

A NEW VARIFOLD SOLUTION CONCEPT FOR MEAN CURVATURE FLOW: CONVERGENCE OF THE ALLEN–CAHN EQUATION AND WEAK-STRONG UNIQUENESS

SEBASTIAN HENSEL AND TIM LAUX

ABSTRACT. We propose a new weak solution concept for (two-phase) mean curvature flow which enjoys both (unconditional) existence and (weak-strong) uniqueness properties. These solutions are evolving varifolds, just as in Brakke’s formulation, but are coupled to the phase volumes by a simple transport equation. First, we show that, in the exact same setup as in Ilmanen’s proof [J. Differential Geom. 38, 417–461, (1993)], any limit point of solutions to the Allen–Cahn equation is a varifold solution in our sense. Second, we prove that any calibrated flow in the sense of Fischer et al. [arXiv:2003.05478]—and hence any classical solution to mean curvature flow—is unique in the class of our new varifold solutions. This is in sharp contrast to the case of Brakke flows, which a priori may disappear at any given time and are therefore fatally non-unique. Finally, we propose an extension of the solution concept to the multi-phase case which is at least guaranteed to satisfy a weak-strong uniqueness principle.

Keywords: Mean curvature flow, gradient flows, varifolds, weak solutions, weak-strong uniqueness, calibrated geometry, gradient-flow calibrations.

Mathematical Subject Classification: 53E10 (primary), 49Q20, 35K57, 35Q49, 28A75.

1. INTRODUCTION

Weak solution concepts for mean curvature flow have been investigated since the seminal work of Brakke [8]. Remarkably, Brakke constructed his weak solution more than a decade before Gage and Hamilton published their treatise of closed convex curves [25]. His solution concept is based on the theory of varifolds, a measure-theoretic generalization of embedded surfaces also used in this work. It has been applied in a broad context as it enjoys excellent compactness properties, and—thanks to Huisken’s monotonicity formula [27]—a partial regularity theory. However, it is well-known that a priori Brakke’s formulation lacks any uniqueness or continuity in time: at any given time the solution (or parts of it) might instantly disappear. More recently, Kim and Tonegawa [32] showed that a variant of Brakke’s original construction not only yields a Brakke flow, but that in addition the enclosed volume changes continuously, so that at least the sudden disappearance of the solution can be ruled out for the particular Brakke flow constructed there.

In the present work, we introduce a new notion of weak solutions to mean curvature flow, which is also based on varifolds, enjoys basically the same compactness and existence properties, but in addition does not allow the un-physical time-discontinuities and non-uniqueness of Brakke’s formulation. In some sense, our weak formulation lives between the concepts of Brakke’s solution and the distributional (or BV) solution introduced by Luckhaus and Sturzenhecker [38]. The latter also enjoys a uniqueness theory as our new concept (see [24]), but all existence proofs are either conditional (see e.g. [38, 35, 37]), or rest on additional geometric properties [15].

There is yet another weak solution concept, the viscosity (or levelset) solution [19, 11], which satisfies both existence and uniqueness (if one allows for “fattening” of levelsets) and is compatible with all other solution concepts mentioned above in the sense that a.e. levelset is a solution in all three senses [20, 33]. The main drawback of this concept is that it relies on the comparison principle, which only applies in the simple case of two phases. All other notions mentioned above (including our new notion) can be formulated in the more general case of multiple phases

(or surface clusters) as they solely rely on the gradient-flow structure of (multi-phase) mean curvature flow.

Of all the weak solution concepts for mean curvature flow, the one which we propose in the present work is the concept which exploits the gradient-flow structure in the strongest form as it mirrors the concepts known for general gradient flows as introduced by De Giorgi [14, 4], and Sandier and Serfaty [49]. We refer the reader interested in the general framework to the book by Ambrosio, Gigli, and Savaré [5], and Mielke's lecture notes [40]. Ever since the seminal work of Jordan, Kinderlehrer, and Otto [31], the study of (Wasserstein) gradient-flows in the context of partial differential equations has received continuous attention. Formally speaking, a gradient flow is the steepest descent in an energy landscape, which means that for a given energy functional E on a (usually infinitely dimensional curved) space \mathcal{M} equipped with a metric tensor g , one considers the equation

$$(1) \quad \frac{du}{dt} = -\nabla_g E(u),$$

where $\nabla_g E(u) \in T_u \mathcal{M}$ is the tangent vector dual to the differential $dE(u)$, i.e., $g_u(\nabla_g E(u), v) = dE(u).v$ for all tangent vectors $v \in T_u \mathcal{M}$. If \mathcal{M} is a Riemannian manifold, it is easy to construct solutions to this system of ordinary differential equations, but this formal picture becomes more subtle in the infinite dimensional case as the nonlinearity $-\nabla_g E(u)$ may not interact well with weak convergence. However, it is easily seen that, when \mathcal{M} , E , and u are sufficiently smooth so that one can apply the chain rule to $E(u(t))$, that the validity of (1) on a given time interval $(0, T)$ is equivalent to the optimal energy-dissipation inequality

$$(2) \quad E(u(T)) + \frac{1}{2} \int_0^T g_{u(t)} \left(\frac{du}{dt}, \frac{du}{dt} \right) dt + \frac{1}{2} \int_0^T g_{u(t)} \left(\nabla_g E(u(t)), \nabla_g E(u(t)) \right) dt \leq E(u(0)).$$

The beauty of this observation is that one only needs to prove an *inequality* for the nonlinear terms in the equation.

In the case of mean curvature flow, still speaking formally, the energy E is the area functional; the space \mathcal{M} is the space of all (embedded) n -dimensional surfaces in some fixed manifold, say, \mathbb{R}^d with $d > n$; the tangent space $T_\Sigma \mathcal{M}$ at a given configuration $\Sigma \in \mathcal{M}$ can be identified with normal velocity fields V on Σ ; and the metric tensor is given by the standard L^2 -structure $g_\Sigma(V, V) = \int_\Sigma |V|^2 d\omega^n$. To turn this formal picture into rigorous analysis is particularly subtle in the case of mean curvature flow as this metric is completely degenerate, a phenomenon which was first recorded by Michor and Mumford [39]. Geometrically speaking, it is the invariance with respect to reparametrizations which causes this degeneracy and allows to move surfaces at arbitrarily small cost when introducing infinitesimal wrinkles. From a PDE viewpoint, one of the key differences between mean curvature flow and more regular Wasserstein gradient flows is that in the latter case, the evolution equation is based on the continuity equation, while the evolution equation for mean curvature flow is based on the transport equation. This degeneracy is the reason why Almgren, Taylor, and Wang [3], and Luckhaus and Sturzenhecker [38] used a proxy for the induced distance of surfaces in their minimizing movements approximation.

Together with Felix Otto [36] and in the general multi-phase case with Jona Lelmi [34], the second author introduced a concept similar to the one here, but still in the framework of sets of finite perimeter, which intrinsically lacks compactness: the perimeter functional is only lower semi-continuous, and in general not continuous. This is why these two convergence proofs are only conditional in the sense that one has to *assume* that the time-integrated energies converge to those of the limit. Under certain assumptions on the geometry in the two-phase case, e.g. mean convexity $H \geq 0$, one can in fact verify this assumption [15]. Geometrically, this assumption rules out the piling up of several layers of surfaces and therefore guarantees the limit varifold to have unit density. Such an assumption is not necessary in our case here.

In the present work, we prove existence and (weak-strong) uniqueness of our solution concept. To the best of our knowledge, this is the first concept which satisfies these two properties and does not rely on the comparison principle. More precisely, we show that solutions to the Allen–Cahn

equation, up to passing to subsequences, converge to solutions of our new concept. This part is inspired by Ilmanen's fundamental work [28], which shows that any such limit is a Brakke solution. However, in our case, a new major difficulty arises since one needs to construct a normal velocity field in the sense of the new notion introduced in the present paper. Let us also mention that our natural square integrability of this normal velocity field, which appears in a transport-type equation for the indicator function of the enclosed volume automatically implies the (Hölder-) continuity of this volume. This is to be compared with the already mentioned work of Kim and Tonegawa [32], which requires a substantial effort to provide the construction of a Brakke flow supported on a network of evolving interfaces associated with a continuously evolving family of phase volumes.

The relation between the Allen–Cahn equation and mean curvature flow has been an ongoing investigation for several decades. We will only mention a few key results and refer to the introduction of [37] for more details. Bronsard and Kohn [9] pointed out the gradient-flow structure of the Allen–Cahn equation, and investigated the radially symmetric case. Evans, Soner, and Souganidis [18] proved the convergence to the viscosity solution. Ilmanen's convergence proof to a Brakke flow [28], which was fundamental to many following works on the Allen–Cahn equation (including the construction in the present work), has been generalized to several different settings, for example to the case of boundary contact [41, 42] or when the underlying space is a manifold with lower bounds on the Ricci curvature, see [47, 46]. Furthermore, Tonegawa [52] showed that the limiting varifold in Ilmanen's proof has integer multiplicity. Together with Fischer and Simon, one of the authors has recently derived the optimal convergence rate for the Allen–Cahn equation to two-phase mean curvature flow solely relying on the gradient-flow structure. At the moment, the only convergence proof in the vector-valued case is the conditional convergence to a BV solution by Simon and one of the authors [37]. Recent advances in the static case include Del Pino, Kowalczyk, and Wei [16] who derived counter-examples to a conjecture of De Giorgi by constructing entire stationary solutions to the Allen–Cahn equation satisfying a monotonicity condition. Chodosh and Mantoulidis [12] have settled a question on the unit density (and on the lower bound of the index) of limits of stationary solutions to the Allen–Cahn equation on 3-dimensional manifolds with generic metrics.

In the second part of the paper, we show that our varifold solution does not allow for unphysical non-uniqueness before the onset of singularities. More precisely, we establish that as long as a classical solution to mean curvature flow (in the sense of a smooth family of parametrized surfaces moving along their mean curvature) exists, all varifold solutions in the sense of the present work coincide with this classical flow. Our proof of this second theorem generalizes our recent result with Fischer and Simon [24] to the case of such varifold solutions for (two-phase) mean curvature flow, and is based on a relative entropy technique.

The basic idea of this approach is to construct a quantity which on one side captures the distance between a weak and a classical solution in a sufficiently strong sense, and on the other side allows for an estimate on its time evolution. The latter requirement makes this task non-trivial due to the typically limited regularity of weak solutions. However, many mathematical models in continuum mechanics are equipped with a natural (non-linear) functional of the solution: the associated energy. Moreover, control over the time evolution of the energy is often provided by means of an energy dissipation principle. In such a setting, the general recipe for the construction of a relative entropy is to start with the energy of the weak solution and to subtract from it a suitable non-linear functional of the strong solution which only depends in a linear way on the weak solution. This particular structure indeed gives hope to compute its time evolution as one (at least in principle) only relies on the sharp energy dissipation inequality, the weak formulation of the problem, and finally the higher regularity of a classical solution.

The classical setup in which this approach has been applied concerns continuum mechanical models equipped with a strictly convex and dissipated energy functional. In fact, in this context the method of relative entropies dates back to the works of Dafermos [13] and DiPerna [17] on conservation laws in the late 70's, and has, for instance, also been used in the context of

mathematical fluid mechanics (see, e.g., [51, 21, 22, 53]). In these models, one subtracts from the energy functional its first-order Taylor expansion with base point located at the classical solution. By strict convexity of the energy, one indeed obtains in this way a coercive error functional.

The relative entropy approach was recently extended to cover basic models from curvature-driven interface evolution. For instance, Fischer and the first author established in [23] a weak-strong uniqueness principle for Navier–Stokes two-phase flow with surface tension based on a suitably constructed relative entropy functional. Their ideas for the treatment of the interfacial energy contribution can (at least in principle) be generalized to multi-phase evolution problems; see our recent work with Fischer and Simon [24] concerning multi-phase mean curvature flow in \mathbb{R}^2 or our recent result [26] for the mean curvature flow of double bubbles in \mathbb{R}^3 . Let us mention that even before these recent developments, Jerrard and Smets [30] already employed the analogue of our relative entropy technique for a problem in codimension 2, namely the evolution of curves in \mathbb{R}^3 by their binormal curvature.

Finally, let us mention that there are other interface evolution problems for which varifold solutions have been considered. Abels proved the existence of varifold solutions to the two-phase Navier–Stokes equation with surface tension [1]. More recently, Chambolle and the second author [10] showed that varifold solutions, which also satisfy the PDE in a distributional sense, exist for the Hele–Shaw or one-phase Mullins–Sekerka equation. Let us stress that in both of these works, the velocity appears in a continuity equation, not a transport equation.

In spirit, our solution concept is related to Serfaty’s solution [50] (which is based on a definition of velocity due to Mugnai and Röger [44]), Ilmanen’s enhanced Brakke flow [29], and the weak solution defined by Bellettini and Mugnai [6]. However, to the best of our knowledge, for these three solution concepts, a weak-strong uniqueness principle is not known.

2. MAIN RESULTS

A key ingredient of the present work is our novel weak solution concept for two-phase mean curvature flow. Roughly speaking, it defines a solution only in terms of the optimal energy-dissipation relation

$$(3) \quad \frac{d}{dt} \mathcal{H}^{d-1}(\partial A(t)) \leq -\frac{1}{2} \int_{\partial A(t)} V^2 d\mathcal{H}^{d-1} - \frac{1}{2} \int_{\partial A(t)} |\mathbf{H}|^2 d\mathcal{H}^{d-1}.$$

The key is that both the normal velocity V as well as the mean curvature vector \mathbf{H} can be defined in a very low regularity setting. This is well-known for the latter via an integration by parts along the closed surface $\partial A(t)$. But for the former, to the best of our knowledge, we give the first definition in this general setting. Let us briefly explain the structure of this definition of V . It is well-known that for a smoothly evolving open set $A(t)$, the volume changes according to

$$\frac{d}{dt} \mathcal{L}^d(A(t)) = - \int_{\partial A(t)} V d\mathcal{H}^{d-1}.$$

(Here and throughout we use the sign convention $V > 0$ for shrinking $A(t)$.) This identity can be localized with a smooth test function, and the resulting PDE in terms of the characteristic function $\chi(x, t) = \chi_{A(t)}(x)$ reads

$$\partial_t \chi + V |\nabla \chi| = 0.$$

This simply means that χ is transported by the vector field Vn , where $n = \frac{\nabla \chi}{|\nabla \chi|}$. Clearly, the term $|\nabla \chi|$ is not stable under weak convergence, which is why we need to relax this definition. When approximating with such smooth sets, we need to take the limit in the sense of varifolds, not currents, which yields

$$\partial_t \chi + V \omega = 0,$$

where ω is the mass measure of the limit varifold μ , and V now is measurable with respect to ω (instead of $|\nabla\chi|$). This is exactly the “transport” equation we will use in our definition.

Finally, let us state (the formal version of) Brakke’s inequality to compare it to our new concept: For any test function $\phi \in C_c^1(\mathbb{R}^d \times [0, \infty))$ with $\phi \geq 0$ Brakke requires

$$\frac{d}{dt} \int_{\partial A(t)} \phi \, d\mathcal{H}^{d-1} \leq \int_{\partial A(t)} (-\phi|\mathbf{H}|^2 + \mathbf{H} \cdot \nabla\phi + \partial_t\phi) \, d\mathcal{H}^{d-1}.$$

Clearly, Brakke’s formulation seems to contain much more information as it requires a whole family of inequalities. In other words, Brakke asks for localized versions of the optimal energy dissipation relation, both in space and time. However, Brakke’s formulation in and of itself cannot exclude the sudden loss of (parts of) the solution. Our solution concept prevents this from happening as the inequality (3) itself provides a certain regularity in time: the phase volume only varies Hölder-continuously in time, see Lemma 1. Even more, we will show in Theorem 2 that, for smooth initial conditions, our weak solution is unique and hence agrees with the classical solution as long as the latter exists, i.e., until the first topological change.

