \int_0^T \mathcal{E}[u(t)|\xi(t)]dt &= c_0 \int_0^T \int_\Omega (\sigma(\nu) - |\xi|\psi(|\xi|)D\sigma(\xi) \cdot \nu) |\nabla\chi(t)|dt \\ &= c_0 \int_0^T \int_\Omega (\sigma(\nu) - |\xi|\psi(|\xi|)D\sigma(\xi) \cdot \nu) |\nabla\chi|. \end{aligned}$$

If  $\xi$  is chosen to be a smooth approximation of the normal, then these relative entropies serve as tilt excesses and can be used to control quadratic error terms: Let  $\xi \in C(\Omega \times [0, T])^d$  such that  $|\xi| \leq 1$  on  $\Omega \times [0, T]$ . By Lemma 2.4(i), we have

$$\int_0^T \int_\Omega |\xi - \nu|^2 |\nabla\chi| \leq \frac{1}{c_0 c_\sigma} \int_0^T \mathcal{E}[u(t)|\xi(t)]dt. \tag{4.30}$$

The link between the relative entropy  $\mathcal{E}$  and the phase-field version  $\mathcal{E}_\varepsilon$  is the following convergence statement:

**Lemma 4.7.** *Under the assumptions of Theorem 4.1, including the energy convergence (4.3), we have*

$$\lim_{\varepsilon \searrow 0} \int_0^T \mathcal{E}_\varepsilon [u_\varepsilon(t)|\xi(t)]dt = \int_0^T \mathcal{E} [u(t)|\xi(t)]dt$$

for every vector field  $\xi \in C(\Omega \times [0, T])^d$  such that  $|\xi| \leq 1$  on  $\Omega \times [0, T]$ , where  $u = \chi$  is the limiting function constructed in Lemma 4.2.

**Proof.** A direct computation yields

$$\begin{aligned} \int_0^T \mathcal{E}_\varepsilon [u_\varepsilon(t)|\xi(t)]dt &= \int_0^T \int_\Omega (\sigma(-\nabla u_\varepsilon) + |\xi|\psi(|\xi|)D\sigma(\xi) \cdot \nabla u_\varepsilon) \sqrt{W(u_\varepsilon)} dxdt \\ &= \int_0^T \int_\Omega \sigma(-\nabla(\phi \circ u_\varepsilon)) dxdt \\ &\quad + \int_0^T \int_\Omega |\xi|\psi(|\xi|)D\sigma(\xi) \cdot \nabla(\phi \circ u_\varepsilon) dxdt \\ &\rightarrow c_0 \int_0^T \sigma(v)|\nabla\chi| + c_0 \int_0^T \int_\Omega |\xi|\psi(|\xi|)D\sigma(\xi) \cdot \nabla\chi \\ &= \int_0^T \mathcal{E} [u(t)|\xi(t)]dt, \end{aligned} \tag{4.31}$$

where the first term converges due to Theorem 4.3(ii). The convergence of the second term follows from the observation that  $\nabla(\phi \circ u_\varepsilon) \xrightarrow{*} c_0 \nabla\chi$  in  $\mathcal{M}(\Omega \times [0, T])$ , which is a consequence of the bound in (4.7) and the  $L^1$ -convergence  $\phi \circ u_\varepsilon \rightarrow c_0\chi$ .  $\square$

While we have only used the first inequality in Lemma 2.4 so far, the second inequality helps us to show that the tilt excess can be made arbitrarily small by approximating the normal  $\nu$  with suitable vector fields  $\xi$ .

**Lemma 4.8.** *For every  $\delta > 0$ , there exists a smooth vector field  $\xi \in C^1(\Omega \times [0, T])^d$  such that  $|\xi| \leq 1$  in  $\Omega \times [0, T]$  and*

$$\int_0^T \mathcal{E} [u(t)|\xi(t)]dt < \delta.$$

**Proof.** The outer unit normal  $\nu$  is  $|\nabla\chi|$ -measurable, and the estimate

$$\int_0^T \int_\Omega |\nu|^2 |\nabla\chi| = \int_0^T \int_\Omega |\nabla\chi| \leq \frac{1}{c_0 \min_{\mathbb{S}^{d-1}} \sigma} \int_0^T E[u(t)]dt < \infty$$

shows that  $\nu \in L^2(|\nabla\chi|)$ . Since  $|\nabla\chi|$  is a Radon measure on  $\Omega \times [0, T]$ , there exists an approximating sequence  $\{\xi_n\}_{n \in \mathbb{N}} \subset C^\infty(\Omega \times [0, T])$  such that  $\xi_n \rightarrow \nu$  in  $L^2(|\nabla\chi|)$  as  $n \rightarrow \infty$ .

Clearly,  $\nu$  takes values in the closed convex set  $\bar{B}_1 = \{p \in \mathbb{R}^d \mid |p| \leq 1\}$  almost everywhere with respect to  $|\nabla\chi|$ , so we can choose  $\xi_n$  in a way that ensures that  $|\xi_n| \leq 1$  on  $\Omega \times [0, T]$  for all  $n \in \mathbb{N}$ .

By Lemma 2.4(ii) and the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}
 \int_0^T \mathcal{E}[u(t)|\xi_n(t)] dt &\leq c_0 C_\sigma \int_0^T \int_\Omega (|v - \xi_n|^2 + (|v| + |\xi_n|)) |\nabla \chi| \\
 &\leq c_0 C_\sigma \int_0^T \int_\Omega |v - \xi_n|^2 |\nabla \chi| \\
 &\quad + c_0 C_\sigma \left( \int_0^T \int_\Omega |\nabla \chi| \right)^{\frac{1}{2}} \left( \int_0^T \int_\Omega |v - \xi_n|^2 |\nabla \chi| \right)^{\frac{1}{2}} \\
 &\leq c_0 C_\sigma \|v - \xi_n\|_{L^2(\nabla \chi)}^2 \\
 &\quad + \sqrt{\frac{c_0}{\min_{S^{d-1}} \sigma}} C_\sigma \left( \int_0^T E[u(t)] dt \right)^{\frac{1}{2}} \|v - \xi_n\|_{L^2(\nabla \chi)} \\
 &\rightarrow 0
 \end{aligned}$$

as  $n \rightarrow \infty$ . □

### 4.5 Convergence of the curvature term

The following two subsections are dedicated to proving the distributional law of anisotropic mean curvature flow (2.19) for the limit function  $u$ . The strategy is to use  $B \cdot \varepsilon \nabla u_\varepsilon$  as a test function in the distributional formulation of the anisotropic Allen–Cahn equation (3.3), where  $B \in C^1(\Omega \times [0, T])^d$ , and then pass to the limit as  $\varepsilon \searrow 0$  on both sides separately.

Here and in the following subsections,  $C$  denotes a generic positive constant that may depend on the pair of anisotropies  $(\sigma, \mu)$ , the double-well potential  $W$ , the time horizon  $T$ , and the cutoff function  $\psi$ . The constant  $C$  is not necessarily the same on every occurrence.

**Theorem 4.9.** *Under the assumptions of Theorem 4.1, including the energy convergence (4.3), we have*

$$\begin{aligned}
 \lim_{\varepsilon \searrow 0} \int_0^T \int_\Omega \frac{1}{2} \left( Df(-\nabla u_\varepsilon) \cdot \nabla(B \cdot \varepsilon \nabla u_\varepsilon) - \frac{1}{\varepsilon^2} W'(u_\varepsilon) B \cdot \varepsilon \nabla u_\varepsilon \right) dxdt \\
 = c_0 \int_0^T \int_\Omega \nabla B : (\sigma(v)I_d - v \otimes D\sigma(v)) |\nabla \chi|
 \end{aligned} \tag{4.32}$$

for all  $B \in C^1(\Omega \times [0, T])^d$ .

This theorem was first proved by Cicalese, Nagase, and Pisante in [18, Theorem 3.3]. However, we proceed with the alternative strategy of proof proposed in [39, Proposition 4.5].

We define the energy-stress tensor  $T_\varepsilon$  by

$$T_\varepsilon := \frac{1}{2} \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) I_d + \varepsilon \nabla u_\varepsilon \otimes \frac{1}{2} Df(-\nabla u_\varepsilon). \tag{4.33}$$

With this definition, the statement of Theorem 4.9 can be rewritten as a weak-\* convergence claim for the energy-stress tensor: Indeed, an integration by parts on the flat torus  $\Omega$  yields

$$\begin{aligned}
 \int_0^T \int_\Omega \nabla B : T_\varepsilon dxdt &= - \int_0^T \int_\Omega B \cdot \operatorname{div} T_\varepsilon dxdt \\
 &= - \frac{1}{2} \int_0^T \int_\Omega B \cdot \nabla \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) dxdt
 \end{aligned}$$

$$\begin{aligned}
 & -\frac{1}{2} \int_0^T \int_{\Omega} \operatorname{div}(Df(-\nabla u_\varepsilon))B \cdot \varepsilon \nabla u_\varepsilon \, dxdt \\
 & -\frac{1}{2} \int_0^T \int_{\Omega} B \cdot \varepsilon \nabla^2 u_\varepsilon Df(-\nabla u_\varepsilon) \, dxdt \\
 & = -\frac{1}{2} \int_0^T \int_{\Omega} B \cdot \frac{1}{\varepsilon} \nabla (W(u_\varepsilon)) \, dxdt \\
 & -\frac{1}{2} \int_0^T \int_{\Omega} \operatorname{div}(Df(-\nabla u_\varepsilon))B \cdot \varepsilon \nabla u_\varepsilon \, dxdt \\
 & = \int_0^T \int_{\Omega} \frac{1}{2} \left( Df(-\nabla u_\varepsilon) \cdot \nabla (B \cdot \varepsilon \nabla u_\varepsilon) - \frac{1}{\varepsilon^2} W'(u_\varepsilon) B \cdot \varepsilon \nabla u_\varepsilon \right) \, dxdt. \tag{4.34}
 \end{aligned}$$

We observe that the right-hand side is exactly the term appearing in (4.32). Thus, in order to prove Theorem 4.9, it suffices to show that  $T_\varepsilon \xrightarrow{*} c_0 (\sigma(v)I_d - v \otimes D\sigma(v)) |\nabla \chi|$  as  $\mathbb{R}^{d \times d}$ -valued Radon measures on  $\Omega \times [0, T]$ .

As for the first summand of the energy-stress tensor  $T_\varepsilon$ , it follows immediately from the equipartition of energy statement in Theorem 4.3(i) that

$$\frac{1}{2} \left( \varepsilon f(-\nabla u_\varepsilon) + \frac{1}{\varepsilon} W(u_\varepsilon) \right) I_d \xrightarrow{*} c_0 \sigma(v) I_d |\nabla \chi|.$$

The following lemma covers the convergence of the second summand of  $T_\varepsilon$ , thereby completing the proof of Theorem 4.9.

**Lemma 4.10.** *Under the assumptions of Theorem 4.1, including the energy convergence (4.3), we have*

$$\varepsilon \nabla u_\varepsilon \otimes \frac{1}{2} Df(-\nabla u_\varepsilon) \xrightarrow{*} -c_0 v \otimes D\sigma(v) |\nabla \chi| \quad \text{in } \mathcal{M}(\Omega \times [0, T])^{d \times d} \tag{4.35}$$

as  $\varepsilon \searrow 0$ .

To prove Lemma 4.10, let us first fix a vector field  $\xi \in C(\Omega \times [0, T])^d$  such that  $|\xi| \leq 1$  in  $\Omega \times [0, T]$ . This vector field will serve as an approximation for the measure-theoretic normal  $v$ . Furthermore let  $A \in C(\Omega \times [0, T])^{d \times d}$  be a test function.