Definition 1 (De Giorgi type varifold solutions for two-phase mean curvature flow). Let $\mu = \mathcal{L}^1 \otimes (\mu_t)_{t \in (0, \infty)}$ be a family of oriented varifolds $\mu_t \in \mathcal{M}(\mathbb{R}^d \times \mathbb{S}^{d-1})$, $t \in (0, \infty)$, such that the map $(0, \infty) \ni t \mapsto \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \eta(\cdot, \cdot, t) \, d\mu_t$ is measurable for all $\eta \in L^1((0, \infty); C_0(\mathbb{R}^d \times \mathbb{S}^{d-1}))$. Consider also a family $A = (A(t))_{t \in (0, \infty)}$ of open subsets of \mathbb{R}^d with finite perimeter in \mathbb{R}^d such that the associated indicator function $\chi(\cdot, t) := \chi_{A(t)}$, $t \in (0, \infty)$, satisfies $\chi \in L^\infty((0, \infty); \text{BV}(\mathbb{R}^d; \{0, 1\}))$. Let $\sigma > 0$ be a surface tension constant.

Given an initial oriented varifold $\mu_0 \in \mathcal{M}(\mathbb{R}^d \times \mathbb{S}^{d-1})$ and an initial phase indicator function $\chi_0 \in \text{BV}(\mathbb{R}^d; \{0, 1\})$, we call the pair (μ, χ) a *De Giorgi type varifold solution for two-phase mean curvature flow with initial data* (μ_0, χ_0) if the following holds.

- (Existence of a normal speed) Writing $\mu_t = \omega_t \otimes (\lambda_{x,t})_{x \in \mathbb{R}^d}$ for the disintegration of μ_t , $t \in (0, \infty)$, we require the existence of some $V \in L^2((0, \infty); L^2(\mathbb{R}^d, d\omega_t))$ encoding a normal speed in the sense of

$$(4a) \quad \sigma \int_{\mathbb{R}^d} \chi(\cdot, T) \zeta(\cdot, T) \, dx - \sigma \int_{\mathbb{R}^d} \chi_0 \zeta(\cdot, 0) \, dx = \sigma \int_0^T \int_{\mathbb{R}^d} \chi \partial_t \zeta \, dx dt - \int_0^T \int_{\mathbb{R}^d} V \zeta \, d\omega_t dt$$

for almost every $T \in (0, \infty)$ and all $\zeta \in C_c^\infty(\mathbb{R}^d \times [0, \infty))$.

- (Existence of a generalized mean curvature vector) We require the existence of some $\mathbf{H} \in L^2((0, \infty); L^2(\mathbb{R}^d, d\omega_t; \mathbb{R}^d))$ encoding a generalized mean curvature vector by

$$(4b) \quad \int_0^\infty \int_{\mathbb{R}^d} \mathbf{H} \cdot B \, d\omega_t dt = - \int_0^\infty \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B \, d\mu_t dt$$

for all $B \in C_c^\infty(\mathbb{R}^d \times [0, \infty); \mathbb{R}^d)$.

- (De Giorgi type inequality for mean curvature flow) A sharp energy dissipation principle à la De Giorgi holds true in form of

$$(4c) \quad \int_{\mathbb{R}^d} 1 \, d\omega_T + \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V|^2 \, d\omega_t dt + \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 \, d\omega_t dt \leq \int_{\mathbb{R}^d} 1 \, d\omega_0$$

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

- (Compatibility) For almost every $t \in (0, \infty)$ and all $\xi \in C_c^\infty(\mathbb{R}^d)$ it holds

$$(4d) \quad \int_{\mathbb{R}^d} \xi \cdot \sigma d\nabla\chi(\cdot, t) = \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \xi \cdot p \, d\mu_t.$$

We observe that due to the compatibility condition (4d) it holds for almost every $t \in (0, \infty)$ that $\sigma|\nabla\chi(\cdot, t)| \leq \omega_t$ in the sense of measures. In particular, for almost every $t \in (0, \infty)$ the

Radon–Nikodým derivative $\rho(\cdot, t) := \frac{\sigma d|\nabla\chi(\cdot, t)|}{d\omega_t}$ exists and satisfies

$$(5) \quad \rho(x, t) \in [0, 1] \quad \text{for } \omega_t \text{ almost every } x \in \mathbb{R}^d,$$

$$(6) \quad \sigma \int_{\partial^* A(t)} f d\mathcal{H}^{d-1} = \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} f \rho(\cdot, t) d\omega_t \quad \text{for all } f \in C_c^\infty(\mathbb{R}^d).$$

In the following first main result, we show that any limit of the Allen–Cahn equation

$$(7) \quad \partial_t u_\varepsilon = \Delta u_\varepsilon - \frac{1}{\varepsilon^2} W'(u_\varepsilon) \quad \text{in } \mathbb{R}^d \times (0, \infty)$$

is a weak solution in the above sense. Here $W: \mathbb{R} \rightarrow [0, \infty)$ is a double-well potential, which for simplicity, we will assume to be of the standard form $W(u) = \frac{1}{4}u^2(u-1)^2$. Recall that the Allen–Cahn equation is the L^2 -gradient flow of the Cahn–Hilliard energy

$$(8) \quad E_\varepsilon(u) := \int_{\mathbb{R}^d} \left(\frac{\varepsilon}{2} |\nabla u|^2 + \frac{1}{\varepsilon} W(u) \right) dx$$

on the slow time scale. This gradient-flow structure can be recognized at the identity

$$\frac{d}{dt} E_\varepsilon(u_\varepsilon(\cdot, t)) = - \int_{\mathbb{R}^d} \varepsilon (\partial_t u_\varepsilon(\cdot, t))^2 dx \leq 0,$$

which follows from a straight-forward computation. In view of our approach here, cf. (3), we should rather measure the energy-dissipation in the balanced way

$$(9) \quad \frac{d}{dt} E_\varepsilon(u_\varepsilon(\cdot, t)) = - \frac{1}{2} \int_{\mathbb{R}^d} \varepsilon (\partial_t u_\varepsilon(\cdot, t))^2 dx - \frac{1}{2} \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \left(\varepsilon \Delta u_\varepsilon(\cdot, t) - \frac{1}{\varepsilon} W'(u_\varepsilon(\cdot, t)) \right)^2 dx,$$

where we have used equation (7) once more.

Theorem 1 (Convergence of the Allen–Cahn equation to De Giorgi type varifold solution). *Let u_ε denote the solution to the Allen–Cahn equation with well-prepared initial conditions $u_{\varepsilon,0}$ in the sense of Ilmanen (40)–(42). Then there exist a measurable function $\chi: \mathbb{R}^d \times (0, \infty) \rightarrow \{0, 1\}$ and a family of oriented varifolds $\mu = \mathcal{L}^1 \otimes (\mu_t)_{t \in (0, \infty)} \in \mathcal{M}(\mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1})$ with associated mass measure $\omega \in \mathcal{M}(\mathbb{R}^d \times (0, \infty))$ such that*

$$(10) \quad \lim_{\varepsilon \downarrow 0} u_\varepsilon = \chi \quad \text{strongly in } L_{loc}^1(\mathbb{R}^d \times (0, \infty)),$$

$$(11) \quad \lim_{\varepsilon \downarrow 0} \left(\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \mathcal{L}^d \otimes \mathcal{L}^1 = \omega \quad \text{weakly-* as Radon measures.}$$

Moreover, the pair (χ, μ) is a solution of mean curvature flow in the sense of Definition 1 with initial conditions (χ_0, μ_0) given by

$$(12) \quad \lim_{\varepsilon \downarrow 0} u_{\varepsilon,0} = \chi_0 \quad \text{strongly in } L_{loc}^1(\mathbb{R}^d),$$

$$(13) \quad \lim_{\varepsilon \downarrow 0} \left(\frac{\varepsilon}{2} |\nabla u_{\varepsilon,0}|^2 + \frac{1}{\varepsilon} W(u_{\varepsilon,0}) \right) \mathcal{L}^d = \omega_0 \quad \text{weakly-* as finite Radon measures.}$$

In addition, if the initial energy measures $(\frac{\varepsilon}{2} |\nabla u_{\varepsilon,0}|^2 + \frac{1}{\varepsilon} W(u_{\varepsilon,0})) \mathcal{L}^d$ have uniformly bounded second moments (e.g. if $u_{\varepsilon,0}$ has compact support), the energy-dissipation inequality (4c) holds also with $(0, T)$ replaced by (s, t) for all $t \in (0, \infty)$ and almost every $s < t$. In particular, the total mass is non-increasing, i.e., $\omega_t(\mathbb{R}^d) \leq \omega_s(\mathbb{R}^d)$ for all $0 < s < t$.

Finally, for a.e. $t \in (0, \infty)$, the corresponding unoriented varifold $\frac{1}{\sigma} \hat{\mu}_t$ (which identifies antipodal points $\pm p$ on \mathbb{S}^{d-1}) is integer rectifiable.

We turn to the question of weak-strong uniqueness for our concept of De Giorgi type varifold solutions for two-phase mean curvature flow. Our strategy is based on lifting the recently developed relative entropy approach for curvature driven interface evolution problems to the present varifold setting by constructing two suitable distance measures between a De Giorgi type varifold solution and a classical solution for two-phase mean curvature flow. To fix notation, let $T_{\text{strong}} \in (0, \infty)$ be a finite time horizon, and let $\mathcal{A} = (\mathcal{A}(t))_{t \in [0, T_{\text{strong}}]}$ be a smoothly evolving

family of bounded domains in \mathbb{R}^d whose associated family of interfaces $\mathcal{I} = (\mathcal{I}(t))_{t \in [0, T_{\text{strong}}]}$ is assumed to evolve smoothly by their mean curvature. Moreover, let (μ, χ) be a De Giorgi type varifold solution in the sense of Definition 1.

The distance measures between \mathcal{A} and (μ, χ) then more precisely consist of a relative entropy

$$(14) \quad E[\mu, \chi | \mathcal{A}](t) := \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} 1 - p \cdot \xi(x, t) d\mu_t(x, p) \geq 0, \quad t \in [0, T_{\text{strong}}],$$

as well as a bulk error

$$(15) \quad E_{\text{bulk}}[\chi | \mathcal{A}](t) := \sigma \int_{\mathbb{R}^d} |\chi_{\mathcal{A}(t)}(x) - \chi_{\mathcal{A}(t)}(x)| |\vartheta(x, t)| dx \geq 0, \quad t \in [0, T_{\text{strong}}].$$

The vector field ξ appearing in the definition of the relative entropy (14) represents a suitable extension of the (inward pointing) normal vector field associated with the classical solution \mathcal{A} , whereas the weight $|\vartheta|$ should be thought of as a smooth version of $(x, t) \mapsto \min\{1, \text{dist}(x, \mathcal{I}(t))\}$. For a rigorous construction of the required data (ξ, ϑ) , we refer the reader to the beginning of Section 4. Let us only mention at this point that the length of the vector field ξ will be subject to a condition of the form $\min\{1, \text{dist}^2(x, \mathcal{I}(t))\} \leq 1 - |\xi(x, t)|$ for all $(x, t) \in \mathbb{R}^d \times [0, T_{\text{strong}}]$. In particular, the definition (14) thus provides an error control in terms of the tilt-excess as well as the distance to the classical solution in form of

$$\int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \frac{1}{2} |p - \xi(\cdot, t)|^2 d\mu_t + \int_{\mathbb{R}^d} \min\{1, \text{dist}^2(\cdot, \mathcal{I}(t))\} d\omega_t \leq 2E[\mu, \chi | \mathcal{A}](t).$$

Finally, we remark that (15) represents nothing more than a version of the well-known Luckhaus–Sturzenhecker [38] or Almgren–Taylor–Wang [3] type distance.

With these ingredients in place, we may now turn to the formulation of the second main result of the present work.

Theorem 2 (Weak-strong uniqueness and quantitative stability for two-phase De Giorgi type varifold solutions). *Let $T_{\text{strong}} \in (0, \infty)$ be a finite time horizon, and let $\mathcal{A} = (\mathcal{A}_t)_{t \in [0, T_{\text{strong}}]}$ be a smoothly evolving family of bounded domains in \mathbb{R}^d whose associated family of interfaces $(\mathcal{I}(t))_{t \in [0, T_{\text{strong}}]}$ evolves smoothly by their mean curvature. Furthermore, let (μ, χ) be a De Giorgi type varifold solution for two-phase mean curvature flow with initial data (μ_0, χ_0) in the sense of Definition 1.*

Defining the relative entropy $E[\mu, \chi | \mathcal{A}]$ and the bulk error $E_{\text{bulk}}[\chi | \mathcal{A}]$ by (14) and (15), respectively, we then have stability estimates for these quantities in form of

$$(16) \quad E[\mu, \chi | \mathcal{A}](T) \leq E[\mu_0, \chi_0 | \mathcal{A}](0) + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt$$

$$(17) \quad E_{\text{bulk}}[\chi | \mathcal{A}](T) \leq E_{\text{bulk}}[\chi_0 | \mathcal{A}](0) + E[\mu_0, \chi_0 | \mathcal{A}](0) + \int_0^T E_{\text{bulk}}[\chi | \mathcal{A}](t) + E[\mu, \chi | \mathcal{A}](t) dt$$

for some constant $C = C(\mathcal{A}) > 0$ and almost every $T \in [0, T_{\text{strong}}]$.

In particular, if $\chi_0 = \chi_{\mathcal{A}(0)}$ and $\mu_0 = \sigma |\nabla \chi_0| \otimes (\delta_{\frac{\nabla \chi_0}{|\nabla \chi_0|}}(x))_{x \in \mathbb{R}^d}$ hold true, then for almost every $t \in [0, T_{\text{strong}}]$

$$(18) \quad \chi(\cdot, t) = \chi_{\mathcal{A}(t)} \text{ almost everywhere in } \mathbb{R}^d,$$

$$(19) \quad \mu_t = \sigma |\nabla \chi(\cdot, t)| \otimes (\delta_{\frac{\nabla \chi(\cdot, t)}{|\nabla \chi(\cdot, t)|}}(x))_{x \in \mathbb{R}^d}.$$

In words, the De Giorgi type varifold solution (μ, χ) reduces to the smooth evolution \mathcal{A} .

In addition to our two main theorems (Theorems 1 and 2 above), one can also derive further properties of our new solution concept from Definition 1. We collect these properties in the following lemma. Its proof is relegated to a separate Section 5.

Lemma 1. *(i) (Classical solutions are weak solutions) If $(\partial \mathcal{A}(t))_{t \in [0, T]}$ is a classical solution to mean curvature flow for some $T \in (0, \infty]$, then the pair consisting of $\chi(x, t) := \chi_{\mathcal{A}(t)}(x)$*

and $\mu_t := |\nabla\chi(\cdot, t)| \otimes \delta_{\frac{\nabla\chi(\cdot, t)}{|\nabla\chi(\cdot, t)|}}$ is a De Giorgi type varifold solution in the sense of Definition 1 (with obvious modifications to Definition 1 to restrict the required conditions to the possibly finite open time interval $(0, T)$).

- (ii) (Regular weak solutions are classical solutions) If (χ, μ) is a De Giorgi type varifold solution such that there exists $T \in (0, \infty]$ so that $\chi(x, t) = \chi_{\mathcal{A}(t)}(x)$ and $\mu_t = |\nabla\chi(\cdot, t)| \otimes \delta_{\frac{\nabla\chi(\cdot, t)}{|\nabla\chi(\cdot, t)|}}$ for all $t \in (0, T)$ and some smoothly evolving family $(\mathcal{A}(t))_{t \in [0, T]}$, then $(\partial\mathcal{A}(t))_{t \in [0, T]}$ is a classical solution to mean curvature flow.
- (iii) (Regularity properties of weak solutions) Any De Giorgi type varifold solution (χ, μ) according to Definition 1 has a certain regularity. First, for a.e. $t \in (0, \infty)$, the corresponding un-oriented varifold $\hat{\mu}_t$ is $(d-1)$ -rectifiable. Second, the volumes change (Hölder-)continuously in time:

$$(20) \quad \sigma \mathcal{L}^d(A(t)\Delta A(s)) \leq \sqrt{2}\omega_0(\mathbb{R}^d)\sqrt{t-s} \quad \text{for all } 0 \leq s < t < \infty.$$

- (iv) (Sequential compactness of solution space) Let (χ_k, μ_k) be a sequence of De Giorgi type varifold solutions such that the energies are non-increasing in time, i.e., for all k , the map $t \mapsto (\omega_k)_t(\mathbb{R}^d)$ is non-increasing. Moreover, we assume that the initial energy $(\omega_k)_0(\mathbb{R}^d)$ is uniformly bounded in $k \in \mathbb{N}$, and that the sequence of $(\omega_k)_0$ is tight. Then there exists a subsequence $(k_\ell)_\ell$ and a De Giorgi type varifold solution (χ, μ) such that $\chi_{k_\ell} \rightarrow \chi$ in L^1_{loc} and $\mu_{k_\ell} \xrightarrow{*} \mu$ as Radon measures.

Let us finally comment on higher multiplicity in our solution concept.

Remark 1. While the De Giorgi type varifold solution allows for arbitrary multiplicity of the varifolds μ_t , some peculiarities may occur in general for De Giorgi varifold solutions when μ_t has higher multiplicity. Let us note that, as long as a classical solution exists, we can control the multiplicity of our solutions by our relative entropy functional (14), see (89) below for the details. However, there exist general De Giorgi varifold solutions, which have higher multiplicity, and in this case it is easy to see that the inequality in the energy-dissipation relation (4c) becomes strict. For example, a self-similarly shrinking sphere with higher multiplicity, e.g. $\chi(\cdot, t) = 0$ and $\omega_t = 2\mathcal{H}^{d-1} \llcorner \partial B_{r(t)}$, or $\chi(\cdot, t) = \chi_{B_{r(t)}}$ and $\omega_t = 3\mathcal{H}^{d-1} \llcorner \partial B_{r(t)}$, where $r(t) = \sqrt{1 - 2(d-1)t^2}$ can be seen to satisfy Definition 1. But here De Giorgi's inequality is strict " $<$ ", because $V = 0$ in the first case (since the phase volume does not change) and $V = -\frac{1}{3}\frac{dr}{dt}$ in the second case, instead of the expected normal velocity $V = -\frac{dr}{dt}$, which would naturally appear in the case of multiplicity 1. Theorem 2 shows that for smooth initial conditions with unit density and for short time, the density of ω_t is 1, which rules out the two examples above. However, it is not clear whether higher multiplicity interfaces could in principle appear after singularities. Since our construction applies in the same framework as Ilmanen's construction of a Brakke flow [28], one could also combine both our De Giorgi inequality and Brakke's inequality. Solutions satisfying both inequalities (in particular any limit of solutions to the Allen-Cahn equation) would then of course enjoy all properties of our and of Brakke's solution.

Structure of the paper. The rest of the paper is structured as follows. In Section 3, we prove Theorem 1. Section 4 is devoted to the proof of Theorem 2. In Section 5, we provide a short proof of Lemma 1. Finally, in Section 6, we propose a generalization of our solution concept to the multi-phase case. Here we also state and prove a weak-strong uniqueness statement analogous to Theorem 2.

3. EXISTENCE OF DE GIORGI TYPE VARIFOLD SOLUTIONS BY ALLEN-CAHN APPROXIMATION

Let u_ε be the solution to (7) with given initial conditions $u_\varepsilon(\cdot, 0) = u_{\varepsilon,0}$. Throughout we will assume the uniform energy bound for the initial conditions

$$(21) \quad E_0 := \sup_{\varepsilon > 0} E_\varepsilon(u_{\varepsilon,0}) < \infty.$$

The basis of our proof of Theorem 1 is the energy-dissipation inequality for u_ε , which states

$$(22) \quad E_\varepsilon(u_\varepsilon(\cdot, T)) + \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \varepsilon (\partial_t u_\varepsilon)^2 dxdt + \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \frac{1}{\varepsilon} \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} W'(u_\varepsilon) \right)^2 dxdt = E_\varepsilon(u_{\varepsilon, 0})$$

for all $T \in (0, \infty)$. This identity (which we will only use as an inequality “ \leq ”) is an immediate consequence of integrating (9) and plugging in (7) to replace half of the dissipation by the right-hand side of the PDE. Furthermore, we will crucially use the well-known quantity

$$(23) \quad \psi_\varepsilon := \phi \circ u_\varepsilon, \quad \text{where } \phi(u) := \int_0^u \sqrt{2W(s)} ds.$$

Recall the Modica–Mortola/Bogomol’nyi-trick [43, 7], i.e., the elegant combination of the chain rule and Young’s inequality $|\nabla \psi_\varepsilon| = \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \leq \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon)$, which implies

$$(24) \quad \sup_{t \in (0, \infty)} \int_{\mathbb{R}^d} |\nabla \psi_\varepsilon| dx \leq E_\varepsilon(u_{\varepsilon, 0}).$$

This motivates to associate to the solution u_ε the oriented space-time varifold

$$(25) \quad \mu_\varepsilon := \left(|\nabla \psi_\varepsilon| \mathcal{L}^d \llcorner \mathbb{R}^d \right) \otimes (\mathcal{L}^1 \llcorner (0, \infty)) \otimes \left(\delta_{\frac{\nabla \psi_\varepsilon}{|\nabla \psi_\varepsilon|} \llcorner \mathbb{S}^{d-1}} \right).$$

In simpler words, μ_ε is the Radon measure on $\mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1}$ such that for any test function $\varphi \in C_c(\mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1})$

$$(26) \quad \int_{\mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1}} \varphi(x, t, p) d\mu_\varepsilon(x, t, p) = \int_{\mathbb{R}^d \times (0, \infty)} \varphi \left(x, t, \frac{\nabla \psi_\varepsilon}{|\nabla \psi_\varepsilon|} \right) |\nabla \psi_\varepsilon(x, t)| dxdt.$$

We denote the associated mass measure as the (x, t) -marginal of μ_ε

$$(27) \quad \omega_\varepsilon := \left(|\nabla \psi_\varepsilon| \mathcal{L}^d \llcorner \mathbb{R}^d \right) \otimes (\mathcal{L}^1 \llcorner (0, \infty)),$$

which simply means that for any test function $\zeta \in C_c(\mathbb{R}^d \times (0, \infty))$

$$(28) \quad \int_{\mathbb{R}^d \times (0, \infty)} \zeta(x, t) d\omega_\varepsilon(x, t) = \int_{\mathbb{R}^d \times (0, \infty)} \zeta(x, t) |\nabla \psi_\varepsilon(x, t)| dxdt.$$

Next, we define the approximate normal velocity V_ε and the approximate mean curvature vector \mathbf{H}_ε to be the ω_ε -measurable functions

$$(29) \quad V_\varepsilon := - \frac{\varepsilon \partial_t u_\varepsilon}{\sqrt{2W(u_\varepsilon)}},$$

$$(30) \quad \mathbf{H}_\varepsilon := - \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} W'(u_\varepsilon) \right) \frac{\nabla \psi_\varepsilon}{|\nabla \psi_\varepsilon|}.$$

On a technical note, in all of the above formulas, we define $\frac{\nabla \psi_\varepsilon}{|\nabla \psi_\varepsilon|} := e_1$ on the set $\{|\nabla \psi_\varepsilon| = 0\}$ and $\frac{\varepsilon \partial_t u_\varepsilon}{\sqrt{2W(u_\varepsilon)}} := 0$ on the set $\{\sqrt{2W(u_\varepsilon)} = 0\} \subset \{|\nabla \psi_\varepsilon| = 0\}$, which anyways are outside the support of the measure ω_ε .

The definitions of μ_ε and \mathbf{H}_ε are very much in the spirit of Ilmanen’s fundamental contribution [28], where the energy density $\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon)$ is used instead of $|\nabla \psi_\varepsilon|$ in the definition of the varifold. We will see that our definition (25) is very handy, but this is only a minor change. The normal velocity V_ε , however, does not appear at all in the work [28]—as the velocity does not appear in Brakke’s formulation. The precise choice (29) will be handy as well, which will become evident in Lemma 3 and Proposition 3.

The following simple compactness statement can be found in the literature.

Proposition 1 (Compactness). *Let the initial conditions $u_\varepsilon(\cdot, 0)$ satisfy the energy bound (21). Then*

$$(31) \quad \sup_{t \in (0, \infty)} \int_{\mathbb{R}^d} |\nabla \psi_\varepsilon| dx \leq E_0,$$

$$(32) \quad \int_0^T \int_{\mathbb{R}^d} |\partial_t \psi_\varepsilon| dx dt \leq (1 + 2T)E_0 \quad \text{for all } T < \infty.$$

Therefore, after passage to a subsequence $\varepsilon \downarrow 0$, there exists a family of sets of finite perimeter $(A(t))_{t \in (0, \infty)}$ such that

$$(33) \quad \psi_\varepsilon \rightarrow \psi \quad \text{in } L^1_{loc}(\mathbb{R}^d \times (0, \infty)),$$

where $\psi(x, t) := \sigma \chi_{A(t)}(x)$ with $\sigma := \phi(1) = \int_0^1 \sqrt{2W(s)} ds$,

$$(34) \quad \operatorname{ess\,sup}_{t \in (0, \infty)} \sigma \mathcal{H}^{d-1}(\partial^* A(t)) \leq E_0$$

and for any $0 \leq s < t$ it holds

$$(35) \quad \sigma \mathcal{L}^d(A(t) \Delta A(s)) \leq E_0 \sqrt{t - s}.$$

Furthermore, there exist Radon measures μ and ω on $\mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1}$ and $\mathbb{R}^d \times (0, \infty)$, respectively, such that

$$(36) \quad \mu_\varepsilon \xrightarrow{*} \mu \quad \text{as Radon measures in } \mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1},$$

$$(37) \quad \omega_\varepsilon \xrightarrow{*} \omega \quad \text{as Radon measures in } \mathbb{R}^d \times (0, \infty).$$

Remark 2. Ilmanen [28, Sections 5.3–5.5] showed even more, namely that

$$(38) \quad (\omega_\varepsilon)_t \xrightarrow{*} \omega_t \quad \text{as finite Radon measures in } \mathbb{R}^d$$

for a.e. $t \in (0, \infty)$. However, this is not needed in the main part of our proof as the disintegration measure ω_t only enters into our solution concept through the *total* mass/energy $\omega_t(\mathbb{R}^d)$. This term is treated in Lemma 2 below. We will however make use of (38) to show the slight improvement of our statement at the end of Theorem 1.

Sketch of the proof of Proposition 1. Although the statements can be found in the literature, let us briefly sketch the simple proofs here.

The estimates (31) and (32) are immediate consequences of the Modica–Mortola/Bogomol’nyi-trick, see the discussion in the paragraph before (24). The same trick applies to the time derivative instead of the spatial gradient. Then the space-time integral of $|\nabla \psi_\varepsilon|$ is controlled using the energy-dissipation estimate in form of

$$\sup_{T \in (0, \infty)} \left\{ E_\varepsilon(u_\varepsilon(\cdot, T)) + \int_{\mathbb{R}^d \times (0, T)} \varepsilon (\partial_t u_\varepsilon)^2 dx dt \right\} \leq E_\varepsilon(u_{\varepsilon, 0}).$$

The Hölder-continuity follows (with a suboptimal constant) from the more general statement [37, Lemma 2.9]. To get the optimal constant in our statement, we estimate by Cauchy–Schwarz

$$\begin{aligned} \int_{\mathbb{R}^d} |\psi_\varepsilon(x, t) - \psi_\varepsilon(x, s)| dx &\leq \int_s^t \int_{\mathbb{R}^d} |\partial_t \psi_\varepsilon| dx dt' \\ &= \int_{\mathbb{R}^d \times (s, t)} \sqrt{2W(u_\varepsilon)} |\partial_t u_\varepsilon| dx dt' \\ &\leq \left(\int_{\mathbb{R}^d \times (0, \infty)} \varepsilon (\partial_t u_\varepsilon)^2 dx dt' \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d \times (s, t)} \frac{2}{\varepsilon} W(u_\varepsilon) dx dt' \right)^{\frac{1}{2}}. \end{aligned}$$

The left-hand side of our inequality converges to the desired $\sigma \mathcal{L}^d(A(t) \Delta A(s))$, at least for a.e. $0 < s < t$. The first right-hand side integral is bounded by E_0 and the second one by $2(t - s)E_0$. This gives the estimate with an additional prefactor $\sqrt{2}$. Treating the second integral with

Proposition 2 below instead, one gets rid of this prefactor. After possibly redefining the limit χ on a negligible set of (positive) times, we obtain (35) for *all* pairs $0 \leq s < t$. \square

The next proposition is at the heart of Ilmanen's fundamental contribution [28]. It states that for well-chosen initial conditions, the discrepancy $\frac{\varepsilon}{2}|\nabla u_\varepsilon(x, t)|^2 - \frac{1}{\varepsilon}W(u_\varepsilon(x, t))$ between the two terms of the energy will remain non-positive for all time. Furthermore, in the limit $\varepsilon \downarrow 0$, this discrepancy vanishes in the sense that

$$(39) \quad \left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2 - \frac{1}{\varepsilon}W(u_\varepsilon) \right) \mathcal{L}^d \otimes \mathcal{L}^1 \xrightarrow{*} 0 \quad \text{as measures.}$$

In other words, asymptotically as $\varepsilon \downarrow 0$, there is equipartition of energy. We also include a simple post-processed version of this statement for later use.