For shorter notation, we introduce the approximate outer unit normal  $v_\varepsilon$  via

$$v_\varepsilon(x) := \begin{cases} -\frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|}(x) & \text{if } \nabla u_\varepsilon(x) \neq 0, \\ e_1 & \text{if } \nabla u_\varepsilon(x) = 0. \end{cases} \tag{4.36}$$

Applying the product rule to  $f = \sigma^2$  and exploiting the positive 0-homogeneity of  $D\sigma$ , one can rewrite the left-hand side of (4.35) as

$$\varepsilon \nabla u_\varepsilon \otimes \frac{1}{2} Df(-\nabla u_\varepsilon) = \varepsilon \sigma(-\nabla u_\varepsilon) \nabla u_\varepsilon \otimes D\sigma(-\nabla u_\varepsilon) = \varepsilon \sigma(-\nabla u_\varepsilon) \nabla u_\varepsilon \otimes D\sigma(v_\varepsilon), \tag{4.37}$$

where one can now conveniently insert the vector field  $\xi$  at the cost of two errors controlled by the tilt excess  $\mathcal{E}$ . Let us add zero twice and compute

$$\begin{aligned}
 & \int_0^T \int_{\Omega} \left( \varepsilon \nabla u_{\varepsilon} \otimes \frac{1}{2} Df(-\nabla u_{\varepsilon}) \right) : A \, dxdt + c_0 \int_0^T \int_{\Omega} (v \otimes D\sigma(v)) : A \, |\nabla \chi| \\
 &= \int_0^T \int_{\Omega} (\varepsilon \sigma(-\nabla u_{\varepsilon}) \nabla u_{\varepsilon} \otimes (D\sigma(v_{\varepsilon}) - \psi(|\xi|)D\sigma(\xi))) : A \, dxdt \\
 &+ \int_0^T \int_{\Omega} (\varepsilon \sigma(-\nabla u_{\varepsilon}) \nabla u_{\varepsilon} \otimes \psi(|\xi|)D\sigma(\xi)) : A \, dxdt \\
 &+ c_0 \int_0^T \int_{\Omega} (v \otimes \psi(|\xi|)D\sigma(\xi)) : A \, |\nabla \chi| \\
 &+ c_0 \int_0^T \int_{\Omega} (v \otimes (D\sigma(v) - \psi(|\xi|)D\sigma(\xi))) : A \, |\nabla \chi|. \tag{4.38}
 \end{aligned}$$

As a consequence of the smoothness and homogeneity of  $\sigma$  and since  $\psi \equiv 0$  in a neighborhood of zero, the map  $p \mapsto \psi(|p|)D\sigma(p)$  is Lipschitz continuous, i.e.,

$$|\psi(|q|)D\sigma(q) - \psi(|p|)D\sigma(p)| \leq C|q - p| \quad \text{for all } p, q \in \mathbb{R}^d. \tag{4.39}$$

The first right-hand side integral of (4.38) can now be estimated as follows:

$$\begin{aligned}
 & \left| \int_0^T \int_{\Omega} (\varepsilon \sigma(-\nabla u_{\varepsilon}) \nabla u_{\varepsilon} \otimes (D\sigma(v_{\varepsilon}) - \psi(|\xi|)D\sigma(\xi))) : A \, dxdt \right| \\
 & \leq C \|A\|_{L^{\infty}} \int_0^T \int_{\Omega} \varepsilon \sigma(-\nabla u_{\varepsilon}) |\nabla u_{\varepsilon}| |v_{\varepsilon} - \xi| \, dxdt \\
 & \leq \frac{\sqrt{2} C \|A\|_{L^{\infty}}}{\min_{S^{d-1}} \sigma} \left( \int_0^T \int_{\Omega} \frac{\varepsilon}{2} f(-\nabla u_{\varepsilon}) \, dxdt \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} \varepsilon f(-\nabla u_{\varepsilon}) |v_{\varepsilon} - \xi|^2 \, dxdt \right)^{\frac{1}{2}} \\
 & \leq \frac{\sqrt{2} C \|A\|_{L^{\infty}}}{\min_{S^{d-1}} \sigma} \int_0^T E_{\varepsilon}[u_{\varepsilon}(t)] \, dt \left( \int_0^T \int_{\Omega} \varepsilon f(-\nabla u_{\varepsilon}) |v_{\varepsilon} - \xi|^2 \, dxdt \right)^{\frac{1}{2}},
 \end{aligned}$$

and with the help of the energy convergence assumption (4.3) and Lemma 4.11 below, we obtain

$$\begin{aligned}
 & \limsup_{\varepsilon \searrow 0} \left| \int_0^T \int_{\Omega} (\varepsilon \sigma(-\nabla u_{\varepsilon}) \nabla u_{\varepsilon} \otimes (D\sigma(v_{\varepsilon}) - \psi(|\xi|)D\sigma(\xi))) : A \, dxdt \right| \\
 & \leq C \|A\|_{L^{\infty}} \left( \int_0^T E[u(t)] \, dt \right)^{\frac{1}{2}} \left( \int_0^T \mathcal{E}[u(t)|\xi(t)] \, dt \right)^{\frac{1}{2}}. \tag{4.40}
 \end{aligned}$$

Similarly recalling (4.30), the last integral on the right-hand side of (4.38) can be controlled by the tilt excess via

$$\begin{aligned}
 & \left| c_0 \int_0^T \int_{\Omega} (v \otimes (D\sigma(v) - \psi(|\xi|)D\sigma(\xi))) : A \, |\nabla \chi| \, dt \right| \\
 & \leq c_0 C \|A\|_{L^{\infty}} \int_0^T \int_{\Omega} |v - \xi| |\nabla \chi| \\
 & \leq c_0 \frac{C \|A\|_{L^{\infty}}}{\sqrt{\min_{S^{d-1}} \sigma}} \left( \int_0^T \int_{\Omega} \sigma(v) |\nabla \chi| \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} |v - \xi|^2 |\nabla \chi| \right)^{\frac{1}{2}} \\
 & \leq \frac{C \|A\|_{L^{\infty}}}{\sqrt{c_{\sigma} \min_{S^{d-1}} \sigma}} \left( \int_0^T E[u(t)] \, dt \right)^{\frac{1}{2}} \left( \int_0^T \mathcal{E}[u(t)|\xi(t)] \, dt \right)^{\frac{1}{2}}. \tag{4.41}
 \end{aligned}$$



As for the remaining terms in (4.38), we have to show that

$$\int_0^T \int_{\Omega} (\varepsilon \sigma(-\nabla u_{\varepsilon}) \nabla u_{\varepsilon} \otimes \psi(|\xi|) D\sigma(\xi)) : A \, dxdt + c_0 \int_0^T \int_{\Omega} (v \otimes \psi(|\xi|) D\sigma(\xi)) : A |\nabla \chi| \rightarrow 0$$

as  $\varepsilon \searrow 0$ . Since  $\psi(|\xi|)A D\sigma(\xi)$  is a continuous vector field on  $\Omega \times [0, T]$ , this problem reduces to proving that  $\varepsilon \sigma(-\nabla u_{\varepsilon}) \nabla u_{\varepsilon} \xrightarrow{*} -c_0 v |\nabla \chi|$  in  $\mathcal{M}(\Omega \times [0, T])^d$ .

To prove this weak-\* convergence statement, we rewrite

$$\begin{aligned} \varepsilon \sigma(-\nabla u_{\varepsilon}) \nabla u_{\varepsilon} &= \left( \varepsilon \sigma(-\nabla u_{\varepsilon}) - \sqrt{W(u_{\varepsilon})} \right) \nabla u_{\varepsilon} + \sqrt{W(u_{\varepsilon})} \nabla u_{\varepsilon} \\ &= \left( \sqrt{\varepsilon f(-\nabla u_{\varepsilon})} - \sqrt{\frac{1}{\varepsilon} W(u_{\varepsilon})} \right) \sqrt{\varepsilon} \nabla u_{\varepsilon} + \nabla(\phi \circ u_{\varepsilon}). \end{aligned}$$

The first summand converges strongly to 0 in  $L^1(\Omega \times (0, T))$ : This follows by the Cauchy–Schwarz inequality if we recall that  $\sqrt{\varepsilon} \nabla u_{\varepsilon}$  is bounded in  $L^2(\Omega \times (0, T))$  as  $\varepsilon \searrow 0$  and  $\sqrt{\varepsilon f(-\nabla u_{\varepsilon})} - \sqrt{\frac{1}{\varepsilon} W(u_{\varepsilon})} \rightarrow 0$  in  $L^2(\Omega \times (0, T))$  by Theorem 4.3(iii).

It has been shown in (4.6)–(4.8) that  $\phi \circ u_{\varepsilon}$  is bounded in  $W^{1,1}$  as  $\varepsilon \searrow 0$ . In particular, the total variation  $\int \int |\nabla u_{\varepsilon}|$  is uniformly bounded as  $\varepsilon \searrow 0$ . Since  $\phi \circ u_{\varepsilon} \rightarrow c_0 \chi$  in  $L^1(\Omega \times (0, T))$ , it follows that  $\nabla(\phi \circ u_{\varepsilon}) \xrightarrow{*} c_0 \nabla \chi = -c_0 v |\nabla \chi|$  in  $\mathcal{M}(\Omega \times [0, T])^d$ .

In total, we have

$$\varepsilon \sigma(-\nabla u_{\varepsilon}) \nabla u_{\varepsilon} \xrightarrow{*} -c_0 v |\nabla \chi| \quad \text{in } \mathcal{M}(\Omega \times [0, T])^d. \tag{4.42}$$

By plugging in (4.40), (4.41), and (4.42) into (4.38), one arrives at

$$\begin{aligned} \limsup_{\varepsilon \searrow 0} \left| \int_0^T \int_{\Omega} \left( \varepsilon \nabla u_{\varepsilon} \otimes \frac{1}{2} Df(-\nabla u_{\varepsilon}) \right) : A \, dxdt + c_0 \int_0^T \int_{\Omega} (v \otimes D\sigma(v)) : A |\nabla \chi| \right| \\ \leq C \|A\|_{L^\infty} \left( \int_0^T E[u(t)] dt \right)^{\frac{1}{2}} \left( \int_0^T \mathcal{E}[u(t)|\xi(t)] dt \right)^{\frac{1}{2}}. \end{aligned}$$

This concludes the proof of Lemma 4.10 since, appealing to Lemma 4.8, the time-integrated relative entropy can be made arbitrarily small.

**Lemma 4.11.** *Under the assumptions of Theorem 4.1, including the energy convergence (4.3), and with  $v_{\varepsilon}$  defined in (4.36), we have*

$$\limsup_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \varepsilon f(-\nabla u_{\varepsilon}) |v_{\varepsilon} - \xi|^2 dxdt \leq \frac{\max_{S^{d-1}} \sigma}{c_{\sigma}} \int_0^T \mathcal{E}[u(t)|\xi(t)] dt. \tag{4.43}$$

**Proof.** We apply Lemma 2.4(i), the estimate  $|v_{\varepsilon} - \xi| \leq 2$ , and the Cauchy–Schwarz inequality in order to control the occurring integrals by the anisotropic Cahn–Hilliard energy and the  $\varepsilon$ -relative entropy. Precisely,

$$\begin{aligned}
 & \int_0^T \int_{\Omega} \varepsilon f(-\nabla u_{\varepsilon}) |v_{\varepsilon} - \xi|^2 dx dt \\
 &= \int_0^T \int_{\Omega} \sqrt{W(u_{\varepsilon})} \sigma(-\nabla u_{\varepsilon}) |v_{\varepsilon} - \xi|^2 dx dt \\
 & \quad + \int_0^T \int_{\Omega} \left( \sqrt{\varepsilon f(-\nabla u_{\varepsilon})} - \sqrt{\frac{1}{\varepsilon} W(u_{\varepsilon})} \right) \sqrt{\varepsilon} \sigma(-\nabla u_{\varepsilon}) |v_{\varepsilon} - \xi|^2 dx dt \\
 &\leq \frac{\max_{S^{d-1}} \sigma}{c_{\sigma}} \int_0^T \int_{\Omega} \sqrt{W(u_{\varepsilon})} |\nabla u_{\varepsilon}| (\sigma(v_{\varepsilon}) - |\xi| \psi(|\xi|) D\sigma(\xi) \cdot v_{\varepsilon}) dx dt \\
 & \quad + 4 \left( \int_0^T \int_{\Omega} \varepsilon f(-\nabla u_{\varepsilon}) dx dt \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} \left( \sqrt{\varepsilon f(-\nabla u_{\varepsilon})} - \sqrt{\frac{1}{\varepsilon} W(u_{\varepsilon})} \right)^2 dx dt \right)^{\frac{1}{2}} \\
 &\leq \frac{\max_{S^{d-1}} \sigma}{c_{\sigma}} \int_0^T \mathcal{E}_{\varepsilon}[u_{\varepsilon}(t) | \xi(t)] dt \\
 & \quad + 4\sqrt{2} \left( \int_0^T E_{\varepsilon}[u_{\varepsilon}(t)] dt \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} \left( \sqrt{\varepsilon f(-\nabla u_{\varepsilon})} - \sqrt{\frac{1}{\varepsilon} W(u_{\varepsilon})} \right)^2 dx dt \right)^{\frac{1}{2}}.
 \end{aligned}$$

One can now take the limit superior on both sides and note that, by assumption, we have  $\limsup_{\varepsilon \searrow 0} \int_0^T E_{\varepsilon}[u_{\varepsilon}(t)] dt < \infty$ . Thus, applying Lemma 4.7 in the first term and Theorem 4.3(iii) in the second term yields

$$\limsup_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \varepsilon f(-\nabla u_{\varepsilon}) |v_{\varepsilon} - \xi|^2 dx dt \leq \frac{\max_{S^{d-1}} \sigma}{c_{\sigma}} \int_0^T \mathcal{E}[u(t) | \xi(t)] dt. \quad \square$$

**4.6 Convergence of the velocity term**

The analogue of Theorem 4.9 for the left-hand side terms is

**Theorem 4.12.** *Under the assumptions of Theorem 4.1, including the energy convergence (4.3), we have*

$$\lim_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} g(-\nabla u_{\varepsilon}) \partial_t u_{\varepsilon} B \cdot \varepsilon \nabla u_{\varepsilon} dx dt = -c_0 \int_0^T \int_{\Omega} \frac{1}{\mu(v)} VB \cdot v |\nabla \chi| \tag{4.44}$$

for all  $B \in C^1(\Omega \times [0, T])^d$ .

Similarly to the argument for the curvature term, we replace the approximate outer normal  $v_{\varepsilon}$  by a  $v_{\varepsilon}$  by a vector field  $\xi \in C(\Omega \times [0, T])^d$  such that  $|\xi| \leq 1$  on terms can be controlled by the relative entropy. An additional error term arises as we replace the term  $g(-\nabla u_{\varepsilon})$  by its ‘asymptotic’ version  $\frac{\sigma(v_{\varepsilon})}{\mu(v_{\varepsilon})}$ : Expanding both sides of (4.44), we obtain

$$\begin{aligned}
 & \int_0^T \int_{\Omega} g(-\nabla u_{\varepsilon}) \partial_t u_{\varepsilon} B \cdot \varepsilon \nabla u_{\varepsilon} dx dt + c_0 \int_0^T \int_{\Omega} \frac{1}{\mu(v)} VB \cdot v |\nabla \chi| \\
 &= \int_0^T \int_{\Omega} \left( g(-\nabla u_{\varepsilon}) - \frac{\sigma(v_{\varepsilon})}{\mu(v_{\varepsilon})} \right) \partial_t u_{\varepsilon} B \cdot \varepsilon \nabla u_{\varepsilon} dx dt
 \end{aligned}$$

$$\begin{aligned}
 & - \int_0^T \int_{\Omega} \partial_t u_{\varepsilon} B \cdot \varepsilon \sigma(-\nabla u_{\varepsilon}) \left( \frac{1}{\mu(v_{\varepsilon})} v_{\varepsilon} - \frac{\psi(|\xi|)}{\mu(\xi)} \xi \right) dxdt \\
 & - \int_0^T \int_{\Omega} \partial_t u_{\varepsilon} B \cdot \varepsilon \sigma(-\nabla u_{\varepsilon}) \frac{\psi(|\xi|)}{\mu(\xi)} \xi dxdt + c_0 \int_0^T \int_{\Omega} VB \cdot \frac{\psi(|\xi|)}{\mu(\xi)} \xi |\nabla \chi| \\
 & + c_0 \int_0^T \int_{\Omega} VB \cdot \left( \frac{1}{\mu(v)} v - \frac{\psi(|\xi|)}{\mu(\xi)} \xi \right) |\nabla \chi|,
 \end{aligned} \tag{4.45}$$

where we have used the identity  $-\sigma(v_{\varepsilon})\nabla u_{\varepsilon} = \sigma(v_{\varepsilon})|\nabla u_{\varepsilon}|v_{\varepsilon} = \sigma(-\nabla u_{\varepsilon})v_{\varepsilon}$ .

To see that the third line on the right-hand side converges to zero as  $\varepsilon \searrow 0$ , it suffices to show that

$$\int_0^T \int_{\Omega} \varepsilon \partial_t u_{\varepsilon} \sigma(-\nabla u_{\varepsilon}) \zeta dxdt \longrightarrow c_0 \int_0^T \int_{\Omega} V \zeta |\nabla \chi| \tag{4.46}$$

as  $\varepsilon \searrow 0$  for all  $\zeta \in C^1(\Omega \times [0, T])$ . A density argument then yields the same statement for all  $\zeta \in C(\Omega \times [0, T])$ , so that one can choose  $\zeta = B \cdot \frac{\psi(|\xi|)}{\mu(\xi)} \xi$ .

In order to prove (4.46), we integrate by parts and write

$$\begin{aligned}
 \int_0^T \int_{\Omega} \varepsilon \partial_t u_{\varepsilon} \sigma(-\nabla u_{\varepsilon}) \zeta dxdt &= \int_0^T \int_{\Omega} \sqrt{\varepsilon} \partial_t u_{\varepsilon} \left( \sqrt{\varepsilon} \sigma(-\nabla u_{\varepsilon}) - \sqrt{\frac{1}{\varepsilon} W(u_{\varepsilon})} \right) \zeta dxdt \\
 &\quad - \int_0^T \int_{\Omega} (\phi \circ u_{\varepsilon}) \partial_t \zeta dxdt \\
 &\quad + \int_{\Omega} \phi(u_{\varepsilon}(x, T)) \zeta(x, T) dx - \int_{\Omega} \phi(u_{\varepsilon,0}(x)) \zeta(x, 0) dx.
 \end{aligned}$$

The first integral converges to zero as  $\varepsilon \searrow 0$  due to the equipartition of energy, Theorem 4.3(iii), and the optimal dissipation. For the other three integrals, we apply the convergence statements for the subsequence  $\varepsilon \searrow 0$  given in Lemma 4.2, which leads to

$$\begin{aligned}
 \lim_{\varepsilon \searrow 0} \int_0^T \int_{\Omega} \varepsilon \partial_t u_{\varepsilon} \sigma(-\nabla u_{\varepsilon}) \zeta dxdt \\
 = -c_0 \int_0^T \int_{\Omega} \chi \partial_t \zeta dxdt + c_0 \int_{\Omega} \chi(x, T) \zeta(x, T) dx - c_0 \int_{\Omega} \chi_0(x) \zeta(x, 0) dx.
 \end{aligned}$$

The distributional criterion (2.18) for the normal velocity yields (4.46), and so it follows that the third line on the right-hand side of (4.45) converges to zero.

It remains to estimate the “error” terms in (4.45). For the first integral on the right-hand side, one uses the asymptotic description of  $g$  provided by Lemma 3.2(iii): For every  $\delta > 0$ , there exists a positive constant  $R$  such that  $\left| g(p) - \frac{\sigma(p)}{\mu(p)} \right| < \delta$  whenever  $|p| \geq R$ . Therefore, we obtain

$$\begin{aligned}
 & \left| \int_0^T \int_{\Omega} \left( g(-\nabla u_{\varepsilon}) - \frac{\sigma(v_{\varepsilon})}{\mu(v_{\varepsilon})} \right) \partial_t u_{\varepsilon} B \cdot \varepsilon \nabla u_{\varepsilon} dxdt \right| \\
 & \leq \int_0^T \int_{\Omega} \chi_{\{|\nabla u_{\varepsilon}| \leq R\}} \left| g(-\nabla u_{\varepsilon}) - \frac{\sigma(v_{\varepsilon})}{\mu(v_{\varepsilon})} \right| |\partial_t u_{\varepsilon}| |B \cdot \varepsilon \nabla u_{\varepsilon}| dxdt \\
 & \quad + \int_0^T \int_{\Omega} \chi_{\{|\nabla u_{\varepsilon}| > R\}} \left| g(-\nabla u_{\varepsilon}) - \frac{\sigma(v_{\varepsilon})}{\mu(v_{\varepsilon})} \right| |\partial_t u_{\varepsilon}| |B \cdot \varepsilon \nabla u_{\varepsilon}| dxdt
 \end{aligned}$$

$$\begin{aligned}
 &\leq 2R \left(\sup_{\mathbb{R}^d} g\right) \|B\|_{L^\infty} \int_0^T \int_\Omega \varepsilon |\partial_t u_\varepsilon| \, dxdt \\
 &\quad + \delta \|B\|_{L^\infty} \int_0^T \int_\Omega |\partial_t u_\varepsilon| \varepsilon |\nabla u_\varepsilon| \, dxdt \\
 &\leq 2R \left(\sup_{\mathbb{R}^d} g\right) \|B\|_{L^\infty} \sqrt{\varepsilon T} \left(\int_0^T \int_\Omega \varepsilon (\partial_t u_\varepsilon)^2 \, dxdt\right)^{\frac{1}{2}} \\
 &\quad + \delta \|B\|_{L^\infty} \left(\int_0^T \int_\Omega \varepsilon (\partial_t u_\varepsilon)^2 \, dxdt\right)^{\frac{1}{2}} \left(\int_0^T \int_\Omega \varepsilon |\nabla u_\varepsilon|^2 \, dxdt\right)^{\frac{1}{2}} \\
 &\leq \sqrt{\varepsilon} \frac{2R\sqrt{T} \sup_{\mathbb{R}^d} g}{\sqrt{c_g}} \|B\|_{L^\infty} E_\varepsilon[u_{\varepsilon,0}]^{\frac{1}{2}} \\
 &\quad + \delta \frac{\sqrt{2}}{\sqrt{c_g} \min_{S^{d-1}} \sigma} \|B\|_{L^\infty} E_\varepsilon[u_{\varepsilon,0}]^{\frac{1}{2}} \left(\int_0^T E_\varepsilon[u_\varepsilon(t)] \, dt\right)^{\frac{1}{2}}, \tag{4.47}
 \end{aligned}$$

from which it follows that

$$\begin{aligned}
 &\limsup_{\varepsilon \searrow 0} \left| \int_0^T \int_\Omega \left( g(-\nabla u_\varepsilon) - \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} \right) \partial_t u_\varepsilon B \cdot \varepsilon \nabla u_\varepsilon \, dxdt \right| \\
 &\leq \delta C \|B\|_{L^\infty} E_0^{\frac{1}{2}} \left(\int_0^T E[u(t)] \, dt\right)^{\frac{1}{2}}. \tag{4.48}
 \end{aligned}$$

Since  $\delta > 0$  is arbitrary, we have shown that the first integral on the right-hand side of (4.45) converges to zero as  $\varepsilon \searrow 0$ .

As for the second integral on the right-hand side, we use that the map  $p \mapsto \frac{\psi(|p|)}{\mu(p)} p$  is Lipschitz continuous on  $\mathbb{R}^d$ , so that

$$\begin{aligned}
 &\left| \int_0^T \int_\Omega \partial_t u_\varepsilon B \cdot \varepsilon \sigma(-\nabla u_\varepsilon) \left( \frac{1}{\mu(v_\varepsilon)} v_\varepsilon - \frac{\psi(|\xi|)}{\mu(\xi)} \xi \right) \, dxdt \right| \\
 &\leq C \|B\|_{L^\infty} \int_0^T \int_\Omega |\partial_t u_\varepsilon| \varepsilon \sigma(-\nabla u_\varepsilon) |v_\varepsilon - \xi| \, dxdt \\
 &\leq C \|B\|_{L^\infty} \left(\int_0^T \int_\Omega \varepsilon (\partial_t u_\varepsilon)^2 \, dxdt\right)^{\frac{1}{2}} \left(\int_0^T \int_\Omega \varepsilon f(-\nabla u_\varepsilon) |v_\varepsilon - \xi|^2 \, dxdt\right)^{\frac{1}{2}} \\
 &\leq \frac{C}{\sqrt{c_g}} \|B\|_{L^\infty} E_\varepsilon[u_{\varepsilon,0}]^{\frac{1}{2}} \left(\int_0^T \int_\Omega \varepsilon f(-\nabla u_\varepsilon) |v_\varepsilon - \xi|^2 \, dxdt\right)^{\frac{1}{2}}.
 \end{aligned}$$

It follows with the help of Lemma 4.11 that

$$\begin{aligned}
 &\limsup_{\varepsilon \searrow 0} \left| \int_0^T \int_\Omega \partial_t u_\varepsilon B \cdot \varepsilon \sigma(-\nabla u_\varepsilon) \left( \frac{1}{\mu(v_\varepsilon)} v_\varepsilon - \frac{\psi(|\xi|)}{\mu(\xi)} \xi \right) \, dxdt \right| \\
 &\leq C \|B\|_{L^\infty} E_0^{\frac{1}{2}} \left(\int_0^T \mathcal{E}[u(t)|\xi(t)] \, dt\right)^{\frac{1}{2}}. \tag{4.49}
 \end{aligned}$$

Again making use of the Lipschitz continuity and Lemma 2.4(i), the last integral on the right-hand side of (4.45) can be estimated as

$$\begin{aligned} & \left| c_0 \int_0^T \int_{\Omega} VB \cdot \left( \frac{1}{\mu(v)} v - \frac{\psi(|\xi|)}{\mu(\xi)} \xi \right) |\nabla \chi| \right| \\ & \leq c_0 C \|B\|_{L^\infty} \int_0^T \int_{\Omega} |V| |v - \xi| |\nabla \chi| \\ & \leq c_0 C \|B\|_{L^\infty} \left( \int_0^T \int_{\Omega} V^2 |\nabla \chi| \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} |v - \xi|^2 |\nabla \chi| \right)^{\frac{1}{2}} \\ & \leq \frac{C \sqrt{c_0}}{\sqrt{c_\sigma}} \|B\|_{L^\infty} \|V\|_{L^2(\nabla \chi)} \left( \int_0^T \mathcal{E}[u(t)|\xi(t)] dt \right)^{\frac{1}{2}}. \end{aligned} \tag{4.50}$$

Plugging in the results for each right-hand side integral into (4.45) leads to

$$\begin{aligned} \limsup_{\varepsilon \searrow 0} & \left| \int_0^T \int_{\Omega} g(-\nabla u_\varepsilon) \partial_t u_\varepsilon B \cdot \varepsilon \nabla u_\varepsilon \, dx dt + c_0 \int_0^T \int_{\Omega} \frac{1}{\mu(v)} VB \cdot v |\nabla \chi| \right| \\ & \leq C \|B\|_{L^\infty} \left( E_0^{\frac{1}{2}} + \|V\|_{L^2(\nabla \chi)} \right) \left( \int_0^T \mathcal{E}[u(t)|\xi(t)] dt \right)^{\frac{1}{2}}. \end{aligned}$$

The proof of Theorem 4.12 is complete once we take into account that, by Lemma 4.8, the time-integrated relative entropy can be made arbitrarily small for suitable choices of vector fields  $\xi$ .

### 4.7 Optimal energy dissipation inequality

**Lemma 4.13.** *Under the assumptions of Theorem 4.1, including the energy convergence (4.3), the inequality*

$$E[u(T')] + c_0 \int_0^{T'} \int_{\Omega} \frac{1}{\mu(v)} V^2 |\nabla \chi| \leq E_0$$

holds true for all  $T' \in [0, T]$ .