Proposition 2 (Ilmanen [28]). *If the initial conditions $u_{\varepsilon,0}$ are well-prepared in the sense of Ilmanen, i.e.,*

$$(40) \quad \frac{\varepsilon}{2}|\nabla u_{\varepsilon,0}(x)|^2 \leq \frac{1}{\varepsilon}W(u_{\varepsilon,0}(x)) \quad \text{for all } x \in \mathbb{R}^d,$$

$$(41) \quad \sup_{\varepsilon > 0} E_\varepsilon(u_{\varepsilon,0}) < \infty,$$

and denoting $\omega_0 := \text{weak-}^* \lim \left(\frac{\varepsilon}{2}|\nabla u_{\varepsilon,0}|^2 + \frac{1}{\varepsilon}W(u_{\varepsilon,0}) \right) \mathcal{L}^d$, it holds

$$(42) \quad \lim_{\varepsilon \downarrow 0} E_\varepsilon(u_{\varepsilon,0}) = \omega_0(\mathbb{R}^d).$$

Then the analogous statements to (40) and (41) hold for all $t > 0$:

$$(43) \quad \frac{\varepsilon}{2}|\nabla u_\varepsilon(x, t)|^2 \leq \frac{1}{\varepsilon}W(u_\varepsilon(x, t)) \quad \text{for all } x \in \mathbb{R}^d \text{ and all } t \geq 0, \text{ and}$$

$$(44) \quad \sup_{\varepsilon > 0} E_\varepsilon(u_\varepsilon(\cdot, t)) < \infty \quad \text{for all } t \geq 0.$$

Moreover, for all test functions $\zeta \in C_c(\mathbb{R}^d \times (0, \infty))$, it holds

$$(45) \quad \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} \zeta \left(\frac{\varepsilon}{2}|\nabla u_\varepsilon(x, t)|^2 - \frac{1}{\varepsilon}W(u_\varepsilon(x, t)) \right) dx dt = 0$$

and

$$(46) \quad \begin{aligned} & \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} \frac{1}{2} \zeta \left(\sqrt{\varepsilon}|\nabla u_\varepsilon| - \frac{1}{\sqrt{\varepsilon}}\sqrt{2W(u_\varepsilon)} \right)^2 dx dt \\ & = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} \zeta \left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon}W(u_\varepsilon) - |\nabla \psi_\varepsilon| \right) dx dt = 0. \end{aligned}$$

The assertion (43) follows from the maximum principle applied either to the quotient (as is done in [28]) or to the difference of the two terms in the energy. It is the statement (45) which is highly non-trivial. In fact, this is one of the major achievements of the fundamental contribution [28]. Ilmanen's proof is based on a monotonicity formula for the Allen–Cahn equation in analogy to Huisken's celebrated monotonicity formula [27] for the mean curvature flow. We refer to [28] for more detail.

Proof. Since (43), (44), and (45) are derived in Ilmanen's work [28], we will only show how to post-process the statement (45) to the second version (46). We first make the trivial observation

$$\begin{aligned} & \limsup_{\varepsilon \downarrow 0} \left| \int_{\mathbb{R}^d \times (0, \infty)} \zeta \left(\sqrt{\varepsilon}|\nabla u_\varepsilon| - \sqrt{2W(u_\varepsilon)} \right)^2 dx dt \right| \\ & \leq \limsup_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} |\zeta| \left(\sqrt{\varepsilon}|\nabla u_\varepsilon| - \sqrt{2W(u_\varepsilon)} \right)^2 dx dt. \end{aligned}$$

Once we show that the right-hand side is non-positive, we are done. So we may assume $\zeta \geq 0$ and only need to prove the upper bound

$$(47) \quad \limsup_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} \zeta \left(\sqrt{\varepsilon} |\nabla u_\varepsilon| - \sqrt{2W(u_\varepsilon)} \right)^2 dx dt \leq 0.$$

Using (43) in the form of $\sqrt{2W(u_\varepsilon)} \geq \varepsilon |\nabla u_\varepsilon|$ we can estimate pointwise

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

By (45), when tested with ζ , the right-hand side vanishes in the limit $\varepsilon \downarrow 0$. Finally, by the chain rule $|\nabla \psi_\varepsilon| = \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon|$, we see that the first identity in (46) holds even for fixed ε . This concludes the proof of the proposition. \square

The following lemma provides a sharp lower bound for the energy in terms of the mass of the time slices of $\omega_\varepsilon = \lim_\varepsilon \omega_\varepsilon$ constructed in Proposition 1.

Lemma 2. *The measures ω and μ from Proposition 1 can be written as $\omega = (\mathcal{L}^1 \llcorner (0, \infty)) \otimes (\omega_t)_t$ and $\mu = (\mathcal{L}^1 \llcorner (0, \infty)) \otimes (\omega_t)_t \otimes (\lambda_{x,t})_{x,t}$, where $(\omega_t)_t$ is a weakly-* \mathcal{L}^1 -measurable family of finite Radon measures and $(\lambda_{x,t})_{x,t}$ is a weakly-* $\mathcal{L}^1 \otimes (\omega_t)_t$ -measurable family of Radon probability measures. Furthermore, for a.e. $t \in (0, \infty)$ it holds*

$$(48) \quad \liminf_{\varepsilon \downarrow 0} E_\varepsilon(u_\varepsilon(\cdot, t)) \geq \omega_t(\mathbb{R}^d).$$

Proof. For each $\varepsilon > 0$, the real function $t \mapsto e_\varepsilon(t) := E_\varepsilon(u_\varepsilon(\cdot, t))$ is monotonically non-increasing, cf. (9). These functions are also uniformly bounded $0 \leq e_\varepsilon(t) \leq E_0 < \infty$ for all $t \in (0, \infty)$ and all $\varepsilon > 0$. Hence, by Helly's theorem we find a subsequence $\varepsilon \downarrow 0$ and a (monotonically non-increasing) real function $t \mapsto e(t)$ such that

$$(49) \quad e_\varepsilon(t) \rightarrow e(t) \quad \text{for a.e. } t \in (0, \infty).$$

Now we want to use a cylindrical test function. Let $\eta \in C_c((0, \infty))$ with $\eta \geq 0$ and $\zeta \in C_c(\mathbb{R}^d)$ with $\zeta \in [0, 1]$. Using the definition of ω_ε , $\zeta \leq 1$, and the non-negativity (43), we obtain

$$(50) \quad \int_{\mathbb{R}^d \times (0, \infty)} \zeta(x) \eta(t) d\omega_\varepsilon(x, t) \leq \int_0^\infty \eta(t) e_\varepsilon(t) dt.$$

To pass to the limit $\varepsilon \downarrow 0$, we use the compactness (37) for the left-hand side. For the right-hand side, we use (49) and the dominated convergence theorem, which yields

$$(51) \quad \int_{\mathbb{R}^d \times (0, \infty)} \zeta(x) \eta(t) d\omega(x, t) \leq \int_0^\infty \eta(t) e(t) dt.$$

By monotone convergence this can be upgraded to

$$(52) \quad \int_{\mathbb{R}^d \times (0, \infty)} \eta(t) d\omega(x, t) \leq \int_0^\infty \eta(t) e(t) dt.$$

Since the right-hand side is finite for non-negative $\eta \in C_c((0, \infty))$, this means that the projection of ω onto the time variable is a Radon measure. In particular, we can disintegrate ω , i.e., there exists a Radon measure σ on $(0, \infty)$ and a weakly-* σ -measurable family of Radon probability measures $(\tau_t)_t$ on \mathbb{R}^d such that $\omega = \sigma \otimes (\tau_t)_t$. Then (52) means that $\sigma \leq e \mathcal{L}^1 \llcorner (0, \infty)$, so that by the Radon–Nikodým theorem, there exists an \mathcal{L}^1 -measurable function ρ on $(0, \infty)$ with $\rho(t) \leq e(t)$ for a.e. $t \in (0, \infty)$ such that $\sigma = \rho \mathcal{L}^1 \llcorner (0, \infty)$ and hence we obtain the claimed representation for ω with $\omega_t = \rho(t) \tau_t$. Note that also ω_t is weakly-* measurable and, in addition, satisfies the claimed inequality (48). Just another disintegration on the p -variable for μ proves the claimed representation of μ . \square

The following innocent looking but crucial lemma justifies the curious choice (29) for our approximate velocity V_ε . It states that V_ε is uniformly square integrable with respect to our energy measure ω_ε , and that the pair $(\psi_\varepsilon, V_\varepsilon)$ solves an approximate transport equation

$$(53) \quad \partial_t \psi_\varepsilon + V_\varepsilon \omega_\varepsilon \approx 0.$$

We call this a transport equation, since by the definition (27) of ω_ε , this operator simply reads

$$(54) \quad \partial_t \psi_\varepsilon + V_\varepsilon \omega_\varepsilon = \partial_t \psi_\varepsilon + \left(V_\varepsilon \frac{\nabla \psi_\varepsilon}{|\nabla \psi_\varepsilon|} \cdot \nabla \right) \psi_\varepsilon,$$

so that (53) means that ψ_ε is transported by the velocity vector field $V_\varepsilon \frac{\nabla \psi_\varepsilon}{|\nabla \psi_\varepsilon|}$.

Lemma 3. *If the initial conditions $u_{\varepsilon,0}$ are well-prepared in the sense of Ilmanen (40) and (41), then*

$$(55) \quad \int_{\mathbb{R}^d \times (0, \infty)} V_\varepsilon^2 d\omega_\varepsilon \leq E_\varepsilon(u_{\varepsilon,0}).$$

Furthermore, for any test function $\zeta \in C_c(\mathbb{R}^d \times (0, \infty))$

$$(56) \quad \lim_{\varepsilon \downarrow 0} \left(\int_{\mathbb{R}^d \times (0, \infty)} \zeta V_\varepsilon d\omega_\varepsilon + \int_{\mathbb{R}^d \times (0, \infty)} \zeta \partial_t \psi_\varepsilon dxdt \right) = 0.$$

Proof. We first derive the L^2 -estimate (55). By definitions (25) and (29), and the chain rule we have for every $T \in (0, \infty)$

$$\begin{aligned} \int_{\mathbb{R}^d \times (0, T)} V_\varepsilon^2 d\omega_\varepsilon &= \int_{\{W(u_\varepsilon) \neq 0\} \cap (\mathbb{R}^d \times (0, T))} \varepsilon^2 \frac{(\partial_t u_\varepsilon)^2}{2W(u_\varepsilon)} \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| dxdt \\ &= \int_{\{W(u_\varepsilon) \neq 0\} \cap (\mathbb{R}^d \times (0, T))} \varepsilon (\partial_t u_\varepsilon)^2 \frac{\sqrt{\varepsilon} |\nabla u_\varepsilon|}{\frac{1}{\sqrt{\varepsilon}} \sqrt{2W(u_\varepsilon)}} dxdt. \end{aligned}$$

Miraculously enough, by (43) in Ilmanen's Proposition (Proposition 2), the factor $\frac{\sqrt{\varepsilon} |\nabla u_\varepsilon|}{\frac{1}{\sqrt{\varepsilon}} \sqrt{2W(u_\varepsilon)}}$ is less or equal to 1, which shows for every $T \in (0, \infty)$

$$(57) \quad \int_{\mathbb{R}^d \times (0, T)} V_\varepsilon^2 d\omega_\varepsilon \leq \int_{\mathbb{R}^d \times (0, T)} \varepsilon (\partial_t u_\varepsilon)^2 dxdt.$$

Now we conclude the argument for (55) by the energy-dissipation inequality (22) and taking the limit $T \nearrow \infty$.

Next, we want to derive (56). We fix $\zeta \in C_c(\mathbb{R}^d \times (0, \infty))$ and again plug in the definitions (25) and (29), and then apply the chain rule in form of $|\nabla \psi_\varepsilon| = \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon|$, and add zero to obtain

$$\begin{aligned} \int_{\mathbb{R}^d \times (0, \infty)} \zeta V_\varepsilon d\omega_\varepsilon &= - \int_{\{W(u_\varepsilon) \neq 0\}} \zeta \varepsilon \frac{\partial_t u_\varepsilon}{\sqrt{2W(u_\varepsilon)}} |\nabla \psi_\varepsilon| dxdt \\ &= - \int_{\{W(u_\varepsilon) \neq 0\}} \zeta \varepsilon \partial_t u_\varepsilon |\nabla u_\varepsilon| dxdt \\ &= - \int_{\{W(u_\varepsilon) \neq 0\}} \zeta \sqrt{\varepsilon} \partial_t u_\varepsilon \left(\sqrt{\varepsilon} |\nabla u_\varepsilon| - \frac{1}{\sqrt{\varepsilon}} \sqrt{2W(u_\varepsilon)} \right) dxdt \\ &\quad - \int_{\{W(u_\varepsilon) \neq 0\}} \zeta \partial_t u_\varepsilon \sqrt{2W(u_\varepsilon)} dxdt. \end{aligned}$$

Now we may drop the restriction $\{W(u_\varepsilon) \neq 0\}$ in the last integral and recognize the appearing product as a derivative $\partial_t u_\varepsilon \sqrt{2W(u_\varepsilon)} = \partial_t \psi_\varepsilon$. Rearranging terms and applying Cauchy-Schwarz

yields

$$\begin{aligned} & \left| \int_{\mathbb{R}^d \times (0, \infty)} \zeta V_\varepsilon d\omega_\varepsilon + \int_{\mathbb{R}^d \times (0, \infty)} \zeta \partial_t \psi_\varepsilon dxdt \right| \\ & \leq \left(\int_{\mathbb{R}^d \times (0, \infty)} \varepsilon (\partial_t u_\varepsilon)^2 dxdt \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d \times (0, \infty)} \zeta^2 \left(\sqrt{\varepsilon} |\nabla u_\varepsilon| - \frac{1}{\sqrt{\varepsilon}} \sqrt{2W(u_\varepsilon)} \right)^2 dxdt \right)^{\frac{1}{2}}. \end{aligned}$$

By the energy-dissipation inequality (22), the first factor is uniformly bounded in ε , while the second factor vanishes in the limit $\varepsilon \downarrow 0$ thanks to the equipartition of energy (46). This concludes the proof. \square

Let us first state the remaining two ingredients for Theorem 1. The first proposition will finish our analysis of the velocity. In it we construct a velocity V for the limit, which solves the “transport equation”

$$(58) \quad \partial_t \psi + V\omega = 0.$$

Note that this is not quite a transport equation, as we only know that $\omega \geq |\nabla \psi|$ in the sense of measures. Geometrically speaking, the transport equation (58) means that V is the normal velocity of the evolving family of finite perimeter sets $\Omega(t)$ —with the novelty in this work that it might also have a support away from the reduced boundary $\partial^* \Omega(t)$. Furthermore, the proposition provides a sharp inequality between the dissipation term $\int \varepsilon (\partial_t u_\varepsilon)^2 dxdt$ appearing in the energy-dissipation inequality of the Allen–Cahn equation and the corresponding term $\int V^2 d\omega$ in the energy-dissipation inequality for the mean curvature flow (4c), where V is precisely the velocity field in the transport equation (58)

Proposition 3. *For $\varepsilon > 0$, let u_ε be the solution to the Allen–Cahn equation (7) with well-prepared initial conditions in the sense of (43) and (44), let V_ε be defined by (29), and let μ and ω be given by Proposition 1. Then there exists an ω -measurable function $V: \mathbb{R}^d \times (0, \infty) \rightarrow \mathbb{R}$, which is the normal velocity of χ in the precise sense of (4a). Furthermore, for a.e. $T \in (0, \infty)$ we have the following lower semi-continuity-type inequality*

$$(59) \quad \liminf_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} V_\varepsilon^2 d\omega_\varepsilon \geq \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} V^2 d\omega.$$

Moreover, for a.e. $T \in (0, \infty)$ the velocity V provides a sharp lower bound for the dissipation functional

$$(60) \quad \liminf_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \varepsilon (\partial_t u_\varepsilon)^2 dxdt \geq \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} V^2 d\omega.$$

The following second proposition gives us a sharp inequality between the gradient-terms in the energy-dissipation inequalities. More precisely, but still only formally, it gives a sharp lower bound for the term $\int \frac{1}{\varepsilon} (\delta E_\varepsilon)^2 dxdt$ in terms of the corresponding term in the sharp-interface limit $\int (\delta \omega)^2 d\omega$. Although the proof is contained in Ilmanen’s work [28], we reproduce the statement and will, for the reader’s convenience, give a short proof at the end of this section.