The strategy to prove this lemma is to take  $\varepsilon \searrow 0$  in the optimal energy dissipation identity (3.4) for the anisotropic Allen–Cahn equation. By the well-preparedness of the initial data (4.1), we have  $\lim_{\varepsilon \searrow 0} E_\varepsilon[u_{\varepsilon,0}] = E_0$ . Furthermore, the  $\Gamma$ -convergence  $E_\varepsilon \xrightarrow{\Gamma} E$  on  $L^1(\Omega)$  (see [9, Theorem 3.5]) yields  $\liminf_{\varepsilon \searrow 0} E_\varepsilon[u_\varepsilon(T')] \geq E[u(T')]$  since it follows from the conditions in Lemma 4.2 that  $u_\varepsilon(t) \rightarrow u(t)$  in all  $t \in [0, T]$ .

We remark that Bouchitté [9] defines  $\text{dom}(E_\varepsilon) = \text{Lip}(\Omega)$ , so that the liminf inequality for our definition in (1.7) does not immediately follow. However, the liminf inequality is the easier part of the  $\Gamma$ -convergence statement and can be shown by combining the Modica–Mortola argument in Theorem 4.3(i) above with the truncation of  $W$  that can be found in [44, Proof of Theorem 1.6].

Therefore, the proof of the lemma reduces to proving the following inequality:

**Claim 2.** *For every  $T' \in [0, T]$ , we have*

$$\liminf_{\varepsilon \searrow 0} \int_0^{T'} \int_{\Omega} \varepsilon g(-\nabla u_\varepsilon) (\partial_t u_\varepsilon)^2 \, dx dt \geq c_0 \int_0^{T'} \int_{\Omega} \frac{1}{\mu(v)} V^2 |\nabla \chi|. \tag{4.51}$$

Let  $\zeta \in C(\Omega \times [0, T])$  and  $\xi \in C(\Omega \times [0, T])^d$  such that  $|\xi| \leq 1$  in  $\Omega \times [0, T]$ . We apply Young’s inequality in the form  $a^2 \geq 2ab - b^2$  and add zero multiple times in order to replace  $g$  with its asymptotic

version and  $v_\varepsilon$  with  $\xi$ , similarly to the approach in (4.45). In this case, we obtain

$$\begin{aligned}
 & \int_0^{T'} \int_\Omega \varepsilon g(-\nabla u_\varepsilon) (\partial_t u_\varepsilon)^2 dx dt \\
 & \geq 2 \int_0^{T'} \int_\Omega \varepsilon g(-\nabla u_\varepsilon) \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta dx dt - \int_0^{T'} \int_\Omega \varepsilon g(-\nabla u_\varepsilon) f(-\nabla u_\varepsilon) \zeta^2 dx dt \\
 & = 2 \int_0^{T'} \int_\Omega \varepsilon \left( g(-\nabla u_\varepsilon) - \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} \right) \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta dx dt \\
 & \quad - \int_0^{T'} \int_\Omega \varepsilon \left( g(-\nabla u_\varepsilon) - \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} \right) f(-\nabla u_\varepsilon) \zeta^2 dx dt \\
 & \quad + 2 \int_0^{T'} \int_\Omega \varepsilon \left( \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} - \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \right) \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta dx dt \\
 & \quad - \int_0^{T'} \int_\Omega \varepsilon \left( \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} - \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \right) f(-\nabla u_\varepsilon) \zeta^2 dx dt \\
 & \quad + 2 \int_0^{T'} \int_\Omega \varepsilon \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta dx dt \\
 & \quad - \int_0^{T'} \int_\Omega \varepsilon \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} f(-\nabla u_\varepsilon) \zeta^2 dx dt. \tag{4.52}
 \end{aligned}$$

To show that the first integral on the right-hand side converges to zero, let  $\delta > 0$ . By Lemma 3.2(iii), there exists some  $R > 0$  such that  $\left| g(p) - \frac{\sigma(p)}{\mu(p)} \right| < \delta$  for all  $p \in \mathbb{R}^d$  with  $|p| \geq R$ . Arguing as in (4.47), we find

$$\begin{aligned}
 & \left| 2 \int_0^{T'} \int_\Omega \varepsilon \left( g(-\nabla u_\varepsilon) - \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} \right) \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta dx dt \right| \\
 & \leq 4R \frac{(\max_{S^{d-1}} \sigma)(\sup_{\mathbb{R}^d} g)}{\sqrt{c_g}} \|\zeta\|_{L^\infty} \sqrt{\varepsilon T} E_\varepsilon[u_{\varepsilon,0}]^{\frac{1}{2}} \\
 & \quad + 2\sqrt{2} \delta \frac{\max_{S^{d-1}} \sigma}{\sqrt{c_g} \min_{S^{d-1}} \sigma} \|\zeta\|_{L^\infty} E_\varepsilon[u_{\varepsilon,0}]^{\frac{1}{2}} \left( \int_0^{T'} E_\varepsilon[u_\varepsilon(t)] dt \right)^{\frac{1}{2}}.
 \end{aligned}$$

Taking first  $\varepsilon \searrow 0$ , then  $\delta \searrow 0$  leads to

$$\lim_{\varepsilon \searrow 0} \left| 2 \int_0^{T'} \int_\Omega \varepsilon \left( g(-\nabla u_\varepsilon) - \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} \right) \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta dx dt \right| = 0. \tag{4.53}$$

For the second line on the right-hand side of (4.52), we argue analogously and obtain

$$\lim_{\varepsilon \searrow 0} \left| - \int_0^{T'} \int_\Omega \varepsilon \left( g(-\nabla u_\varepsilon) - \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} \right) f(-\nabla u_\varepsilon) \zeta^2 dx dt \right| = 0. \tag{4.54}$$

Due to the Lipschitz continuity of the map  $p \mapsto \psi(|p|) \frac{\sigma(p)}{\mu(p)}$ , the third integral can be estimated as

$$\begin{aligned} & \left| 2 \int_0^{T'} \int_{\Omega} \varepsilon \left( \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} - \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \right) \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta \, dx dt \right| \\ & \leq 2C \|\zeta\|_{L^\infty} \int_0^{T'} \int_{\Omega} \varepsilon |v_\varepsilon - \xi| |\partial_t u_\varepsilon| \sigma(-\nabla u_\varepsilon) \, dx dt \\ & \leq 2C \|\zeta\|_{L^\infty} \left( \int_0^{T'} \int_{\Omega} \varepsilon (\partial_t u_\varepsilon)^2 \, dx dt \right)^{\frac{1}{2}} \left( \int_0^{T'} \int_{\Omega} \varepsilon |v_\varepsilon - \xi|^2 f(-\nabla u_\varepsilon) \, dx dt \right)^{\frac{1}{2}} \\ & \leq 2 \frac{C}{\sqrt{c_g}} \|\zeta\|_{L^\infty} E_\varepsilon[u_{\varepsilon,0}]^{\frac{1}{2}} \left( \int_0^{T'} \int_{\Omega} \varepsilon |v_\varepsilon - \xi|^2 f(-\nabla u_\varepsilon) \, dx dt \right)^{\frac{1}{2}}, \end{aligned}$$

so that, by appealing to Lemma 4.11, we obtain

$$\begin{aligned} & \limsup_{\varepsilon \searrow 0} \left| 2 \int_0^{T'} \int_{\Omega} \varepsilon \left( \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} - \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \right) \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta \, dx dt \right| \\ & \leq C \|\zeta\|_{L^\infty} E_0^{\frac{1}{2}} \left( \int_0^T \mathcal{E}[u(t)|\xi(t)] \, dt \right)^{\frac{1}{2}}. \end{aligned} \tag{4.55}$$

For the fourth line of the right-hand side of (4.52), a very similar computation yields

$$\begin{aligned} & \limsup_{\varepsilon \searrow 0} \left| - \int_0^{T'} \int_{\Omega} \varepsilon \left( \frac{\sigma(v_\varepsilon)}{\mu(v_\varepsilon)} - \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \right) f(-\nabla u_\varepsilon) \zeta^2 \, dx dt \right| \\ & \leq C \|\zeta\|_{L^\infty}^2 \left( \int_0^T E[u(t)] \, dt \right)^{\frac{1}{2}} \left( \int_0^T \mathcal{E}[u(t)|\xi(t)] \, dt \right)^{\frac{1}{2}}. \end{aligned} \tag{4.56}$$

Finally, it follows from (4.46) and the equipartition of energy, Theorem 4.3(v), that

$$\begin{aligned} & \lim_{\varepsilon \searrow 0} \left( 2 \int_0^{T'} \int_{\Omega} \varepsilon \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \partial_t u_\varepsilon \sigma(-\nabla u_\varepsilon) \zeta \, dx dt - \int_0^{T'} \int_{\Omega} \varepsilon \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} f(-\nabla u_\varepsilon) \zeta^2 \, dx dt \right) \\ & = 2c_0 \int_0^{T'} \int_{\Omega} \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} V \zeta |\nabla \chi| - c_0 \int_0^{T'} \int_{\Omega} \sigma(v) \psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \zeta^2 |\nabla \chi|. \end{aligned} \tag{4.57}$$

It is now desirable to let  $\xi \rightarrow v$  in  $L^2(|\nabla \chi|)$  so that  $\int_0^T \mathcal{E}[u(t)|\xi(t)] \, dt \rightarrow 0$ . This is possible by Lemma 4.8 and Lemma 2.4(i). By the dominated convergence theorem, we also have  $\psi(|\xi|) \frac{\sigma(\xi)}{\mu(\xi)} \rightarrow \frac{\sigma(v)}{\mu(v)}$  in  $L^2(|\nabla \chi|)$ . Under this convergence, it follows from (4.52) and the computations (4.53)–(4.57) that

$$\begin{aligned} & \liminf_{\varepsilon \searrow 0} \int_0^{T'} \int_{\Omega} \varepsilon g(-\nabla u_\varepsilon) (\partial_t u_\varepsilon)^2 \, dx dt \\ & \geq 2c_0 \int_0^{T'} \int_{\Omega} \frac{\sigma(v)}{\mu(v)} V \zeta |\nabla \chi| - c_0 \int_0^{T'} \int_{\Omega} \sigma(v) \frac{\sigma(v)}{\mu(v)} \zeta^2 |\nabla \chi|. \end{aligned}$$

In a last step we let  $\zeta \rightarrow \frac{V}{\sigma(v)}$  in  $L^2(|\nabla \chi|)$ , which yields the lower bound as stated in the claim.

### 4.8 Proof of Theorem 4.1

We simply wrap up the proof of the theorem.

**Proof.** The compactness follows from Lemma 4.2. Under the energy convergence assumption, the distributional formulation for the time derivative (2.18) follows from Lemma 4.4. The optimal energy dissipation relation (2.20) follows from Lemma 4.13. To obtain the distributional formulation for the curvature in (2.19), we use  $B \cdot \varepsilon \nabla u_\varepsilon$  as a test function in the distributional formulation of the anisotropic Allen–Cahn equation (3.3), where  $B \in C^1(\Omega \times [0, T])^d$ , where we recall that by Theorem 3.10, we have  $B \cdot \varepsilon \nabla u_\varepsilon \in L^2(0, T; H^1(\Omega))$ , and that these functions are admissible test functions in (3.3). To pass to the limit as  $\varepsilon \searrow 0$  in the left-hand side, one applies Theorem 4.12, and likewise for the right-hand side, apply Theorem 4.9.  $\square$

## 5. Weak–strong uniqueness for anisotropic mean curvature flow

The goal of this section is to prove that, as long as a strong solution to anisotropic mean curvature flow (1.1) exists, any BV solution with the same initial data coincides with the strong solution. One needs to require additional regularity for  $(\sigma, \mu)$  to make sure that strong solutions will be sufficiently smooth, and we will also rely on the higher regularity of  $\sigma$  in the proof of the weak–strong uniqueness statement. Thus, we assume that  $\sigma, \mu \in C^\infty(\mathbb{R}^d \setminus \{0\})$ . Further, without loss of generality, we let  $c_0 = 1$ . Following Hensel and Moser [37, Definition 10], we define strong solutions to (1.1) as follows:

**Definition 5.1.** Let  $T > 0$  be a finite time horizon. A family  $\{\mathcal{A}(t)\}_{t \in [0, T]}$  of open subsets of  $\Omega$  is a strong solution to anisotropic mean curvature flow if

- $\partial \mathcal{A}(0)$  is an embedded  $C^\infty$ -submanifold,
- there exists a  $C^\infty$ -map  $\Phi : \Omega \times [0, T] \rightarrow \Omega$  such that  $\Phi(\cdot, t)$  is a diffeomorphism for all  $t \in [0, T]$ , furthermore  $\Phi(\cdot, 0) = \text{id}_\Omega$  and

$$\Phi(\mathcal{A}(0), t) = \mathcal{A}(t), \quad \Phi(\partial \mathcal{A}(0), t) = \partial \mathcal{A}(t)$$

for all  $t \in [0, T]$ , and

- $\{\partial \mathcal{A}(t)\}_{t \in [0, T]}$  evolves by (1.1) in the classical sense.

The setup in [37, Definition 10 and Remark 15] also suggests that the  $C^\infty$ -regularity for  $(\sigma, \mu)$ ,  $\Phi$ , and  $\partial \mathcal{A}(0)$  can be relaxed.

We are now ready to formulate the central theorem of this section:

**Theorem 5.2.** Let  $\{\mathcal{A}(t)\}_{t \in [0, T]}$  be a solution of anisotropic mean curvature flow (1.1), and further, let  $\{A(t)\}_{t \in [0, T]}$  be time parametrized collection of sets with  $\chi := \chi_A$  a distributional solution of anisotropic mean curvature flow as in Definition 2.8. If

$$|\mathcal{A}(0) \Delta A(0)| = 0,$$

then

$$|\mathcal{A}(t) \Delta A(t)| = 0$$

for all  $t \in [0, T]$ .

The proof for this theorem is modeled after [35, Sections 2.2, 4], where Hensel and the first author derive an analogous result for multiphase isotropic mean curvature flow. The key step in this argument is to find a gradient flow calibration (see below) for the smooth evolution  $\{\mathcal{A}(t)\}_{t \in [0, T]}$ . While the existence of a gradient flow calibration is nontrivial in the multiphase case (see [37, Theorem 4] for a gradient flow calibration for multiphase mean curvature flow in  $d = 2$  with constant contact angle), such a calibration can always be constructed explicitly for a smooth two-phase evolution.



### 5.1 Gradient flow calibrations

In the remainder of this chapter,  $C < \infty$  and  $c > 0$  denote positive constants ('large' and 'small', respectively) that may depend on the pair of anisotropies  $(\sigma, \mu)$ , on the time horizon  $T$ , and on the smooth evolution  $\{\mathcal{A}(t)\}_{t \in [0, T]}$ . These constants need not be the same on every occurrence.

**Definition 5.3.** Let  $\{\mathcal{A}(t)\}_{t \in [0, T]}$  be a strong solution to anisotropic mean curvature flow consisting of nonempty open proper subsets of  $\Omega$ . A gradient flow calibration for  $\{\mathcal{A}(t)\}_{t \in [0, T]}$  is a triple

$$(\xi, B, \vartheta) \in C_1^2(\Omega \times [0, T])^d \times C([0, T]; C^2(\Omega))^d \times C^1(\Omega \times [0, T]; [-1, 1])$$

satisfying

- the approximate evolution equations

$$|\partial_t \xi + (B \cdot \nabla) \xi + (\nabla B)^T \xi| (x, t) \leq C \operatorname{dist}(x, \partial \mathcal{A}(t)) \quad \text{in } \Omega \times [0, T], \quad (5.1)$$

$$|\xi \cdot (\partial_t \xi + (B \cdot \nabla) \xi)| (x, t) \leq C \operatorname{dist}^2(x, \partial \mathcal{A}(t)) \quad \text{in } \Omega \times [0, T], \quad (5.2)$$

$$|\partial_t \vartheta + (B \cdot \nabla) \vartheta| (x, t) \leq C \operatorname{dist}(x, \partial \mathcal{A}(t)) \quad \text{in } \Omega \times [0, T], \quad (5.3)$$

- the compatibility condition

$$|B \cdot \xi + \mu(\xi) \operatorname{div}(|\xi| \psi(|\xi|) D\sigma(\xi))| (x, t) \leq C \operatorname{dist}(x, \partial \mathcal{A}(t)) \quad \text{in } \Omega \times [0, T], \quad (5.4)$$

- and with  $\nu_{\partial \mathcal{A}(t)}$  denoting the outer unit normal of the set  $\mathcal{A}(t)$ , we have the coercivity conditions

$$\xi(x, t) = \nu_{\partial \mathcal{A}(t)}(x) \quad \text{on } \bigcup_{t \in [0, T]} (\partial \mathcal{A}(t) \times \{t\}), \quad (5.5)$$

$$|\xi(x, t)| \leq 1 - c \operatorname{dist}^2(x, \partial \mathcal{A}(t)) \quad \text{in } \Omega \times [0, T], \quad (5.6)$$

$$\vartheta(x, t) > c \operatorname{dist}(x, \partial \mathcal{A}(t)) \quad \text{in } \bigcup_{t \in [0, T]} (\overline{\mathcal{A}(t)^c} \times \{t\}), \quad (5.7)$$

$$\vartheta(x, t) < -c \operatorname{dist}(x, \partial \mathcal{A}(t)) \quad \text{in } \bigcup_{t \in [0, T]} (\mathcal{A}(t) \times \{t\}). \quad (5.8)$$

Intuitively,  $\xi$  is an extension of the outer unit normal  $\nu_{\partial \mathcal{A}(t)}$  (with an additional coercivity property), whereas  $B$  extends the normal velocity vector and  $\vartheta$  is comparable to a signed distance function. The compatibility condition (5.4) encodes the motion by anisotropic mean curvature. The space  $C_1^2(\Omega \times [0, T])$  is defined as

$$C_1^2(\Omega \times [0, T]) := \left\{ f \in C(\Omega \times [0, T]) \mid \partial f, \nabla f, D^2 f \text{ are continuous on } \Omega \times [0, T] \right\}.$$

Let us collect the key inequalities for gradient flow calibrations that will be used in the proof of Theorem 5.2.

**Lemma 5.4.** *Let  $(\xi, B, \vartheta)$  be a gradient flow calibration for  $\{\mathcal{A}(t)\}_{t \in [0, T]}$ . Then the following estimates hold true for all  $(x, t) \in \Omega \times [0, T]$ :*

(i)

$$c \operatorname{dist}^2(x, \partial \mathcal{A}(t)) \leq 1 - |\xi| (x, t) \leq C \operatorname{dist}^2(x, \partial \mathcal{A}(t)),$$

(ii)

$$c \operatorname{dist}^2(x, \partial \mathcal{A}(t)) \leq 1 - |\xi|^2 (x, t) \leq C \operatorname{dist}^2(x, \partial \mathcal{A}(t)),$$

(iii)

$$|\xi \cdot (\xi \cdot \nabla) B| (x, t) \leq C \operatorname{dist}(x, \partial \mathcal{A}(t)).$$

If  $\nu : \Omega \times [0, T] \rightarrow \mathbb{R}^d$  such that  $|\nu| \equiv 1$ , then

(iv)

$$|v - \xi|^2 \leq C (\sigma(v) - |\xi| \psi(|\xi|) D\sigma(\xi) \cdot v),$$

(v)

$$1 - |\xi| \leq C (\sigma(v) - |\xi| \psi(|\xi|) D\sigma(\xi) \cdot v), \text{ and}$$

(vi)

$$|\xi \cdot (v - \xi)| \leq C (\sigma(v) - |\xi| \psi(|\xi|) D\sigma(\xi) \cdot v).$$

**Proof.**

(i) The lower bound is precisely the coercivity condition (5.6).

For the upper bound, we observe that, due to (5.5) and (5.6),  $|\xi|$  attains its maximum on  $\partial\mathcal{A}(t)$  for every  $t \in [0, T]$ . In particular, we have  $\nabla|\xi| = 0$  on  $\partial\mathcal{A}(t)$ . A second-order Taylor expansion yields  $1 - |\xi|(x, t) \leq C \text{dist}^2(x, \partial\mathcal{A}(t))$ . For this argument, we use the smoothness of  $\xi$  (i.e., the continuity of  $\nabla\xi$  and  $\nabla^2\xi$  on  $\Omega \times [0, T]$ ) and the compactness of the domain  $\Omega \times [0, T]$ .

(ii) This follows immediately from (i) since the condition  $|\xi| \leq 1$  allows us to compute

$$1 - |\xi| \leq 1 - |\xi|^2 = (1 - |\xi|)(1 + |\xi|) \leq 2(1 - |\xi|),$$

i.e.,  $1 - |\xi|^2$  is bounded from below and above by a rescaled  $1 - |\xi|$ .

(iii) This inequality is a consequence of (5.1) and (5.2): We have

$$\begin{aligned} |\xi \cdot (\xi \cdot \nabla)B|(x, t) &= \left| \xi \cdot (\partial_t \xi + (B \cdot \nabla)\xi + (\nabla B)^T \xi) - \xi \cdot (\partial_t \xi + (B \cdot \nabla)\xi) \right|(x, t) \\ &\leq C \text{dist}(x, \partial\mathcal{A}(t)) + C \text{dist}^2(x, \partial\mathcal{A}(t)) \\ &\leq C \text{dist}(x, \partial\mathcal{A}(t)). \end{aligned}$$

(iv) This is the statement of Lemma 2.4(i).

(v) This is another elementary estimate: With the help of Lemma 2.3(ii), (vi), one computes

$$\begin{aligned} \sigma(v) - |\xi| \psi(|\xi|) D\sigma(\xi) \cdot v &\geq \sigma(v) - |\xi| \psi(|\xi|) \sigma(v) \\ &\geq (\min_{S^{d-1}} \sigma) (1 - |\xi| \psi(|\xi|)) \\ &\geq (\min_{S^{d-1}} \sigma) (1 - |\xi|). \end{aligned}$$

(vi) follows from the previous two estimates since

$$\xi \cdot (v - \xi) = -\frac{1}{2} (1 - 2 v \cdot \xi + |\xi|^2) + \frac{1}{2} (1 - |\xi|^2) = -\frac{1}{2} |v - \xi|^2 + \frac{1}{2} (1 - |\xi|^2). \quad \square$$

**Lemma 5.5.** *Every strong solution  $\{\mathcal{A}(t)\}_{t \in [0, T]}$  to anisotropic mean curvature flow (1.1) that consists of nonempty open proper subsets of  $\Omega$  admits a gradient flow calibration  $(\xi, B, \vartheta)$ .*

**Proof.** There exists a positive  $\delta > 0$  such that in the neighborhood

$$U := \bigcup_{t \in [0, T]} (B_\delta(\partial\mathcal{A}(t)) \times \{t\}),$$

the signed distance function  $\text{sdist} : U \rightarrow (-\delta, \delta)$  and the orthogonal projection

$$p : U \rightarrow \bigcup_{t \in [0, T]} (\partial\mathcal{A}(t) \times \{t\})$$

with respect to  $\partial\mathcal{A}(t)$  are well-defined and regular (see [34, Lemmas 14.16, 14.17]). We use the sign convention that  $\text{sdist}(x, t) < 0$  for  $x \in \mathcal{A}(t) \cap B_\delta(\partial\mathcal{A}(t))$ .

Let  $\zeta \in C_c^\infty((-\delta, \delta))$  be a cutoff function such that  $\zeta(s) = 1 - s^2$  for  $|s| \leq \frac{\delta}{2}$  and  $s\zeta'(s) \leq 0$  on  $(-\delta, \delta)$ . Furthermore, we define another truncation  $f \in C^\infty(\mathbb{R})$  such that  $f(s) = s$  for  $|s| \leq \frac{\delta}{2}$ ,  $f(s) = \text{sgn}(s)\delta$  for  $|s| \geq \frac{3}{4}\delta$ , and  $f' \geq 0$ . Let now

$$\xi(x, t) := \zeta(\text{sdist}(x, t))\nabla\text{sdist}(x, t).$$

Second, we define  $\vartheta := f \circ \text{sdist}$  on  $U$ , which can be extended locally constantly to  $\Omega \times [0, T]$ . Third, let

$$B(x, t) = -\mu(\xi(x, t))\text{div}(D\sigma(\nabla\text{sdist}(\cdot, t))) \Big|_{p(x,t)} \xi(x, t) \quad \text{for } (x, t) \in U, \tag{5.9}$$

which can be extended to  $\Omega \times [0, T]$  by zero using a cutoff function.

To show that  $(\xi, B, \vartheta)$  is a gradient flow calibration for  $\{\mathcal{A}(t)\}_{t \in [0, T]}$  as in Definition 5.3, we rely on a generic equation for the signed distance of a smoothly evolving surface with normal velocity vector  $B$ , valid in the neighborhood  $U$ . Precisely, for a smoothly evolving surface, we have

$$\partial_t \text{sdist}(x, t) + B(p(x, t), t) \cdot \nabla \text{sdist}(x, t) = 0 \quad \text{for all } (x, t) \in U. \tag{5.10}$$

Supposing (5.10), we may differentiate this equation in space and replace  $\nabla \text{sdist}$  by  $\xi$  when it makes an explicit appearance but also in  $B$  (recalling (5.9)) to recover (5.1); note the error  $O(\text{dist}(x, \partial\mathcal{A}(t)))$  comes from the replacement of  $\nabla^2 \text{sdist}$  by  $\nabla \xi$  and  $\partial_t(\nabla \text{sdist})$ . Unsurprisingly, given the definition of  $\vartheta$  above and (5.10), we immediately obtain (5.3). To recover (5.2), we note that  $|\xi| \equiv 1$  on  $\cup_{t \in [0, T]} (\partial\mathcal{A}(t) \times \{t\})$  so that

$$(\partial_t + B(p(x, t), t) \cdot \nabla)|\xi|^2(x, t) = 0 \quad \text{for all } x \in \partial\mathcal{A}(t).$$

One can then explicitly verify that

$$\xi \cdot \nabla \left( (\partial_t + B(p(x, t), t) \cdot \nabla)|\xi|^2 \right) (x, t) = 0 \quad \text{for all } x \in \partial\mathcal{A}(t),$$

which allows one to recover the quadratic error in (5.2). Finally, the remaining properties of the calibration are easily verified by the choice of cutoff functions.