Proposition 4 (Ilmanen [28]). *For $\varepsilon > 0$, let u_ε be the solution to the Allen–Cahn equation (7) with well-prepared initial conditions in the sense of Ilmanen (43) and (44), let \mathbf{H}_ε be defined by (30), and let μ and ω be given by Proposition 1. Then there exists a μ -measurable vector field $\mathbf{H}: \mathbb{R}^d \times (0, \infty) \rightarrow \mathbb{R}^d$, which is the mean curvature vector of the oriented space-time varifold μ in the precise sense of (4b). Furthermore, for a.e. $T \in (0, \infty)$ we have the following sharp lower bound*

$$(61) \quad \liminf_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \frac{1}{\varepsilon} |\mathbf{H}_\varepsilon|^2 dxdt \geq \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} |\mathbf{H}|^2 d\omega.$$

Proof of Proposition 3. We start by proving the existence of V . It is enough to show the absolute continuity

$$(62) \quad \sigma \partial_t \chi = \partial_t \psi \ll \omega.$$

Indeed, then we can define V as the Radon–Nikodým derivative $\frac{d(\partial_t \psi)}{d\omega}$. Then automatically, V satisfies (4a) for any $T < \infty$ and any test function $\zeta \in C_0^1(\mathbb{R}^d \times (0, T))$. To now post-process this to the actual statement, we need to approximate in (4a) a given test function $\zeta \in C_c^1(\mathbb{R}^d \times [0, \infty))$ with a sequence $\zeta_n \in C_0^1(\mathbb{R}^d \times (0, T))$. This is standard: take for concreteness $\zeta_n := \eta_n \zeta$ with ζ the piecewise linear cutoff function $\eta(t) = t/n$ for $t \in [0, 1/n]$, $\eta(t) = (T-t)/n$ for $t \in [T-1/n, T]$ and $\eta(t) = 1$ otherwise. Plugging ζ_n into (4a) and applying the product rule yields

$$(63) \quad \sigma n \int_{T-\frac{1}{n}}^T \int_{\mathbb{R}^d} \chi \zeta \, dx dt - \sigma n \int_0^{\frac{1}{n}} \int_{\mathbb{R}^d} \chi \zeta \, dx dt = \sigma \int_0^T \int_{\mathbb{R}^d} \chi \eta_n \partial_t \zeta \, dx dt - \int_0^T \int_{\mathbb{R}^d} V \eta_n \zeta \, d\omega_t dt.$$

By (35), the two average integrals on the left-hand side converge to the desired limits. The terms on the right-hand side converge by absolute continuity.

In order to show (62), assume that $U \subset \mathbb{R}^d \times (0, \infty)$ is open such that

$$(64) \quad \omega(U) = \mu(U \times \mathbb{S}^{d-1}) < \varepsilon.$$

Now we claim that there exists a continuous function $\delta: [0, \infty) \rightarrow [0, \infty)$ with $\delta(0) = 0$ such that

$$(65) \quad |\partial_t \psi|(U) \leq \delta(\varepsilon).$$

Indeed, for any $\zeta \in C_c^1(U)$, using (33), integration by parts, (56), and Cauchy–Schwarz, we may write

$$\begin{aligned} \int_{\mathbb{R}^d \times (0, \infty)} \psi \partial_t \zeta \, dx dt &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} \psi_\varepsilon \partial_t \zeta \, dx dt \\ &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} \zeta V_\varepsilon \, d\omega_\varepsilon \\ &\leq \left(\liminf_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} V_\varepsilon^2 \, d\omega_\varepsilon \right)^{\frac{1}{2}} \left(\lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} \zeta^2 \, d\omega_\varepsilon \right)^{\frac{1}{2}}. \end{aligned}$$

Applying the compactness (55) and (37) (and then using the symmetry $\zeta \mapsto -\zeta$ to get the modulus on the left-hand side)

$$(66) \quad \left| \int_{\mathbb{R}^d \times (0, \infty)} \psi \partial_t \zeta \, dx dt \right| \leq \left(\liminf_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \times (0, \infty)} V_\varepsilon^2 \, d\omega_\varepsilon \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d \times (0, \infty)} \zeta^2 \, d\omega \right)^{\frac{1}{2}}.$$

Taking the supremum over all such ζ , this yields in particular (65) with module $\delta(s) = C\sqrt{s}$ for some constant $C < \infty$ due to (55). Hence, we have proven the absolute continuity (62) and can now define V as the density. Finally, turning back to (66) once more, the Riesz representation theorem yields the L^2 -bound (59).

The sharp lower bound on the dissipation (60) now follows from the already established estimates (59) and (57). \square

Proof of Proposition 4. Although this statement is contained in [28], we give a short argument here for the reader's convenience. To this end, we fix $T \in (0, \infty]$ and $\xi \in C_c(\mathbb{R}^d \times (0, T); \mathbb{R}^d)$.

First, we apply the trivial inequality $\frac{1}{2}|a|^2 \geq a \cdot b - \frac{1}{2}|b|^2$ with $a = \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} W'(u_\varepsilon)$ and $b = \varepsilon \xi \cdot \nabla u_\varepsilon$ to obtain

$$\begin{aligned} \frac{1}{2} \int_{\mathbb{R}^d \times (0, \infty)} \frac{1}{\varepsilon} |\mathbf{H}_\varepsilon|^2 dxdt &= \frac{1}{2} \int_{\mathbb{R}^d \times (0, \infty)} \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} W'(u_\varepsilon) \right)^2 \frac{1}{\varepsilon} dxdt \\ &\geq \int_{\mathbb{R}^d \times (0, \infty)} \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} W'(u_\varepsilon) \right) (\varepsilon \xi \cdot \nabla u_\varepsilon) \frac{1}{\varepsilon} dxdt \\ &\quad - \frac{1}{2} \int_{\mathbb{R}^d \times (0, \infty)} (\varepsilon \xi \cdot \nabla u_\varepsilon)^2 \frac{1}{\varepsilon} dxdt. \end{aligned}$$

We first note that the second right-hand side term can be treated as follows

$$-\frac{1}{2} \int_{\mathbb{R}^d \times (0, \infty)} (\varepsilon \xi \cdot \nabla u_\varepsilon)^2 \frac{1}{\varepsilon} dxdt \geq -\frac{1}{2} \int_{\mathbb{R}^d \times (0, \infty)} |\xi|^2 \varepsilon |\nabla u_\varepsilon|^2 dxdt,$$

which converges to $-\frac{1}{2} \int |\xi|^2 d\omega$ as $\varepsilon \downarrow 0$ by the equipartition (45) and the convergence (37) of ω_ε . For the first right-hand side term, a straight-forward integration by parts yields

$$\int_{\mathbb{R}^d \times (0, \infty)} \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} W'(u_\varepsilon) \right) (\varepsilon \xi \cdot \nabla u_\varepsilon) \frac{1}{\varepsilon} dxdt = \int_{\mathbb{R}^d \times (0, \infty)} \mathbf{T}_\varepsilon : \nabla \xi dxdt,$$

where \mathbf{T}_ε denotes the energy-stress tensor given by

$$\mathbf{T}_\varepsilon = \left(\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) I_d - \varepsilon \nabla u_\varepsilon \otimes \nabla u_\varepsilon.$$

Now clearly

$$\begin{aligned} \int_{\mathbb{R}^d \times (0, \infty)} \mathbf{T}_\varepsilon : \nabla \xi dxdt &= \int_{\mathbb{R}^d \times (0, \infty)} (\nabla \cdot \xi) \left(\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) dxdt \\ &\quad - \int_{\mathbb{R}^d \times (0, \infty)} \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} \cdot \nabla \xi \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} \varepsilon |\nabla u_\varepsilon|^2 dxdt. \end{aligned}$$

Here, the first term converges to the desired limit $\int (\nabla \cdot \xi) d\omega$ thanks to the compactness (37) and the vanishing of the discrepancy (46). Since $\frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} = \frac{\nabla \psi_\varepsilon}{|\nabla \psi_\varepsilon|}$ on the set $\{\nabla \psi_\varepsilon \neq 0\} \supset \{\nabla u_\varepsilon \neq 0\}$, the second term can be written as

$$\begin{aligned} \int_{\mathbb{R}^d \times (0, \infty)} \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} \cdot \nabla \xi \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} \varepsilon |\nabla u_\varepsilon|^2 dxdt &= \int_{\mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1}} p \cdot \nabla \xi p d\mu_\varepsilon(x, t, p) \\ &\quad + \int_{\mathbb{R}^d \times (0, \infty)} \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} \cdot \nabla \xi \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} (\varepsilon |\nabla u_\varepsilon|^2 - |\nabla \psi_\varepsilon|) dxdt. \end{aligned}$$

The first right-hand side term converges to $\int p \cdot \nabla \xi p d\mu$ as $\varepsilon \downarrow 0$ due to the compactness (36), while the second right-hand side term is estimated using Cauchy–Schwarz

$$\begin{aligned} &\int_{\mathbb{R}^d \times (0, \infty)} |\nabla \xi| |\varepsilon |\nabla u_\varepsilon|^2 - |\nabla \psi_\varepsilon| dxdt \\ &= \int_{\mathbb{R}^d \times (0, \infty)} |\nabla \xi| \left| \sqrt{\varepsilon} |\nabla u_\varepsilon| \left| \sqrt{\varepsilon} |\nabla u_\varepsilon| - \sqrt{2W(u_\varepsilon)} \right| \right| dxdt \\ &\leq \left(\int_{\mathbb{R}^d \times (0, \infty)} |\nabla \xi|^2 \varepsilon |\nabla u_\varepsilon|^2 dxdt \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d \times (0, \infty)} |\nabla \xi|^2 \left(\sqrt{\varepsilon} |\nabla u_\varepsilon| - \sqrt{2W(u_\varepsilon)} \right)^2 dxdt \right)^{\frac{1}{2}}, \end{aligned}$$

which vanishes in the limit $\varepsilon \downarrow 0$ thanks to the equipartition of energy (46), the uniform bound on the energy (22) and of course the fact that $\sup |\nabla \xi|$ has compact support.

Therefore, we conclude that

$$\liminf_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d \times (0, \infty)} \frac{1}{\varepsilon} |\mathbf{H}_\varepsilon|^2 dxdt \geq \int_{\mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla \xi d\mu - \frac{1}{2} \int_{\mathbb{R}^d \times (0, \infty)} |\xi|^2 d\omega$$

for all $T \in (0, \infty]$ and all $\xi \in C_c(\mathbb{R}^d \times (0, T); \mathbb{R}^d)$.

Let first $T = \infty$. Since ξ was arbitrary and the left-hand side is finite, by the Riesz representation theorem, there exists a ω -measurable vector field $\mathbf{H}: \mathbb{R}^d \times (0, \infty) \rightarrow \mathbb{R}^d$, which is the mean curvature vector of μ in the precise sense of (4b). Second, for a.e. $T \in (0, \infty)$ since again ξ was arbitrary

$$\liminf_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \frac{1}{\varepsilon} |\mathbf{H}_\varepsilon|^2 dx dt \geq \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} |\mathbf{H}|^2 d\omega.$$

This concludes the proof. \square

We may finally combine the results of this section to provide a proof of the first main result of the present work.

Proof of Theorem 1. Step 1: Convergence to solution. The main part of the theorem now follows directly from the previous propositions. Indeed, the compactness follows from Proposition 1. To check that the limit is a De Giorgi type varifold solution, we need to check the four items in Definition 1. First, by Proposition 3, our function V satisfies the defining equation (4a) of the normal velocity. Proposition 4 guarantees that \mathbf{H} is the generalized mean curvature vector field in the sense of (4b).

To verify the De Giorgi type inequality (4c), we fix $T \in (0, \infty)$ and wish to pass to the limit $\varepsilon \downarrow 0$ in the time-integrated version of (22):

$$(67) \quad E_\varepsilon(u_\varepsilon(\cdot, T)) + \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \varepsilon (\partial_t u_\varepsilon)^2 dx dt + \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \frac{1}{\varepsilon} |\mathbf{H}_\varepsilon|^2 dx dt = E_\varepsilon(u_{\varepsilon, 0}).$$

The right-hand side term converges to the desired limit $\omega_0(\mathbb{R}^d)$ by assumption (42). Now we turn to the left-hand side terms, for which we only need the lower-semi continuity, which was established in Lemma 2, Proposition 3, and Proposition 4, respectively.

The compatibility (4d) follows from passing to the limit $\varepsilon \downarrow 0$ in the (time-integrated) linear relation

$$\int_{\mathbb{R}^d \times (0, \infty)} \xi(x, t) \cdot \nabla \psi_\varepsilon(x, t) dx dt = \int_{\mathbb{R}^d \times \mathbb{S}^{d-1} \times (0, \infty)} \xi(x, t) \cdot p d\mu_\varepsilon(x, t, p)$$

for all $\xi \in C_c^1(\mathbb{R}^d \times [0, \infty))$ and then localizing in time.

Step 2: Additional properties. As stated in the theorem, let us now assume that the second moments of the initial energy measure are bounded, i.e.,

$$(68) \quad \sup_{\varepsilon > 0} \int_{\mathbb{R}^d} |x|^2 \left(\frac{\varepsilon}{2} |\nabla u_{\varepsilon, 0}(x)|^2 + \frac{1}{\varepsilon} W(u_{\varepsilon, 0}(x)) \right) dx < \infty.$$

We want to verify the De Giorgi type inequality (4c) for intervals (s, t) instead of $(0, T)$ for all $t \in (0, \infty)$ and a.e. $s < t$. We first observe that the ε -version (67) holds with $(0, T)$ replaced by (s, t) for all $0 \leq s < t < \infty$. Then we may pass to the limit on the left-hand side just as before (after realizing that the analogous lower semi-continuity statements also hold on the time interval (s, t) for a.e. $s < t$). Now we only need to argue that the right-hand side actually converges: $\lim_{\varepsilon \downarrow 0} E_\varepsilon(u_\varepsilon(\cdot, s)) = \omega_s(\mathbb{R}^d)$. This follows as soon as we know that the energy measures have uniformly bounded second moments, i.e.,

$$(69) \quad \sup_{\varepsilon > 0} \int_{\mathbb{R}^d} |x|^2 \left(\frac{\varepsilon}{2} |\nabla u_\varepsilon(x, t)|^2 + \frac{1}{\varepsilon} W(u_\varepsilon(x, t)) \right) dx < \infty$$

for a.e. $t \in (0, \infty)$. (However, not uniformly in t .) Indeed, (69) implies that each measure $(\omega_\varepsilon)_t$ has uniformly bounded second moments and hence is tight as $\varepsilon \downarrow 0$, so that we may test the convergence (38) with the constant test function 1.

To show (69), we compute, dropping the arguments of u_ε ,

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} (1 + |x|^2) \left(\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) dx &= \int_{\mathbb{R}^d} (1 + |x|^2) (\nabla \cdot (\varepsilon \partial_t u_\varepsilon \nabla u_\varepsilon) - \varepsilon (\partial_t u_\varepsilon)^2) dx \\ &= \int_{\mathbb{R}^d} (2x \cdot \nabla u_\varepsilon) \varepsilon \partial_t u_\varepsilon - (1 + |x|^2) \varepsilon (\partial_t u_\varepsilon)^2 dx \\ &\leq -\varepsilon \int_{\mathbb{R}^d} (\partial_t u_\varepsilon)^2 dx + \int_{\mathbb{R}^d} |x|^2 \varepsilon |\nabla u_\varepsilon|^2 dx \\ &\leq 2 \int_{\mathbb{R}^d} (1 + |x|^2) \left(\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) dx \end{aligned}$$

where we have used Young's inequality and have simply used $-(\partial_t u_\varepsilon)^2 \leq 0$, $|x|^2 \leq (1 + |x|^2)$, and $0 \leq W(u_\varepsilon)$ in the last step. Now (69) follows from Gronwall's inequality, (68), and (41).