We briefly outline how to prove (5.10) at a space-time point  $(x, t) \in U$ . Suppose first that  $x \notin \mathcal{A}(t)$  and that surface velocity  $B(x, t) := B(p(x, t), t)$  (a hopefully forgivable abuse of notation) points from the interface in the direction of  $x$ . Straightforward geometric reasoning on the line segment  $x + \mathbb{R}\nabla\text{sdist}(x, t)$  then shows that

$$\partial_t \text{sdist}(x, t) \leq -B(x, t) \cdot \nabla \text{sdist}(x, t).$$

Supposing by contradiction that the previous inequality is strict, we have

$$\frac{\text{sdist}(x, t + \varepsilon) - \text{sdist}(x, t)}{\varepsilon} < -B(x, t) \cdot \nabla \text{sdist}(x, t) - \eta, \tag{5.11}$$

for  $\varepsilon > 0$  sufficiently small, where  $\eta > 0$  is a small fixed constant. Letting  $y_\varepsilon \in \partial\mathcal{A}(t + \varepsilon)$  be such that  $\text{sdist}(x, t + \varepsilon) = |x - y_\varepsilon|$ , we let  $\bar{y}_\varepsilon := p(y_\varepsilon, t)$ . One may argue by smoothness of the interface evolution and the projection operator that for small enough  $\varepsilon > 0$  (depending on  $\eta$ ), we have

$$|y_\varepsilon - \bar{y}_\varepsilon| \leq \varepsilon \left( \sup_{\substack{|x_0 - p(x, t)| \leq C\varepsilon \\ |t_0 - t| \leq \varepsilon}} |B(x_0, t_0)| \right) \leq \varepsilon(B(x, t) \cdot \nabla \text{sdist}(x, t) + \eta/2), \tag{5.12}$$

where on the right-hand side we used that  $B(x, t) \cdot \nabla \text{sdist}(x, t) = |B(x, t)|$  and  $C > 0$  is a constant depending only on the smooth interface. Synthesizing (5.11) and (5.12) we have

$$|x - \bar{y}_\varepsilon| \leq \text{sdist}(x, t + \varepsilon) + |y_\varepsilon - \bar{y}_\varepsilon| \leq \text{sdist}(x, t) - \varepsilon(\eta/2),$$

a contradiction to the definition of  $\text{sdist}(x, t)$ . This shows (5.10) at  $(x, t)$  under the contingent hypotheses; the other cases are verified similarly (reversing the flow in time when necessary).  $\square$

5.2 A Grönwall-type stability estimate

We wish to prove Theorem 5.2 by deriving a Grönwall-type estimate for a suitable quantity. A straightforward way to measure the difference between the calibrated evolution  $\mathcal{A}(t)$  and the weak solution  $A(t)$  ( $\chi_{A(t)}$  satisfying Definition 2.8) at a given time  $t \in [0, T]$  is the bulk error

$$\mathcal{B}[\chi(t)|\vartheta(t)] := \int_{A(t) \Delta \mathcal{A}(t)} |\vartheta(x, t)| dx.$$

Since  $|\vartheta(\cdot, t)|$  is strictly positive outside  $\partial \mathcal{A}(t)$ , and hence almost everywhere in  $\mathbb{R}^d$  this bulk error vanishes if and only if  $|\mathcal{A}(t) \Delta A(t)| = 0$ . However, the available stability estimate for the bulk error is not strong enough to apply Grönwall’s lemma immediately:

**Lemma 5.6.** *Let  $\{\mathcal{A}(t)\}_{t \in [0, T]}$  be a solution of anisotropic mean curvature flow (1.1) that is calibrated in the sense of Definition 5.3. Further, let  $\{A(t)\}_{t \in [0, T]}$  be time parametrized collection of sets with  $\chi := \chi_A$  a distributional solution of anisotropic mean curvature flow as in Definition 2.8. Recalling the definition of  $\mathcal{E}[\cdot|\cdot]$  in (4.28) (with  $c_0 = 1$ ), the following holds:*

(i) *The bulk error  $\mathcal{B}[\chi(T')|\vartheta(T')]$  at an arbitrary time  $T' \in [0, T]$  is given by*

$$\begin{aligned} \mathcal{B}[\chi(T')|\vartheta(T')] &= \mathcal{B}[\chi(0)|\vartheta(0)] + \int_0^{T'} \int_{\Omega} (\chi_{A(t)} - \chi_{\mathcal{A}(t)}) (\partial_t \vartheta + (B \cdot \nabla) \vartheta) dx dt \\ &\quad + \int_0^{T'} \int_{\Omega} (\chi_{A(t)} - \chi_{\mathcal{A}(t)}) \vartheta \operatorname{div} B dx dt \\ &\quad + \int_0^{T'} \int_{\partial^* A(t)} \vartheta B \cdot (\xi - \nu) d\mathcal{H}^{d-1} dt \\ &\quad + \int_0^{T'} \int_{\partial^* A(t)} \vartheta (V - B \cdot \xi) d\mathcal{H}^{d-1} dt. \end{aligned}$$

(ii) *For every  $\delta > 0$ , there exists a constant  $C(\delta) > 0$ , which also depends on the calibrated evolution, such that*

$$\begin{aligned} \mathcal{B}[\chi(T')|\vartheta(T')] &\leq \mathcal{B}[\chi(0)|\vartheta(0)] + C(\delta) \int_0^{T'} (\mathcal{B}[\chi(t)|\vartheta(t)] + \mathcal{E}[\chi(t)|\xi(t)]) dt \\ &\quad + \delta \int_0^{T'} \int_{\partial^* A(t)} |V - B \cdot \xi|^2 d\mathcal{H}^{d-1} dt \end{aligned} \tag{5.13}$$

for all  $T' \in [0, T]$ .

**Proof.** See [35, Sections 4.2, 4.3], where the argument is carried out for isotropic mean curvature flow. The same argument applies in our anisotropic setting since the definition of the bulk error functional has remained unchanged.  $\square$

To compensate for the additional error term on the right-hand side of (5.13), we make use of another stability estimate for the relative entropy:

**Lemma 5.7.** *Let the hypotheses of Lemma 5.6 hold. For all  $T' \in [0, T]$ , we have*

$$\begin{aligned} \mathcal{E}[\chi(T')|\xi(T')] &\leq \mathcal{E}[\chi(0)|\xi(0)] + C \int_0^{T'} \mathcal{E}[\chi(t)|\xi(t)] dt \\ &\quad - \frac{1}{4 \max_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |V - B \cdot \xi|^2 d\mathcal{H}^{d-1} dt. \end{aligned} \tag{5.14}$$

A combination of the two stability estimates yields

$$\mathcal{B}[\chi(T')|\vartheta(T')] + \mathcal{E}[\chi(T')|\xi(T')] \leq \mathcal{B}[\chi(0)|\vartheta(0)] + \mathcal{E}[\chi(0)|\xi(0)] + C \int_0^{T'} (\mathcal{B}[\chi(t)|\vartheta(t)] + \mathcal{E}[\chi(t)|\xi(t)]) dt.$$

In particular, it follows by Grönwall’s lemma that  $\mathcal{B}[\chi(T')|\vartheta(T')] = 0$  for all  $T' \in [0, T]$  if  $\mathcal{B}[\chi(T')|\vartheta(0)] + \mathcal{E}[\chi(T')|\xi(0)] = 0$ , which completes the proof of Theorem 5.2.

### 5.3 Stability of the relative entropy

The goal of this subsection is to prove Lemma 5.7.

With the help of the divergence theorem and the surface energy (1.2) with  $c_0 = 1$ , one can rewrite the relative entropy functional as

$$\mathcal{E}[\chi(t)|\xi(t)] = E[\chi(t)] - \int_{A(t)} \operatorname{div}(|\xi|\psi(|\xi|)D\sigma(\xi)) dx,$$

which using the distributional formulation of the time derivative from (2.19) leads to

$$\begin{aligned} \mathcal{E}[\chi(T')|\xi(T')] &= \mathcal{E}[\chi(0)|\xi(0)] + E[A(T')] - E[A(0)] \\ &\quad - \int_0^{T'} \int_{\partial^*A(t)} \operatorname{div}(|\xi|\psi(|\xi|)D\sigma(\xi)) V d\mathcal{H}^{d-1} dt \\ &\quad - \int_0^{T'} \int_{\partial^*A(t)} \partial_t(|\xi|\psi(|\xi|)D\sigma(\xi)) \cdot \nu d\mathcal{H}^{d-1} dt \\ &\leq \mathcal{E}[\chi(0)|\xi(0)] - \int_0^{T'} \int_{\partial^*A(t)} \frac{1}{\mu(\nu)} V^2 d\mathcal{H}^{d-1} dt \\ &\quad - \int_0^{T'} \int_{\partial^*A(t)} \operatorname{div}(|\xi|\psi(|\xi|)D\sigma(\xi)) V d\mathcal{H}^{d-1} dt \\ &\quad - \int_0^{T'} \int_{\partial^*A(t)} \partial_t(|\xi|\psi(|\xi|)D\sigma(\xi)) \cdot \nu d\mathcal{H}^{d-1} dt. \end{aligned}$$

Note, in the last step, we used the optimal energy dissipation inequality (2.20) for the weak solution  $\chi$ . We introduce the notation  $F(\xi) = |\xi|\psi(|\xi|)D\sigma(\xi)$  and  $M(\xi) := D_\xi F(\xi) = |\xi|\psi(|\xi|)D^2\sigma(\xi) + \left(\frac{\psi(|\xi|)}{|\xi|} + \psi'(|\xi|)\right)D\sigma(\xi) \otimes \xi$  and complete squares three times to write

$$\begin{aligned} &\mathcal{E}[\chi(T')|\xi(T')] \\ &\leq \mathcal{E}[\chi(0)|\xi(0)] + \frac{1}{2} \int_0^{T'} \int_{\partial^*A(t)} \left( - \left| \frac{1}{\sqrt{\mu(\nu)}} V - \frac{1}{\sqrt{\mu(\nu)}} B \cdot \nu \right|^2 \right. \\ &\quad \left. - \left| \sqrt{\mu(\nu)} \operatorname{div}(F(\xi)) + \frac{1}{\sqrt{\mu(\nu)}} V \right|^2 + \left| \sqrt{\mu(\nu)} \operatorname{div}(F(\xi)) + \frac{1}{\sqrt{\mu(\nu)}} B \cdot \nu \right|^2 \right) d\mathcal{H}^{d-1} dt \end{aligned}$$

$$\begin{aligned}
 & - \int_0^{T'} \int_{\partial^* A(t)} \frac{1}{\mu(v)} VB \cdot v d\mathcal{H}^{d-1} \int_0^{T'} \int_{\partial^* A(t)} \operatorname{div}(F(\xi))B \cdot v d\mathcal{H}^{d-1} dt \\
 & - \int_0^{T'} \int_{\partial^* A(t)} \partial_t (|\xi| \psi(|\xi|) D\sigma(\xi)) \cdot v d\mathcal{H}^{d-1} dt.
 \end{aligned} \tag{5.15}$$

The second square can be trivially estimated. We will now deal with the other squared terms as well as the remaining integrals separately:

The first square will later be used to compensate for a term in the stability estimate of the bulk error. To this end, it will be useful to estimate the term with the help of the inequality  $-a^2 \leq -\frac{1}{2}(a+b)^2 + b^2$  and (4.30):

$$\begin{aligned}
 & -\frac{1}{2} \int_0^{T'} \int_{\partial^* A(t)} \left| \frac{1}{\sqrt{\mu(v)}} V - \frac{1}{\sqrt{\mu(v)}} B \cdot v \right|^2 d\mathcal{H}^{d-1} dt \\
 & \leq -\frac{1}{4} \int_0^{T'} \int_{\partial^* A(t)} \left| \frac{1}{\sqrt{\mu(v)}} V - \frac{1}{\sqrt{\mu(v)}} B \cdot \xi \right|^2 d\mathcal{H}^{d-1} dt \\
 & \quad + \frac{1}{2} \int_0^{T'} \int_{\partial^* A(t)} \left| \frac{1}{\sqrt{\mu(v)}} B \cdot (v - \xi) \right|^2 d\mathcal{H}^{d-1} dt \\
 & \leq -\frac{1}{4 \max_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |V - B \cdot \xi|^2 d\mathcal{H}^{d-1} dt \\
 & \quad + \frac{1}{2 \min_{S^{d-1}} \mu} \|B\|_{C^0}^2 \int_0^{T'} \int_{\partial^* A(t)} |v - \xi|^2 d\mathcal{H}^{d-1} dt \\
 & \leq -\frac{1}{4 \max_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |V - B \cdot \xi|^2 d\mathcal{H}^{d-1} dt \\
 & \quad + C \int_0^{T'} \mathcal{E}[\chi(t)|\xi(t)] dt.
 \end{aligned} \tag{5.16}$$

Similarly, for the third square, the inequality  $a^2 \leq 2(a-b)^2 + 2b^2$  (applied twice), the Lipschitz continuity of  $\mu$ , the compatibility condition (5.4), and finally (4.30) yield

$$\begin{aligned}
 & \frac{1}{2} \int_0^{T'} \int_{\partial^* A(t)} \left| \sqrt{\mu(v)} \operatorname{div}(F(\xi)) + \frac{1}{\sqrt{\mu(v)}} B \cdot v \right|^2 d\mathcal{H}^{d-1} dt \\
 & \leq \frac{1}{2 \min_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |\mu(v) \operatorname{div}(F(\xi)) + B \cdot v|^2 d\mathcal{H}^{d-1} dt \\
 & \leq \frac{1}{\min_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |\mu(v) \operatorname{div}(F(\xi)) + B \cdot \xi|^2 d\mathcal{H}^{d-1} dt \\
 & \quad + \frac{1}{\min_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |B \cdot (v - \xi)|^2 d\mathcal{H}^{d-1} dt \\
 & \leq \frac{2}{\min_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |\mu(\xi) \operatorname{div}(F(\xi)) + B \cdot \xi|^2 d\mathcal{H}^{d-1} dt \\
 & \quad + \frac{2}{\min_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |(\mu(v) - \mu(\xi))M(\xi) : \nabla \xi|^2 d\mathcal{H}^{d-1} dt
 \end{aligned}$$

$$\begin{aligned}
 &+ \frac{1}{\min_{\mathbb{S}^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |B \cdot (v - \xi)|^2 d\mathcal{H}^{d-1} dt \\
 &\leq C \int_0^{T'} \int_{\partial^* A(t)} \min\{1, \text{dist}^2(\cdot, \partial^* \mathcal{A}(t))\} d\mathcal{H}^{d-1} dt \\
 &+ C \int_0^{T'} \int_{\partial^* A(t)} |v - \xi|^2 d\mathcal{H}^{d-1} dt \\
 &\leq C \int_0^{T'} \mathcal{E}[\chi(t)|\xi(t)] dt.
 \end{aligned} \tag{5.17}$$