The integrality of limits of the Allen–Cahn equation $\frac{1}{\sigma} \hat{\mu}$ was shown by Tonegawa [52, Theorem 2.2]. We note that, in dimensions $d = 2, 3$, one can alternatively apply the work of Röger and Schätzle. Indeed, (4c) implies that $\int |\mathbf{H}(\cdot, t)|^2 d\omega_t < \infty$ for a.e. t . Fixing such a t , we may apply [48, Theorem 5.1] along a subsequence $\varepsilon_n \downarrow 0$, and along a sequence of times $t_n \rightarrow t$. \square

4. WEAK-STRONG UNIQUENESS FOR DE GIORGI TYPE VARIFOLD SOLUTIONS

Let $T_{\text{strong}} \in (0, \infty)$ be a finite time horizon, and let $(\mathcal{A}(t))_{t \in [0, T_{\text{strong}}]}$ be a smoothly evolving family of bounded domains in \mathbb{R}^d whose associated family of interfaces $(\mathcal{I}(t))_{t \in [0, T_{\text{strong}}]}$ is assumed to evolve smoothly by their mean curvature. Given this data, we will first perform some auxiliary constructions before we provide a rigorous definition for the data entering the distance measures (14) and (15).

Denote for each $t \in [0, T_{\text{strong}}]$ by $n_{\mathcal{I}}(t)$ the unit normal vector along $\mathcal{I}(t)$ pointing inside $\mathcal{A}(t)$. By the tubular neighborhood theorem and the assumed smoothness of the data, there exists a small scale $r_c \in (0, 1)$ such that the map

$$(70) \quad \Psi: \left(\bigcup_{t \in [0, T_{\text{strong}}]} \mathcal{I}(t) \times \{t\} \right) \times (-2r_c, 2r_c) \rightarrow \mathbb{R}^d \times [0, T_{\text{strong}}],$$

$$((x, t), s) \mapsto (x + sn_{\mathcal{I}(t)}(x), t)$$

is a smooth diffeomorphism onto its image $\text{im}(\Psi) = \{(x, t) : \text{dist}(x, \mathcal{I}(t)) < 2r_c\}$. The projection of the inverse Ψ^{-1} onto the last component is further denoted by $s_{\mathcal{I}}: \text{im}(\Psi) \rightarrow (-2r_c, 2r_c)$. Note that $s_{\mathcal{I}}$ is a smooth space-time function encoding a signed distance to the evolving family of interfaces $(\mathcal{I}(t))_{t \in [0, T_{\text{strong}}]}$ oriented in form of $\nabla s_{\mathcal{I}}(x, t) = n_{\mathcal{I}(t)}(x)$ for $(x, t) \in \text{im}(\Psi)$. We also define a smooth space-time map $P_{\mathcal{I}}: \text{im}(\Psi) \rightarrow \bigcup_{t \in [0, T_{\text{strong}}]} \mathcal{I}(t)$ by requiring that $P_{\mathcal{I}}(x, t)$ denotes the unique nearest point on the interface $\mathcal{I}(t)$ for each $x \in \mathbb{R}^d$ satisfying $\text{dist}(x, \mathcal{I}(t)) < 2r_c$.

We next fix a smooth cutoff $\kappa: \mathbb{R} \rightarrow [0, 1]$ such that $\kappa \equiv 1$ on $[-\frac{1}{2}, \frac{1}{2}]$ and $\kappa \equiv 0$ on $\mathbb{R} \setminus (-1, 1)$, and then define a quadratic cutoff

$$(71) \quad \zeta(r) := (1 - r^2)\kappa(r^2), \quad r \in \mathbb{R}.$$

These choices allow to define two smooth space-time vector fields

$$(72) \quad \xi: \mathbb{R}^d \times [0, T_{\text{strong}}] \rightarrow \{|x| \leq 1\}, \quad (x, t) \mapsto \zeta\left(\frac{s_{\mathcal{I}}(x, t)}{r_c}\right) n_{\mathcal{I}(t)}(x),$$

$$(73) \quad B: \mathbb{R}^d \times [0, T_{\text{strong}}] \rightarrow \mathbb{R}^d \quad (x, t) \mapsto \zeta\left(\frac{s_{\mathcal{I}}(x, t)}{r_c}\right) (-\Delta s_{\mathcal{I}})(P_{\mathcal{I}}(x, t)) n_{\mathcal{I}(t)}(x).$$

The interpretation to keep in mind is that ξ represents an extension of the unit normal vector field along the smoothly evolving interfaces $(\mathcal{I}(t))_{t \in [0, T_{\text{strong}}]}$, whereas B represents an extension of the associated normal velocity. In particular, expressing the evolution law of the interfaces $(\mathcal{I}(t))_{t \in [0, T_{\text{strong}}]}$ in form of $\partial_t s_{\mathcal{I}} + (B \cdot \nabla) s_{\mathcal{I}} = 0$ on $\text{supp } \xi = \{(x, t) : \text{dist}(x, \mathcal{I}(t)) \leq r_c\}$ it follows

from the definition of the pair (ξ, B) , the properties of the cutoff function ζ , and straightforward arguments that

$$(74) \quad (\partial_t \xi + (B \cdot \nabla) \xi + (\nabla B)^\top \xi)(x, t) = O(\text{dist}(x, \mathcal{I}(t)))$$

$$(75) \quad (\xi \cdot (\partial_t \xi + (B \cdot \nabla) \xi))(x, t) = O(\text{dist}^2(x, \mathcal{I}(t)))$$

$$(76) \quad (B \cdot \xi + \nabla \cdot \xi)(x, t) = O(\text{dist}(x, \mathcal{I}(t)))$$

for all $(x, t) \in \mathbb{R}^d \times [0, T_{\text{strong}}]$ such that $\text{dist}(x, \mathcal{I}(t)) \leq r_c$. Moreover, it obviously holds

$$(77) \quad \min\{1, r_c^{-2} \text{dist}^2(x, \mathcal{I}(t))\} \leq 1 - |\xi(x, t)|$$

for all $(x, t) \in \mathbb{R}^d \times [0, T_{\text{strong}}]$.

Next, we fix a smooth truncation $\bar{\vartheta}: \mathbb{R} \rightarrow [-1, 1]$ satisfying $\bar{\vartheta} \equiv 1$ on $(-\infty, -1]$, $\bar{\vartheta} \equiv -1$ on $[1, \infty)$, $\bar{\vartheta}' < 0$ in $(-1, 1)$, $\bar{\vartheta}(0) = 0$ and $|\bar{\vartheta}(r)| \geq |r|$ for all $r \in [-1, 1]$. With such a map, we then define a smooth weight

$$(78) \quad \vartheta: \mathbb{R}^d \times [0, T_{\text{strong}}] \rightarrow [-1, 1], \quad (x, t) \mapsto \bar{\vartheta}\left(\frac{s_{\mathcal{I}}(x, t)}{r_c}\right).$$

It is straightforward to check that

$$(79) \quad (\partial_t \vartheta + (B \cdot \nabla) \vartheta)(x, t) = 0,$$

$$(80) \quad \min\{1, r_c^{-1} \text{dist}(x, \mathcal{I}(t))\} \leq |\vartheta(x, t)| \leq C \min\{1, r_c^{-1} \text{dist}(x, \mathcal{I}(t))\}$$

for some $C > 0$ and all $(x, t) \in \mathbb{R}^d \times [0, T_{\text{strong}}]$.

With all of this notation and properties in place, we are ready to recall the definition of the two suitable measures quantifying the distance between a De Giorgi type varifold solution and a classical solution for two-phase mean curvature flow. These consist of a relative entropy

$$(81) \quad E[\mu, \chi | \mathcal{A}](t) := \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} 1 - p \cdot \xi(x, t) d\mu_t(x, p) \geq 0, \quad t \in [0, T_{\text{strong}}],$$

as well as a bulk error

$$(82) \quad E_{\text{bulk}}[\chi | \mathcal{A}](t) := \sigma \int_{\mathbb{R}^d} |\chi_{A(t)}(x) - \chi_{\mathcal{A}(t)}(x)| |\vartheta(x, t)| dx \geq 0, \quad t \in [0, T_{\text{strong}}].$$

By the definition and properties of ξ and ϑ , note that it holds

$$(83) \quad \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} |p - \xi(\cdot, t)|^2 d\mu_t \leq 2E[\mu, \chi | \mathcal{A}](t),$$

$$(84) \quad \int_{\mathbb{R}^d} \min\{1, r_c^{-2} \text{dist}^2(\cdot, \mathcal{I}(t))\} d\omega_t \leq E[\mu, \chi | \mathcal{A}](t),$$

as well as

$$(85) \quad \sigma \int_{\mathbb{R}^d} (\chi_{A(t)} - \chi_{\mathcal{A}(t)}) \vartheta(\cdot, t) dx = E_{\text{bulk}}[\chi | \mathcal{A}](t)$$

$$(86) \quad \sigma \int_{\mathbb{R}^d} |\chi_{A(t)} - \chi_{\mathcal{A}(t)}| \min\{1, r_c^{-1} \text{dist}(\cdot, \mathcal{I}(t))\} dx \leq E_{\text{bulk}}[\chi | \mathcal{A}](t).$$

The relative entropy in fact admits several equivalent expressions. For the computation of its time evolution, it is appropriate to rewrite it based on the compatibility condition (4d) and an integration by parts in form of (with the measure-theoretic inward pointing unit normal along $\partial^* A(t)$ defined by $n(\cdot, t) := \frac{\nabla \chi(\cdot, t)}{|\nabla \chi(\cdot, t)|}$, $t \in (0, \infty)$)

$$(87) \quad \begin{aligned} E[\mu, \chi | \mathcal{A}](t) &= \int_{\mathbb{R}^d} 1 d\omega_t - \sigma \int_{\partial^* A(t)} n(\cdot, t) \cdot \xi(\cdot, t) d\mathcal{H}^{d-1} \\ &= \int_{\mathbb{R}^d} 1 d\omega_t + \sigma \int_{\mathbb{R}^d} \chi(\cdot, t) (\nabla \cdot \xi)(\cdot, t) dx. \end{aligned}$$

To identify further coercivity properties, it is useful to add and subtract the Radon–Nikodým derivative ρ and again make use of the compatibility condition (4d) to obtain

$$(88) \quad E[\mu, \chi|\mathcal{A}](t) = \int_{\mathbb{R}^d} 1 - \rho(\cdot, t) d\omega_t + \sigma \int_{\partial^* A(t)} 1 - n(\cdot, t) \cdot \xi(\cdot, t) d\mathcal{H}^{d-1}.$$

In particular, we may deduce from the properties of ρ and ξ that

$$(89) \quad \int_{\mathbb{R}^d} 1 - \rho(\cdot, t) d\omega_t \leq E[\mu, \chi|\mathcal{A}](t),$$

$$(90) \quad \sigma \int_{\partial^* A(t)} |n(\cdot, t) - \xi(\cdot, t)|^2 d\mathcal{H}^{d-1} \leq 2E[\mu, \chi|\mathcal{A}](t),$$

$$(91) \quad \sigma \int_{\partial^* A(t)} \min\{1, r_c^{-2} \text{dist}^2(\cdot, \mathcal{I}(t))\} d\mathcal{H}^{d-1} \leq E[\mu, \chi|\mathcal{A}](t).$$

Now, we have everything in place to proceed with the

Proof of Theorem 2. We split the proof into three steps.

Step 1: Proof of the stability estimate (16). Plugging in the representation (87) at time $t = T$ and $t = 0$, making use of the De Giorgi type inequality (4c) and the evolution equation (4a), and integrating by parts shows

$$(92) \quad \begin{aligned} E[\mu, \chi|\mathcal{A}](T) &\leq E[\mu_0, \chi_0|\mathcal{A}](0) - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V|^2 d\omega_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t dt \\ &\quad - \int_0^T \int_{\mathbb{R}^d} V(\nabla \cdot \xi) d\omega_t dt - \sigma \int_0^T \int_{\partial^* A(t)} n \cdot \partial_t \xi d\mathcal{H}^{d-1} dt. \end{aligned}$$

Having in mind the approximate transport equations (74) and (75) as well as the coercivity properties (90) and (91), we add zero twice to obtain the estimate

$$(93) \quad \begin{aligned} & - \sigma \int_0^T \int_{\partial^* A(t)} n \cdot \partial_t \xi d\mathcal{H}^{d-1} dt \\ &= -\sigma \int_0^T \int_{\partial^* A(t)} n \cdot (\partial_t \xi + (B \cdot \nabla) \xi + (\nabla B)^\top \xi) d\mathcal{H}^{d-1} dt \\ &\quad + \sigma \int_0^T \int_{\partial^* A(t)} n \cdot ((B \cdot \nabla) \xi + (\nabla B)^\top \xi) d\mathcal{H}^{d-1} dt \\ &= -\sigma \int_0^T \int_{\partial^* A(t)} (n - \xi) \cdot (\partial_t \xi + (B \cdot \nabla) \xi + (\nabla B)^\top \xi) d\mathcal{H}^{d-1} dt \\ &\quad - \sigma \int_0^T \int_{\partial^* A(t)} \xi \cdot (\partial_t \xi + (B \cdot \nabla) \xi) d\mathcal{H}^{d-1} dt + \sigma \int_0^T \int_{\partial^* A(t)} \xi \otimes (n - \xi) : \nabla B d\mathcal{H}^{d-1} dt \\ &\quad + \sigma \int_0^T \int_{\partial^* A(t)} n \cdot (B \cdot \nabla) \xi d\mathcal{H}^{d-1} dt \\ &\leq \sigma \int_0^T \int_{\partial^* A(t)} \xi \otimes (n - \xi) : \nabla B d\mathcal{H}^{d-1} dt + \sigma \int_0^T \int_{\partial^* A(t)} n \cdot (B \cdot \nabla) \xi d\mathcal{H}^{d-1} dt \\ &\quad + C \int_0^T E[\mu, \chi|\mathcal{A}](t) dt. \end{aligned}$$

Based on the compatibility condition (4d), the identity (6), the coercivity property (89), and adding zero we also have

$$\begin{aligned}
& \sigma \int_0^T \int_{\partial^* A(t)} \xi \otimes (n - \xi) : \nabla B \, d\mathcal{H}^{d-1} dt \\
&= \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \xi \otimes (p - \xi) : \nabla B \, d\mu_t dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (\rho - 1)(\xi \otimes \xi : \nabla B) \, d\omega_t dt \\
(94) \quad &\leq \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \xi \otimes (p - \xi) : \nabla B \, d\mu_t dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) \, dt.
\end{aligned}$$

Next, we obtain by appealing to the product rule in form of $n \cdot (B \cdot \nabla) \xi = n \cdot (\nabla \cdot (\xi \otimes B)) - (n \cdot \xi)(\nabla \cdot B)$, adding zero three times, applying again the identity (6) as well as the coercivity property (89), and finally using the representation (88)

$$\begin{aligned}
& \sigma \int_0^T \int_{\partial^* A(t)} n \cdot (B \cdot \nabla) \xi \, d\mathcal{H}^{d-1} dt \\
&= \sigma \int_0^T \int_{\partial^* A(t)} n \cdot (\nabla \cdot (\xi \otimes B)) \, d\mathcal{H}^{d-1} dt - \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot \xi - 1)(\nabla \cdot B) \, d\mathcal{H}^{d-1} dt \\
&\quad - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (\rho - 1)(\nabla \cdot B) \, d\omega_t dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B \, d\mu_t dt \\
&\quad - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \otimes p : \nabla B \, d\mu_t dt \\
(95) \quad &\leq \sigma \int_0^T \int_{\partial^* A(t)} n \cdot (\nabla \cdot (\xi \otimes B)) \, d\mathcal{H}^{d-1} dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B \, d\mu_t dt \\
&\quad - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \otimes p : \nabla B \, d\mu_t dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) \, dt.
\end{aligned}$$