It remains to show that

$$\begin{aligned}
 &\left| \int_0^{T'} \int_{\partial^* A(t)} \frac{1}{\mu(v)} VB \cdot v d\mathcal{H}^{d-1} dt \int_0^{T'} \int_{\partial^* A(t)} \text{div}(F(\xi))B \cdot v d\mathcal{H}^{d-1} dt \right. \\
 &+ \left. \int_0^{T'} \int_{\partial^* A(t)} \partial_t (|\xi| \psi(|\xi|) D\sigma(\xi)) \cdot v d\mathcal{H}^{d-1} dt \right| \\
 &\leq C \int_0^{T'} \mathcal{E}[\chi(t)|\xi(t)] dt.
 \end{aligned} \tag{5.18}$$

The third integral on the left-hand side can be expanded in a way that resembles the approximate evolution equations (5.1) and (5.2):

$$\begin{aligned}
 &\int_0^{T'} \int_{\partial^* A(t)} \partial_t (|\xi| \psi(|\xi|) D\sigma(\xi)) \cdot v d\mathcal{H}^{d-1} dt = \int_0^{T'} \int_{\partial^* A(t)} (v \otimes \partial_t \xi) : M(\xi) d\mathcal{H}^{d-1} dt \\
 &= \int_0^{T'} \int_{\partial^* A(t)} ((v - \xi) \otimes (\partial_t \xi + (B \cdot \nabla)\xi + (\nabla B)^T \xi)) : M(\xi) d\mathcal{H}^{d-1} dt \\
 &+ \int_0^{T'} \int_{\partial^* A(t)} (\xi \otimes (\partial_t \xi + (B \cdot \nabla)\xi)) : M(\xi) d\mathcal{H}^{d-1} dt \\
 &- \int_0^{T'} \int_{\partial^* A(t)} (v \otimes (B \cdot \nabla)\xi) : M(\xi) d\mathcal{H}^{d-1} dt \\
 &- \int_0^{T'} \int_{\partial^* A(t)} (\xi \otimes (v - \xi)) : ((\nabla B)M(\xi)^T) d\mathcal{H}^{d-1} dt.
 \end{aligned} \tag{5.19}$$

Here, the first two integrals in the right-hand side expression are already in a suitable form: For the first integral this can be seen directly from (5.1), whereas for the second term a computation yields

$$\begin{aligned}
 (\xi \otimes (\partial_t \xi + (B \cdot \nabla)\xi)) : M(\xi) &= |\xi| \psi(|\xi|) \xi \cdot D^2 \sigma(\xi) (\partial_t \xi + (B \cdot \nabla)\xi) \\
 &+ \left( \frac{\psi(|\xi|)}{|\xi|} + \psi'(|\xi|) \right) (D\sigma(\xi) \cdot \xi) (\xi \cdot (\partial_t \xi + (B \cdot \nabla)\xi)) \\
 &= |\xi| \psi(|\xi|) (\partial_t \xi + (B \cdot \nabla)\xi) \cdot D^2 \sigma(\xi) \xi \\
 &+ \left( \frac{\psi(|\xi|)}{|\xi|} + \psi'(|\xi|) \right) \sigma(\xi) \xi \cdot (\partial_t \xi + (B \cdot \nabla)\xi) \\
 &= \left( \frac{\psi(|\xi|)}{|\xi|} + \psi'(|\xi|) \right) \sigma(\xi) \xi \cdot (\partial_t \xi + (B \cdot \nabla)\xi),
 \end{aligned}$$

so that one can invoke (5.2). Here we have used a fact that follows from the positive 0-homogeneity of  $D\sigma$ , namely that  $D^2\sigma(\xi)\xi = \frac{d}{ds}\Big|_{s=0} D\sigma(e^s\xi) = \frac{d}{ds}\Big|_{s=0} D\sigma(\xi) = 0$ .

Furthermore, an application of the chain rule shows that

$$(v \otimes (B \cdot \nabla)\xi) : M(\xi) = v \cdot (B \cdot \nabla)(F(\xi)),$$

and we will from now on use this slightly shorter formulation for the third right-hand side integral of (5.19).

By using (5.19) and plugging in the computations for the three right-hand side integrals as well as the weak formulation of anisotropic mean curvature flow in (2.19), we obtain

$$\begin{aligned} & \left| \int_0^{T'} \int_{\partial^* A(t)} \frac{1}{\mu(v)} VB \cdot v d\mathcal{H}^{d-1} dt \int_0^{T'} \int_{\partial^* A(t)} \operatorname{div}(F(\xi))B \cdot v d\mathcal{H}^{d-1} dt \right. \\ & \left. + \int_0^{T'} \int_{\partial^* A(t)} \partial_t (|\xi| \psi(|\xi|) D\sigma(\xi)) \cdot v d\mathcal{H}^{d-1} dt \right| \\ & \leq \left| - \int_0^{T'} \int_{\partial^* A(t)} \nabla B : (\sigma(v)I_d - v \otimes D\sigma(v)) d\mathcal{H}^{d-1} dt \right. \\ & \quad + \int_0^{T'} \int_{\partial^* A(t)} \operatorname{div}(F(\xi))B \cdot v d\mathcal{H}^{d-1} dt \\ & \quad - \int_0^{T'} \int_{\partial^* A(t)} v \cdot (B \cdot \nabla) (F(\xi)) d\mathcal{H}^{d-1} dt \\ & \quad \left. - \int_0^{T'} \int_{\partial^* A(t)} (\xi \otimes (v - \xi)) : ((\nabla B)M(\xi)^T) d\mathcal{H}^{d-1} dt \right| \\ & \quad + C \int_0^{T'} \mathcal{E}[\chi(t)|\xi(t)] dt. \tag{5.20} \end{aligned}$$

It is now advisable to expand the first right-hand side integral and to add zero in order to isolate two more errors controlled by the time-integrated relative entropy:

$$\begin{aligned} & - \int_0^{T'} \int_{\partial^* A(t)} \nabla B : (\sigma(v)I_d - v \otimes D\sigma(v)) d\mathcal{H}^{d-1} dt \\ & = - \int_0^{T'} \int_{\partial^* A(t)} \sigma(v) \operatorname{div} B d\mathcal{H}^{d-1} dt \\ & \quad + \int_0^{T'} \int_{\partial^* A(t)} v \cdot (F(v) \cdot \nabla) B d\mathcal{H}^{d-1} dt \\ & = - \int_0^{T'} \int_{\partial^* A(t)} (\sigma(v) - F(\xi) \cdot v) \operatorname{div} B d\mathcal{H}^{d-1} dt \\ & \quad + \int_0^{T'} \int_{\partial^* A(t)} (v - \xi) \cdot ((F(v) - F(\xi)) \cdot \nabla) B d\mathcal{H}^{d-1} dt \end{aligned}$$



$$\begin{aligned}
 & - \int_0^{T'} \int_{\partial^* A(t)} F(\xi) \cdot \nu \operatorname{div} B \, d\mathcal{H}^{d-1} dt \\
 & + \int_0^{T'} \int_{\partial^* A(t)} \nu \cdot (F(\xi) \cdot \nabla) B \, d\mathcal{H}^{d-1} dt \\
 & + \int_0^{T'} \int_{\partial^* A(t)} \xi \cdot ((F(v) - F(\xi)) \cdot \nabla) B \, d\mathcal{H}^{d-1} dt. \tag{5.21}
 \end{aligned}$$

The first two integrals are controlled by  $\int_0^{T'} \mathcal{E}[\chi(t)|\xi(t)] dt$ ; in the case of the second integral this follows from the local Lipschitz continuity of  $F$ .

We approach the last integral in (5.21) by an application of Taylor’s theorem and show that it cancels with the second to last integral in (5.20). Indeed, one computes that

$$\begin{aligned}
 & \xi \cdot ((F(v) - F(\xi)) \cdot \nabla) B - (\xi \otimes (v - \xi)) : ((\nabla B)M(\xi)^T) \\
 & = \xi \cdot ((M(\xi)(v - \xi)) \cdot \nabla) B - \xi \cdot (\nabla B)M(\xi)^T(v - \xi) \\
 & \quad + \xi \cdot \left( \int_0^1 (1-r) ((v - \xi) \cdot \nabla) M(rv + (1-r)\xi)(v - \xi) dr \cdot \nabla \right) B \\
 & = \xi \cdot (\nabla B) (M(\xi) - M(\xi)^T) (v - \xi) \\
 & \quad + \xi \cdot \left( \int_0^1 (1-r) ((v - \xi) \cdot \nabla) M(rv + (1-r)\xi)(v - \xi) dr \cdot \nabla \right) B \\
 & = \left( \frac{\psi(|\xi|)}{|\xi|} + \psi'(|\xi|) \right) \xi \cdot (\nabla B) (D\sigma(\xi) \otimes \xi - \xi \otimes D\sigma(\xi)) (v - \xi) \\
 & \quad + \xi \cdot \left( \int_0^1 (1-r) ((v - \xi) \cdot \nabla) M(rv + (1-r)\xi)(v - \xi) dr \cdot \nabla \right) B \\
 & = \left( \frac{\psi(|\xi|)}{|\xi|} + \psi'(|\xi|) \right) (\xi \cdot (\nabla B) D\sigma(\xi)) \xi \cdot (v - \xi) \\
 & \quad - \left( \frac{\psi(|\xi|)}{|\xi|} + \psi'(|\xi|) \right) (D\sigma(\xi) \cdot (v - \xi)) \xi \cdot (\xi \cdot \nabla) B \\
 & \quad + \xi \cdot \left( \int_0^1 (1-r) ((v - \xi) \cdot \nabla) M(rv + (1-r)\xi)(v - \xi) dr \cdot \nabla \right) B, \tag{5.22}
 \end{aligned}$$

where now each summand is controlled by  $\sigma(v) - |\xi| \psi(|\xi|) D\sigma(\xi) \cdot \nu$  by Lemma 5.4 (iii) and (vi). Here we use that  $\sigma \in C^3(\mathbb{R}^d \setminus \{0\})$ .

Plugging in (5.21) and (5.22) into (5.20) yields

$$\begin{aligned}
 & \left| \int_0^{T'} \int_{\partial^* A(t)} \frac{1}{\mu(v)} \nu B \cdot \nu \, d\mathcal{H}^{d-1} dt \int_0^{T'} \int_{\partial^* A(t)} \operatorname{div}(F(\xi)) B \cdot \nu \, d\mathcal{H}^{d-1} dt \right. \\
 & \quad \left. + \int_0^{T'} \int_{\partial^* A(t)} \partial_t (|\xi| \psi(|\xi|) D\sigma(\xi)) \cdot \nu \, d\mathcal{H}^{d-1} dt \right|
 \end{aligned}$$

$$\begin{aligned}
 &= \left| - \int_0^{T'} \int_{\partial^* A(t)} F(\xi) \cdot \nu \operatorname{div} B \, d\mathcal{H}^{d-1} dt \right. \\
 &\quad + \int_0^{T'} \int_{\partial^* A(t)} \nu \cdot (F(\xi) \cdot \nabla) B \, d\mathcal{H}^{d-1} dt \\
 &\quad + \int_0^{T'} \int_{\partial^* A(t)} \operatorname{div}(F(\xi)) B \cdot \nu \, d\mathcal{H}^{d-1} dt \\
 &\quad \left. - \int_0^{T'} \int_{\partial^* A(t)} \nu \cdot (B \cdot \nabla) (F(\xi)) \, d\mathcal{H}^{d-1} dt \right| \\
 &\quad + C \int_0^{T'} \mathcal{E}[\chi(t)|\xi(t)] dt. \tag{5.23}
 \end{aligned}$$

The four remaining surface integrals can be written as one integral involving the divergence of a matrix field since

$$\operatorname{div} (B \otimes F(\xi) - F(\xi) \otimes B) = \operatorname{div}(F(\xi))B + (F(\xi) \cdot \nabla)B - \operatorname{div}(B)F(\xi) - (B \cdot \nabla)(F(\xi)).$$

As a result of an integration by parts and the symmetry relation  $\operatorname{div}(\operatorname{div}(a \otimes b)) = \operatorname{div}(\operatorname{div}(b \otimes a))$ , one can then see that

$$\begin{aligned}
 &- \int_0^{T'} \int_{\partial^* A(t)} F(\xi) \cdot \nu \operatorname{div} B \, d\mathcal{H}^{d-1} dt \\
 &+ \int_0^{T'} \int_{\partial^* A(t)} \nu \cdot (F(\xi) \cdot \nabla) B \, d\mathcal{H}^{d-1} dt \\
 &+ \int_0^{T'} \int_{\partial^* A(t)} \operatorname{div}(F(\xi)) B \cdot \nu \, d\mathcal{H}^{d-1} dt \\
 &- \int_0^{T'} \int_{\partial^* A(t)} \nu \cdot (B \cdot \nabla) (F(\xi)) \, d\mathcal{H}^{d-1} dt \\
 &= \int_0^{T'} \int_{\partial^* A(t)} \nu \cdot \operatorname{div} (B \otimes F(\xi) - F(\xi) \otimes B) \, d\mathcal{H}^{d-1} dt \\
 &= \int_0^{T'} \int_{A(t)} \operatorname{div} (\operatorname{div} (B \otimes F(\xi) - F(\xi) \otimes B)) \, dx dt \\
 &= 0. \tag{5.24}
 \end{aligned}$$

Plugging in (5.24) into (5.23) yields (5.18).