We further compute due to an integration by parts, the symmetry relation $\nabla \cdot (\nabla \cdot (\xi \otimes B)) = \nabla \cdot (\nabla \cdot (B \otimes \xi))$, reverting the integration by parts, the product rule in form of $n \cdot (\nabla \cdot (B \otimes \xi)) = (n \cdot B)(\nabla \cdot \xi) + n \otimes \xi : \nabla B$, and the compatibility condition (4d)

$$\begin{aligned}
& \sigma \int_0^T \int_{\partial^* A(t)} n \cdot (\nabla \cdot (\xi \otimes B)) \, d\mathcal{H}^{d-1} dt \\
&= \sigma \int_0^T \int_{\partial^* A(t)} n \cdot (\nabla \cdot (B \otimes \xi)) \, d\mathcal{H}^{d-1} dt \\
&= \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot B)(\nabla \cdot \xi) \, d\mathcal{H}^{d-1} dt + \sigma \int_0^T \int_{\partial^* A(t)} n \otimes \xi : \nabla B \, d\mathcal{H}^{d-1} dt \\
(96) \quad &= \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot B)(\nabla \cdot \xi) \, d\mathcal{H}^{d-1} dt + \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \otimes \xi : \nabla B \, d\mu_t dt.
\end{aligned}$$

The combination of the estimates (93)–(96) thus implies

$$\begin{aligned}
& - \sigma \int_0^T \int_{\partial^* A(t)} n \cdot \partial_t \xi \, d\mathcal{H}^{d-1} dt \\
&\leq \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot B)(\nabla \cdot \xi) \, d\mathcal{H}^{d-1} dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B \, d\mu_t dt \\
&\quad - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p - \xi) \otimes (p - \xi) : \nabla B \, d\mu_t dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) \, dt.
\end{aligned}$$

Based on the coercivity property (83), the previous display in turn upgrades (92) to

$$(97) \quad \begin{aligned} E[\mu, \chi | \mathcal{A}](T) &\leq E[\mu_0, \chi_0 | \mathcal{A}](0) - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V|^2 d\omega_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t dt \\ &\quad - \int_0^T \int_{\mathbb{R}^d} V(\nabla \cdot \xi) d\omega_t dt + \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot B)(\nabla \cdot \xi) d\mathcal{H}^{d-1} dt \\ &\quad - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B d\mu_t dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt. \end{aligned}$$

What remains to be done in the derivation of the stability estimate (16) is to make use of the control provided by the dissipation terms from the De Giorgi inequality (4c). We start by completing squares twice, applying the estimate (76) as well as the identity (6), and exploiting the coercivity properties (84) and (89)

$$(98) \quad \begin{aligned} & - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V|^2 d\omega_t dt - \int_0^T \int_{\mathbb{R}^d} V(\nabla \cdot \xi) d\omega_t dt + \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot B)(\nabla \cdot \xi) d\mathcal{H}^{d-1} dt \\ &= - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V + (\nabla \cdot \xi)|^2 d\omega_t dt + \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} (\nabla \cdot \xi)^2 d\omega_t dt \\ &\quad + \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot B)(\nabla \cdot \xi) d\mathcal{H}^{d-1} dt \\ &= - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V + (\nabla \cdot \xi)|^2 d\omega_t dt + \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\nabla \cdot \xi + (B \cdot \xi)|^2 d\omega_t dt \\ &\quad - \int_0^T \int_{\mathbb{R}^d} (1 - \rho)(B \cdot \xi)(\nabla \cdot \xi) d\omega_t dt \\ &\quad + \sigma \int_0^T \int_{\partial^* A(t)} ((n - \xi) \cdot B)(\nabla \cdot \xi) d\mathcal{H}^{d-1} dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} (B \cdot \xi)^2 d\omega_t dt \\ &\leq - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V + (\nabla \cdot \xi)|^2 d\omega_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} (B \cdot \xi)^2 d\omega_t dt \\ &\quad + \sigma \int_0^T \int_{\partial^* A(t)} ((n - \xi) \cdot B)(\nabla \cdot \xi) d\mathcal{H}^{d-1} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt. \end{aligned}$$

Furthermore, decomposing and completing squares also shows (recalling from definition (72) that $\xi = |\xi| \nabla s_{\mathcal{I}}$ and $\text{supp } \xi(\cdot, t) = \{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}$)

$$(99) \quad \begin{aligned} & - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} (B \cdot \xi)^2 d\omega_t dt \\ &= - \frac{1}{2} \int_0^T \int_{\{\text{dist}(\cdot, \mathcal{I}(t)) > r_c\}} |\mathbf{H}|^2 d\omega_t dt - \frac{1}{2} \int_0^T \int_{\{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}} |(I_d - \nabla s_{\mathcal{I}} \otimes \nabla s_{\mathcal{I}}) \mathbf{H}|^2 d\omega_t dt \\ &\quad - \frac{1}{2} \int_0^T \int_{\{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}} |\nabla s_{\mathcal{I}} \cdot \mathbf{H}|^2 (1 - |\xi|^2) d\omega_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |(\xi \cdot \mathbf{H}) - (B \cdot \xi)|^2 d\omega_t dt \\ &\quad - \int_0^T \int_{\mathbb{R}^d} (\xi \cdot \mathbf{H})(B \cdot \xi) d\omega_t dt. \end{aligned}$$

As a consequence of plugging in test functions of the form ηB into the identity (4b), where $\eta \in C_c^\infty([0, \infty); [0, 1])$ runs through a sequence approximating the indicator function of $(0, T)$

monotonically from above, and a splitting similar to the previous display, it also holds

$$\begin{aligned}
(100) \quad & - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B \, d\mu_t dt \\
& = \frac{1}{2} \int_0^T \int_{\{\text{dist}(\cdot, \mathcal{I}(t)) > r_c\}} \mathbf{H} \cdot B \, d\omega_t dt + \int_0^T \int_{\mathbb{R}^d} (\xi \cdot \mathbf{H})(B \cdot \xi) \, d\omega_t dt \\
& \quad + \int_0^T \int_{\{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}} (1 - |\xi|^2)(\nabla_{s_{\mathcal{I}}} \cdot \mathbf{H})(B \cdot \nabla_{s_{\mathcal{I}}}) \, d\omega_t dt \\
& \quad + \int_0^T \int_{\{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}} (I_d - \nabla_{s_{\mathcal{I}}} \otimes \nabla_{s_{\mathcal{I}}}) \mathbf{H} \cdot (I_d - \nabla_{s_{\mathcal{I}}} \otimes \nabla_{s_{\mathcal{I}}}) B \, d\omega_t dt.
\end{aligned}$$

Summing (99) and (100), by an absorption argument, the bounds $1 - |\xi|^2 \leq 2(1 - |\xi|) \leq 2(1 - n \cdot \xi)$, the fact $\text{supp } \xi(\cdot, t) = \{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}$, and the representation (88), we obtain

$$\begin{aligned}
(101) \quad & - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 \, d\omega_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} (B \cdot \xi)^2 \, d\omega_t dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B \, d\mu_t dt \\
& \leq C \int_0^T \int_{\{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}} |(I_d - \nabla_{s_{\mathcal{I}}} \otimes \nabla_{s_{\mathcal{I}}}) B|^2 \, d\omega_t dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) \, dt.
\end{aligned}$$

Due to $|(n - \xi)\xi| \leq (1 - |\xi|^2) + (1 - n \cdot \xi) \leq 3(1 - n \cdot \xi)$ we also get

$$\begin{aligned}
(102) \quad & \sigma \int_0^T \int_{\partial^* A(t)} ((n - \xi) \cdot B)(\nabla \cdot \xi) \, d\mathcal{H}^{d-1} dt \\
& \leq \sigma \int_0^T \int_{\partial^* A(t)} ((n - \xi) \cdot (I_d - \xi \otimes \xi) B)(\nabla \cdot \xi) \, d\mathcal{H}^{d-1} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) \, dt.
\end{aligned}$$

In summary, the estimates (97), (98), (101), and (102) together with $(I_d - \nabla_{s_{\mathcal{I}}} \otimes \nabla_{s_{\mathcal{I}}}) B = 0$ (recall the definition (73) of B) and $|\nabla_{s_{\mathcal{I}}} \otimes \nabla_{s_{\mathcal{I}}} - \xi \otimes \xi| \leq 1 - |\xi|^2$ throughout $\{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}$ imply

$$\begin{aligned}
(103) \quad & E[\mu, \chi | \mathcal{A}](T) \leq E[\mu_0, \chi_0 | \mathcal{A}(0)] - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V + (\nabla \cdot \xi)|^2 \, d\omega_t dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) \, dt \\
& \quad - \sigma \int_0^T \int_{\partial^* A(t)} ((n - \xi) \cdot (I_d - \xi \otimes \xi) B)(\nabla \cdot \xi) \, d\mathcal{H}^{d-1} dt \\
& \quad + C \int_0^T \int_{\{\text{dist}(\cdot, \mathcal{I}(t)) \leq r_c\}} |(I_d - \nabla_{s_{\mathcal{I}}} \otimes \nabla_{s_{\mathcal{I}}}) B|^2 \, d\omega_t dt \\
& \leq E[\mu_0, \chi_0 | \mathcal{A}(0)] - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |V + (\nabla \cdot \xi)|^2 \, d\omega_t dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) \, dt.
\end{aligned}$$

Of course, the previous display entails the asserted estimate. (Note that only the very last step of this argument relied on the fact that B always points in normal direction, which—at least in principle—is the only point exploiting crucially the simple two-phase setting.)

Step 2: Proof of the stability estimate (17). Plugging in the alternative characterization (82) of the bulk error term at time $t = T$ and $t = 0$, exploiting the evolution equation (4a), and making use of $\vartheta(\cdot, t) = 0$ on $\mathcal{I}(t)$ due to (80), we first get

$$(104) \quad E_{\text{bulk}}[\chi | \mathcal{A}](T) = E_{\text{bulk}}[\chi_0 | \mathcal{A}(0)] - \int_0^T \int_{\mathbb{R}^d} V \vartheta \, d\omega_t dt + \sigma \int_0^T \int_{\mathbb{R}^d} (\chi_{A(t)} - \chi_{\mathcal{A}(t)}) \partial_t \vartheta \, dx dt.$$

Inserting the evolution equation (79), appealing to the product rule in form of $(B \cdot \nabla) \vartheta = \nabla \cdot (\vartheta B) - \vartheta (\nabla \cdot B)$, recalling the definition (82), and finally integrating by parts and using $\rho = 0$

on $\partial\mathcal{A}(t)$ shows

$$\begin{aligned}
\sigma \int_0^T \int_{\mathbb{R}^d} (\chi_{A(t)} - \chi_{\mathcal{A}(t)}) \partial_t \vartheta \, dx dt &= -\sigma \int_0^T \int_{\mathbb{R}^d} (\chi_{A(t)} - \chi_{\mathcal{A}(t)}) \nabla \cdot (\vartheta B) \, dx dt \\
&\quad + \sigma \int_0^T \int_{\mathbb{R}^d} (\chi_{A(t)} - \chi_{\mathcal{A}(t)}) \vartheta (\nabla \cdot B) \, dx dt \\
(105) \qquad \qquad \qquad &\leq \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot B) \vartheta \, d\mathcal{H}^{d-1} dt + C \int_0^T E_{\text{bulk}}[\chi|\mathcal{A}](t) \, dt.
\end{aligned}$$

Next, by adding zero three times as well as making use of the identity (6), Young's inequality and the coercivity properties (80), (82), (84) and (89)–(91) we may further estimate

$$\begin{aligned}
& - \int_0^T V \vartheta \, d\omega_t dt + \sigma \int_0^T \int_{\partial^* A(t)} (n \cdot B) \vartheta \, d\mathcal{H}^{d-1} dt \\
&= - \int_0^T V \vartheta \, d\omega_t dt + \sigma \int_0^T \int_{\partial^* A(t)} ((n - \xi) \cdot B) \vartheta \, d\mathcal{H}^{d-1} dt \\
&\quad + \sigma \int_0^T \int_{\partial^* A(t)} ((\xi \cdot B) + (\nabla \cdot \xi)) \vartheta \, d\mathcal{H}^{d-1} dt \\
&\quad - \int_0^T \int_{\partial^* A(t)} (\rho - 1) (\nabla \cdot \xi) \vartheta \, d\omega_t dt - \int_0^T \int_{\partial^* A(t)} (\nabla \cdot \xi) \vartheta \, d\omega_t dt \\
(106) \qquad \qquad \qquad &\leq \frac{1}{4} \int_0^T \int_{\mathbb{R}^d} |(\nabla \cdot \xi) + V|^2 \, d\omega_t dt + C \int_0^T E[\mu, \chi|\mathcal{A}](t) \, dt.
\end{aligned}$$

Hence, the bounds (104)–(106) in total entail

$$\begin{aligned}
(107) \qquad E_{\text{bulk}}[\chi|\mathcal{A}](T) &\leq E_{\text{bulk}}[\chi_0|\mathcal{A}(0)] + \frac{1}{4} \int_0^T \int_{\mathbb{R}^d} |(\nabla \cdot \xi) + V|^2 \, d\omega_t dt \\
&\quad + C \int_0^T E_{\text{bulk}}[\chi|\mathcal{A}](t) + E[\mu, \chi|\mathcal{A}](t) \, dt.
\end{aligned}$$

We thus infer (17) from adding (107) to (103).

Step 3: Proof of the weak-strong uniqueness principle (18)–(19). This now follows immediately from the previous two steps. \square

5. PROOF OF LEMMA 1

Argument for Item (i): Only the validity of the De Giorgi type inequality (4c) deserves a comment, while the other requirements are immediate consequences of the smoothness of the underlying evolution and the definition of (μ, χ) . With respect to (4c), we note that it is sufficient to establish for a.e. $T \in (0, \infty)$

$$\sigma \mathcal{H}^{d-1}(\partial\mathcal{A}(T)) + \sigma \int_0^T \int_{\mathbb{R}^d} V n \cdot \mathbf{H} \, d\mathcal{H}^{d-1} dt \leq \sigma \mathcal{H}^{d-1}(\partial\mathcal{A}(0)).$$

Indeed, since we know by assumption that $V = n \cdot \mathbf{H}$ throughout $\mathbb{R}^d \times (0, \infty)$, the De Giorgi type inequality (4c) is easily seen to be equivalent to the previous display. On the other side, in the setting considered here the inequality of the previous display holds true even as an equality which in turn is a well-known property for smooth evolutions.

Argument for Item (ii): Since $(\partial\mathcal{A}(t))_{t \in [0, T]}$ evolves smoothly, we have again for all $T' \in (0, T)$

$$\sigma \mathcal{H}^{d-1}(\partial\mathcal{A}(T')) + \sigma \int_0^{T'} \int_{\mathbb{R}^d} V n \cdot \mathbf{H} \, d\mathcal{H}^{d-1} dt = \sigma \mathcal{H}^{d-1}(\partial\mathcal{A}(0)).$$

On the other hand, by virtue of the assumed representation of the given De Giorgi type varifold solution, (4c) and completing squares, we also have for all $T' \in (0, T)$

$$\begin{aligned} & \sigma \mathcal{H}^{d-1}(\partial \mathcal{A}(T')) + \sigma \int_0^{T'} \int_{\mathbb{R}^d} \frac{1}{2} (V - n \cdot \mathbf{H})^2 d\mathcal{H}^{d-1} dt + \sigma \int_0^{T'} \int_{\mathbb{R}^d} V n \cdot \mathbf{H} d\mathcal{H}^{d-1} dt \\ & \leq \sigma \mathcal{H}^{d-1}(\partial \mathcal{A}(0)). \end{aligned}$$

Subtracting the former display from the latter, we obtain as desired that $V = n \cdot \mathbf{H}$ holds true throughout $\mathbb{R}^d \times (0, T)$.