In a last step, plugging in (5.16), (5.17) and (5.18) into (5.15) yields

$$\begin{aligned}
 \mathcal{E}[\chi(T')|\xi(T')] &\leq \mathcal{E}[\chi(0)|\xi(0)] + C \int_0^{T'} \mathcal{E}[\chi(t)|\xi(t)] dt \\
 &\quad - \frac{1}{4 \max_{S^{d-1}} \mu} \int_0^{T'} \int_{\partial^* A(t)} |V - B \cdot \xi|^2 \, d\mathcal{H}^{d-1} dt.
 \end{aligned}$$

This concludes the proof of Lemma 5.7.

## 6. Conclusion

In this paper, we began by constructing solutions of the anisotropic Allen–Cahn equation using relatively elementary arguments, relying on convexity for instance, for a naturally associated minimizing movements scheme. With this, we showed that these solutions are precompact in an appropriate functional space and, as the interface width vanishes, converge to a solution of anisotropic mean curvature flow. In contrast to previous methods, our approach relies on control of a relative entropy, which quantifies the distance between the sharp interface and smooth interface coming from the approximation. Having access to the relative entropy naturally allows for further understanding of the BV-solution concept associated with the anisotropic curvature flow. Precisely, we proved that solutions of the anisotropic mean curvature flow, defined in terms of a distributional formulation of the curvature and an optimal energy dissipation relation, actually possess a weak–strong uniqueness property: meaning that when classical, smooth solutions exist, the weak solution with the same initial condition coincides with the smooth solution. Future questions will involve further investigation of convergence in the multiphase setting and obtaining quantitative convergence rates.

**Competing interests.** None.

**Acknowledgements.** The present paper is an extension of the third author’s master’s thesis at the University of Bonn.

**Financial Support.** This project had received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2047/1 – 390685813 and the Graduiertenkolleg (Research Training Group) 2339 “Interfaces, Complex Structures, and Singular Limits in Continuum Mechanics”.

## References

- [1] Almgren, F., Taylor, J. E. & Wang, L. (1993) Curvature-driven flows: A variational approach. *SIAM J. Control Optim.* **31**(2), 387–438.
- [2] Ambrosio, L., Fusco, N. & Pallara, D. (2006). *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs. Oxford University Press, Oxford, UK – New York. Reprint.
- [3] Angenent, S. & Gurtin, M. (1989) Multiphase thermomechanics with interfacial structure 2. Evolution of an isothermal interface. *Arch. Ration. Mech. Anal.* **108**(3), 323–391.
- [4] Barrett, J. W., Garcke, H. & Nürnberg, R. (2008) A variational formulation of anisotropic geometric evolution equations in higher dimensions. *Numer. Math.* **109**(1), 1–44.
- [5] Bazant, M. Z. (2013) Theory of chemical kinetics and charge transfer based on nonequilibrium thermodynamics. *Acc. Chem. Res.* **46**(5), 03–1160.
- [6] Bellettini, G. & Novaga, M. (2000) Approximation and comparison for nonsmooth anisotropic motion by mean curvature in  $\mathbb{R}^N$ . *Math. Models Methods Appl. Sci.* **10**(01), 1–10.
- [7] Bellettini, G. & Paolini, M. (1996) Anisotropic motion by mean curvature in the context of finler geometry. *Hokkaido Math. J.* **25**(3), 537–566.
- [8] Bogomol’nyi, E. B. (1976) The stability of classical solutions. *Soviet J. Nuclear Phys.* **24**(4), 861–870.
- [9] Bouchitté, G. (1990) Singular perturbations of variational problems arising from a two-phase transition model. *Appl. Math. Optim.* **21**(1), 289–314.
- [10] Braides, A. (2002). *Gamma-Convergence for Beginners*, vol. 22. Oxford University Press, Oxford.
- [11] Chambolle, A., De Gennaro, D. & Morini, M. (2022) Minimizing movements for anisotropic and inhomogeneous mean curvature flows. *arXiv preprint*, arXiv: 2212.05027.
- [12] Chambolle, A., Morini, M., Novaga, M. & Ponsiglione, M. (2019) Existence and uniqueness for anisotropic and crystalline mean curvature flows. *J. Am. Math. Soc.* **32**(3), 779–824.
- [13] Chambolle, A., Morini, M. & Ponsiglione, M. (2017) Existence and uniqueness for a crystalline mean curvature flow. *Comm. Pure Appl. Math.* **70**(6), 1084–1114.
- [14] Chambolle, A. & Novaga, M. (2007) Approximation of the anisotropic mean curvature flow. *Math. Models Methods Appl. Sci.* **17**(06), 833–844.
- [15] Chen, L.-Q. (2002) Phase-field models for microstructure evolution. *Annu. Rev. Mater. Res.* **32**(1), 113–140.
- [16] Chen, X. (1992) Generation and propagation of interfaces for reaction-diffusion equations. *J. Differential Equ.* **96**(1), 116–141.
- [17] Chen, Y.-G., Giga, Y. & Goto, S. (1991) Uniqueness and existence of viscosity solutions of generalized mean curvature equations. *J. Differ. Geom.* **33**(3), 749–786.
- [18] Cicalese, M., Nagase, Y. & Pisante, G. (2010) The Gibbs-Thomson relation for non homogeneous anisotropic phase transitions. *Adv. Calc. Var.* **3**(3), 321–344.

- [19] Maso, G. Dal (1993). *An introduction to  $\Gamma$ -convergence*. *Progress in Nonlinear Differential Equations and their Applications*, vol. **8**, Birkhäuser Boston, Inc, Boston, MA.
- [20] Dziuk, G. (1999) Discrete anisotropic curve shortening flow. *SIAM J. Numer. Anal.* **36**(6), 1808–1830.
- [21] Elliott, C. M. & Schätzle, R. (1997) The limit of the fully anisotropic double-obstacle Allen–Cahn equation in the nonsmooth case. *SIAM J. Math. Anal.* **28**(2), 273–303.
- [22] Fischer, J. & Hensel, S. (2020) Weak-strong uniqueness for the Navier–Stokes equation for two fluids with surface tension. *Arch. Ration. Mech. Anal.* **236**(2), 967–1087.
- [23] Fischer, J., Hensel, S., Laux, T. & Simon, T. M. (2021) The local structure of the energy landscape in multiphase mean curvature flow: Weak-strong uniqueness and stability of evolutions, *arXiv preprint*, arXiv: 2003.05478.
- [24] Fischer, J., Laux, T. & Simon, T. M. (2020) Convergence rates of the Allen–Cahn equation to mean curvature flow: A short proof based on relative entropies. *SIAM J. Math. Anal.* **52**(6), 6222–6233.
- [25] Fischer, J. & Marveggio, A. (2022) Quantitative convergence of the vectorial Allen–Cahn equation towards multiphase mean curvature flow. *Annales De l'Institut Henri Poincaré C, Analyse Non Linéaire* *arXiv preprint*, arXiv: 2203.17143.
- [26] Fonseca, I. & Leoni, G. Modern Methods in the Calculus of Variations: Sobolev Spaces. *In preparation, accepted for publication by Springer*.
- [27] Fonseca, I. & Tartar, L. (1989) The gradient theory of phase transitions for systems with two potential wells. *Proc. R. Soc. Lond. Ser. A* **111**(1-2), 89–102.
- [28] Garcke, H., Nestler, B. & Stoth, B. (1998) On anisotropic order parameter models for multi-phase systems and their sharp interface limits. *Phys. D Nonlinear Phenomena* **115**(1), 87–108.
- [29] Giga, M. & Giga, Y. (2001) Generalized motion by nonlocal curvature in the plane. *Arch. Ration. Mech. Anal.* **159**(4), 295–333.
- [30] Giga, Y. (2006). *Surface Evolution Equations: A Level Set Approach*, vol. 99. *Monographs in Mathematics*. Birkhäuser, Basel, Switzerland – Boston, MA – Berlin, Germany.
- [31] Giga, Y., Ohtsuka, T. & Schätzle, R. (2006) On a uniform approximation of motion by anisotropic curvature by the Allen–Cahn equations. *Interfaces Free Bound* **8**(3), 317–348.
- [32] Giga, Y. & Požár, N. (2016) A level set crystalline mean curvature flow of surfaces. *Adv. Diff. Equ.* **7**(8), 631–698.
- [33] Giga, Y. & Požár, N. (2018) Approximation of general facets by regular facets with respect to anisotropic total variation energies and its application to crystalline mean curvature flow. *Comm. Pure Appl. Math.* **71**(7), 1461–1491.
- [34] Gilbarg, D. & Trudinger, N. S. (2001). *Elliptic Partial Differential Equations of Second Order*, *Classics in Mathematics*. 2nd ed. Springer-Verlag, Berlin, Germany – Heidelberg, Germany, Reprint of the 1998 ed.
- [35] Hensel, S. & Laux, T. (2021) BV solutions for mean curvature flow with constant contact angle: Allen–Cahn approximation and weak-strong uniqueness. *Indiana Univ. Math. J. (online first)*.
- [36] Hensel, S. & Laux, T. (2021) A new varifold solution concept for mean curvature flow: Convergence of the Allen–Cahn equation and weak-strong uniqueness, *arXiv preprint*, arXiv: 2109.04233.
- [37] Hensel, S. & Moser, M. (2022) Convergence rates for the Allen–Cahn equation with boundary contact energy: The non-perturbative regime. *Calc. Var. Partial Differ. Equ.* **61**(201).
- [38] Kim, I., Mellet, A. & Wu, Y. (2022) Density-constrained chemotaxis and Hele-Shaw flow. *arXiv preprint*, arxiv: 2204.11917.
- [39] Laux, T. (2020). A gradient-flow approach for the convergence of the anisotropic Allen–Cahn equation, *Geometric Aspects of Solutions to Partial Differential Equations*, Vol. 2172, *RIMS Kôkyûroku*, Research Institute for Mathematical Sciences, Kyoto, Japan, pp. 32–42.
- [40] Laux, T. (2022) Weak-strong uniqueness for volume-preserving mean curvature flow. *Rev. Mat. Iberoam. (online first)*.
- [41] Laux, T. & Otto, F. (2016) Convergence of the thresholding scheme for multi-phase mean-curvature flow. *Calc. Var. Partial Differ. Equ.* **129**(5).
- [42] Laux, T. & Simon, T. M. (2018) Convergence of the Allen-Cahn equation to multiphase mean curvature flow. *Comm. Pure Appl. Math.* **71**(8), 1597–1647.
- [43] Laux, T. & Stinson, K. (2023) Sharp interface limit of the Cahn–Hilliard reaction model for lithium-ion batteries. *Math. Models Methods Appl. Sci.* **33**(12), 2557–2585.
- [44] Leoni, G. (2013). *Gamma Convergence and Applications to Phase Transitions*, Lecture Notes from Carnegie Mellon University.
- [45] Lindenstrauss, J. & Tzafriri, L. (1979). *Classical Banach spaces II*, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Vol. **97**, Springer-Verlag, Berlin, Germany – Heidelberg, Germany.
- [46] Luckhaus, S. & Modica, L. (1989) The Gibbs–Thompson relation within the gradient theory of phase transitions. *Arch. Ration. Mech. Anal.* **107**(1), 71–83.
- [47] Luckhaus, S. & Sturzenhecker, T. (1995) Implicit time discretization for the mean curvature equation. *Calc. Var. Partial Differ. Equ.* **3**(2), 253–271.
- [48] Modica, L. (1987) The gradient theory of phase transitions and the minimal interface criterion. *Arch. Ration. Mech. Anal.* **98**(2), 123–142.
- [49] Modica, L. & Mortola, S. (1977) Un esempio di  $\Gamma$ -convergenza. *Boll. Unione Mat. Ital.* **14-B**, 285–299.
- [50] Rockafellar, R. T. (1972). *Convex Analysis*, Princeton Mathematical Series., Vol. **28**, Princeton, NJ.: Princeton University Press, 2nd printing.
- [51] Serfaty, S. (2011) Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.* **31**(4), 1427–1451.
- [52] Sternberg, P. (1988) The effect of a singular perturbation on nonconvex variational problems. *Arch. Ration. Mech. Anal.* **101**(3), 209–260.

- [53] Taylor, J. E. (1978) Crystalline variational problems. *Bull. Amer. Math. Soc.* **84**(4), 568–588.
- [54] Taylor, J. E., Cahn, J. W. & Handwerker, C. A. (1992) Geometric models of crystal growth. *Acta Metall. Mater.* **40**(7), 1443–1474.