Argument for Item (iii): The rectifiability of μ_t for a.e. t follows from the fact that (4c) implies $\int |\mathbf{H}(\cdot, t)|^2 d\omega_t < \infty$ for a.e. t and Allard's rectifiability criterion [2], see also [45] for an alternative proof.

Now we show (20). Taking the difference of (4a) with $T = t$ and $T = s$ for cylindrical test functions $(x, t) \mapsto \eta(t)\zeta(x)$, where $\zeta \in C_c^\infty(\mathbb{R}^d; [-1, 1])$ and $\eta \in C_c^\infty([0, \infty); [0, 1])$ such that $\eta \equiv 1$ on $[s, t]$, and applying Cauchy–Schwarz yields

$$\begin{aligned} \sigma \int_{A(t) \Delta A(s)} \zeta(x) dx &= - \int_s^t \int_{\mathbb{R}^d} V(x, t) \zeta(x) d\omega_t dt \\ &\leq \left(\int_{\mathbb{R}^d \times (0, \infty)} V^2 d\omega \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d \times (s, t)} 1 d\omega \right)^{\frac{1}{2}}. \end{aligned}$$

The first right-hand side integral is bounded by $2\omega_0(\mathbb{R}^d)$ while the second one is bounded by $(t - s)\omega_0(\mathbb{R}^d)$; both due to (4c). Taking the supremum over ζ thus entails (20).

Argument for Item (iv): We again only focus on establishing (4c); the remaining conditions in fact essentially follow in the process.

Define $\omega_k = (\mathcal{L}^1 \llcorner (0, \infty)) \otimes ((\omega_k)_t)_{t \in (0, \infty)}$. Up to choosing a suitable subsequence, the sequence $\omega_k \in \mathcal{M}(\mathbb{R}^d \times (0, \infty))$ converges weakly-* to some $\omega \in \mathcal{M}(\mathbb{R}^d \times (0, \infty))$ in the sense of Radon measures. Recall that by assumption, for all k , the map $t \mapsto (\omega_k)_t(\mathbb{R}^d)$ is non-increasing, and that the initial energy $(\omega_k)_0(\mathbb{R}^d)$ is uniformly bounded in $k \in \mathbb{N}$. Hence, a close inspection of the proof of Lemma 2 shows that $\omega = (\mathcal{L}^1 \llcorner (0, \infty)) \otimes (\omega_t)_{t \in (0, \infty)}$, where $(\omega_t)_{t \in (0, \infty)}$ is a weakly-* \mathcal{L}^1 -measurable family of finite Radon measures, and for a.e. $t \in (0, \infty)$ it moreover holds

$$(108) \quad \liminf_{k \downarrow 0} (\omega_k)_t(\mathbb{R}^d) \geq \omega_t(\mathbb{R}^d).$$

As in the proof of Proposition 3, there exists $V \in L^2((0, \infty); L^2(\mathbb{R}^d, d\omega_t))$ arising as the Radon–Nikodým derivative of $\partial_t \chi$ with respect to ω (in particular, the equation (4a) follows) such that in addition for a.e. $T \in (0, \infty)$

$$(109) \quad \liminf_{k \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} V_k^2 d\omega_k \geq \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} V^2 d\omega.$$

By a straightforward argument solely based on the Riesz representation theorem and the sub-sequential weak-* convergence of μ_k to some μ (with ω being the associated mass measure) in the sense of Radon measures in $\mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-1}$, there exists $\mathbf{H} \in L^2((0, \infty); L^2(\mathbb{R}^d, d\omega_t; \mathbb{R}^d))$ such that (4b) holds true and for a.e. $T \in (0, \infty)$, we have

$$(110) \quad \liminf_{k \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \mathbf{H}_k^2 d\omega_k \geq \frac{1}{2} \int_{\mathbb{R}^d \times (0, T)} \mathbf{H}^2 d\omega.$$

All in all, in view of the previous three displays together with the validity of (4c) in terms of the sequence (μ_k, χ_k) , we may deduce for a.e. $T \in (0, \infty)$ that (4c) holds true in terms of (μ, χ) as desired. \square

6. A PROPOSAL IN THE SETTING OF MULTI-PHASE MEAN CURVATURE FLOW

In this section, we propose a varifold solution concept à la De Giorgi for multi-phase mean curvature flow. We only consider the case of equal surface tensions (which we consequently normalize to the unit) meaning that at triple junctions the interfaces intersect at equal angles of 120° . When reduced to the two-phase setting, our proposal turns out to be slightly stronger than what is required in Definition 1 in the sense that, for almost every $t \in (0, T)$ and ω_t almost every $x \in \mathbb{R}^d$, the generalized mean curvature vector $\mathbf{H}(x, t)$ of (4b) points in the direction of the expected value $\langle \lambda_{x,t} \rangle := \int_{\mathbb{S}^{d-1}} p \, d\lambda_{x,t}$ (if the latter does not vanish; see (111i) below). The merit of this additional assumption is that it allows to close the Gronwall argument from the previous section without requiring the velocity B to point in normal direction (cf. the proof of Theorem 3 below). This in turn is of course essential for the multi-phase case since B will necessarily contain tangential components due to the presence of moving triple junctions. Note that this additional assumption on \mathbf{H} is a direct consequence of Brakke's orthogonality theorem [8, Chapter 5] in the two-phase case since we have shown in Theorem 1 that the varifolds $\frac{1}{\sigma}\mu_t$ in our construction have integer multiplicity for a.e. t .

Definition 2 (De Giorgi type varifold solutions for multi-phase mean curvature flow). Let $P \geq 2$ be the number of phases. For each pair of phases $i, j \in \{1, \dots, P\}$, let $\mu^{i,j} = \mathcal{L}^1 \otimes (\mu_t^{i,j})_{t \in (0, \infty)}$ be a family of oriented varifolds $\mu_t^{i,j} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{S}^{d-1})$, $t \in (0, \infty)$, such that the map $(0, \infty) \ni t \mapsto \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \eta(\cdot, \cdot, t) \, d\mu_t^{i,j}$ is measurable for all test functions $\eta \in L^1((0, \infty); C_0(\mathbb{R}^d \times \mathbb{S}^{d-1}))$. Define evolving oriented varifolds $\mu^i = \mathcal{L}^1 \otimes (\mu_t^i)_{t \in (0, \infty)}$, $i \in \{1, \dots, P\}$, and $\mu = \mathcal{L}^1 \otimes (\mu_t)_{t \in (0, \infty)}$ by means of

$$(111a) \quad \mu_t^i := 2\mu_t^{i,i} + \sum_{j=1, j \neq i}^P \mu_t^{j,i}, \quad t \in (0, \infty),$$

$$(111b) \quad \mu_t := \frac{1}{2} \sum_{i=1}^P \mu_t^i, \quad t \in (0, \infty).$$

The disintegration of $\mu_t^{i,j}$ is expressed in form of $\mu_t^{i,j} = \omega_t^{i,j} \otimes (\lambda_{x,t}^{i,j})_{x \in \mathbb{R}^d}$ with expected value $\langle \lambda_{x,t}^{i,j} \rangle := \int_{\mathbb{S}^{d-1}} p \, d\lambda_{x,t}^{i,j}(p)$ for all $i, j \in \{1, \dots, P\}$ and all $t \in (0, \infty)$. Analogous expressions are introduced for the disintegrations of μ_t^i and μ_t , respectively.

Consider also a family $A = (A^1, \dots, A^P)$ such that for each phase $i \in \{1, \dots, P\}$ we have a family $A^i = (A^i(t))_{t \in (0, \infty)}$ of open subsets of \mathbb{R}^d with finite perimeter in \mathbb{R}^d . We also require $(A^1(t), \dots, A^P(t))$ to be a partition of \mathbb{R}^d for all $t \in (0, \infty)$ and that all except for the P -th phase have finite mass for all $t \in (0, \infty)$. Moreover, for each $i \in \{1, \dots, P\}$ the associated indicator function $\chi^i(\cdot, t) := \chi_{A^i(t)}$, $t \in (0, \infty)$, satisfies $\chi^i \in L^\infty((0, \infty); \text{BV}_{\text{loc}}(\mathbb{R}^d; \{0, 1\}))$. We shortly write $\chi = (\chi^1, \dots, \chi^P)$.

Given initial data $((\mu_0^{i,j})_{i,j \in \{1, \dots, P\}}, (\chi_0^i)_{i \in \{1, \dots, P\}})$ of the above form, we call the pair (μ, χ) a *De Giorgi type varifold solution for multi-phase mean curvature flow with initial data* (μ_0, χ_0) if the following requirements hold true.

- (Existence of normal speeds) For each phase $i \in \{1, \dots, P\}$, there exists a normal speed $V^i \in L^2((0, \infty); L^2(\mathbb{R}^d, d\omega_t^i))$ in the sense that

$$(111c) \quad \int_{\mathbb{R}^d} \chi^i(\cdot, T) \zeta(\cdot, T) \, dx - \int_{\mathbb{R}^d} \chi_0^i \zeta(\cdot, 0) \, dx = \int_0^T \int_{\mathbb{R}^d} \chi^i \partial_t \zeta \, dx \, dt - \int_0^T \int_{\mathbb{R}^d} V^i \zeta \, d\omega_t^i \, dt$$

for almost every $T \in (0, \infty)$ and all $\zeta \in C_c^\infty(\mathbb{R}^d \times [0, \infty))$.

- (Existence of a generalized mean curvature vector) There exists a generalized mean curvature vector $\mathbf{H} \in L^2((0, \infty); L^2(\mathbb{R}^d, d\omega_t; \mathbb{R}^d))$ in the sense that

$$(111d) \quad \int_0^T \int_{\mathbb{R}^d} \mathbf{H} \cdot B \, d\omega_t \, dt = - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B \, d\mu_t \, dt$$

for almost every $T \in (0, \infty)$ and all $B \in C_c^\infty(\mathbb{R}^d \times [0, \infty); \mathbb{R}^d)$.

- (Sharp energy dissipation principle à la De Giorgi) It holds

$$(111e) \quad \int_{\mathbb{R}^d} 1 \, d\omega_T + \frac{1}{2} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 \, d\omega_t^i \, dt + \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 \, d\omega_t \, dt \leq \int_{\mathbb{R}^d} 1 \, d\omega_0$$

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

- (Compatibility conditions) For all $i, j \in \{1, \dots, P\}$ with $i \neq j$ we require

$$(111f) \quad \omega_t^{i,j} = \omega_t^{j,i}$$

for almost every $t \in (0, \infty)$ and

$$(111g) \quad \langle \lambda_{x,t}^{i,j} \rangle = -\langle \lambda_{x,t}^{j,i} \rangle, \quad \langle \lambda_{x,t}^{i,i} \rangle = 0,$$

$$(111h) \quad V^i(x, t) = -V^j(x, t),$$

$$(111i) \quad |\langle \lambda_{x,t}^{i,j} \rangle|^2 \mathbf{H}(x, t) = (\mathbf{H}(x, t) \cdot \langle \lambda_{x,t}^{i,j} \rangle) \langle \lambda_{x,t}^{i,j} \rangle$$

for almost every $t \in (0, \infty)$ and $\omega_t^{i,j}$ almost every $x \in \mathbb{R}^d$. Finally, for all $i \in \{1, \dots, P\}$

$$(111j) \quad \int_{\mathbb{R}^d} \xi \cdot d\nabla \chi^i(\cdot, t) = \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \xi \cdot p \, d\mu_t^i$$

for almost every $t \in (0, \infty)$ and all $\xi \in C_c^\infty(\mathbb{R}^d)$.

Analogous to the two-phase case, the condition (111j) ensures that $|\nabla \chi^i(\cdot, t)| \leq \omega_t^i$ in the sense of measures for almost every $t \in (0, \infty)$ and all $i \in \{1, \dots, P\}$, and that the associated Radon–Nikodým derivative satisfies

$$(112) \quad \rho^i(x, t) := \frac{d|\nabla \chi^i(\cdot, t)|}{d\omega_t^i}(x) \in [0, 1] \quad \text{for } \omega_t^i \text{ almost every } x \in \mathbb{R}^d,$$

$$(113) \quad \int_{\partial^* \mathcal{A}^i(t)} f \, d\mathcal{H}^{d-1} = \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} f \rho^i(\cdot, t) \, d\omega_t^i \quad \text{for all } f \in C_c^\infty(\mathbb{R}^d).$$

At the time of this writing, we admittedly do not know of an existence proof for such a solution concept which does not require an energy convergence assumption as in [38, 35, 37]. Moreover, the extension to general classes of surface tension matrices does not seem to be immediate. However, what we will argue in this section is that the above concept of varifold solutions for multi-phase mean curvature flow is at least somehow reasonable in the sense that it is subject to a weak-strong uniqueness principle; a computation which we wanted to share with the community. To this end, we first recall the recently introduced concept of a gradient-flow calibration (with respect to the L^2 gradient-flow of the interface energy).

The underlying data is given by a family $\mathcal{A} = (\mathcal{A}^1, \dots, \mathcal{A}^P)$ such that for each phase $i \in \{1, \dots, P\}$ we have a family $\mathcal{A}^i = (\mathcal{A}^i(t))_{t \in [0, T_{\text{cal}}]}$ of non-empty open subsets of \mathbb{R}^d with finite perimeter in \mathbb{R}^d . We also require $(\mathcal{A}^1(t), \dots, \mathcal{A}^P(t))$ to be a partition of \mathbb{R}^d for all $t \in [0, T_{\text{cal}}]$, that all except for the P th phase have finite mass for all $t \in [0, T_{\text{cal}}]$, and that $\partial \mathcal{A}^i(t) = \overline{\partial^* \mathcal{A}^i(t)}$ as well as $\mathcal{H}^{d-1}(\partial \mathcal{A}^i(t) \setminus \partial^* \mathcal{A}^i(t)) = 0$ for all $i \in \{1, \dots, P\}$ and all $t \in [0, T_{\text{cal}}]$. Finally, we write $\mathcal{I}_{i,j}(t)$ for the associated interfaces $\partial^* \mathcal{A}^i(t) \cap \partial^* \mathcal{A}^j(t)$ between phases $i \neq j \in \{1, \dots, P\}$ at time $t \in [0, T_{\text{cal}}]$, and $n_{\mathcal{I}_{i,j}(t)}$ for the corresponding unit normal vector field pointing from $\mathcal{A}^i(t)$ to $\mathcal{A}^j(t)$. We then assume that \mathcal{A} is a calibrated evolution in the precise sense of [24, Definition 2]. In particular, there exists a family of sufficiently regular and sufficiently integrable

vector fields $((\xi_i)_{i \in \{1, \dots, P\}}, (\xi_{i,j})_{i,j \in \{1, \dots, P\}, i \neq j}, B)$ so that

$$(114) \quad \xi_{i,j}(\cdot, t) = \xi_i(\cdot, t) - \xi_j(\cdot, t),$$

$$(115) \quad \xi_{i,j}(\cdot, t) = n_{\mathcal{I}_{i,j}(t)} \text{ along } \mathcal{I}_{i,j}(t),$$

$$(116) \quad c \min\{1, \text{dist}^2(x, \mathcal{I}_{i,j}(t))\} \leq 1 - |\xi_{i,j}(x, t)|,$$

$$(117) \quad \partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j} + (\nabla B)^\top \xi_{i,j} = O(\min\{1, \text{dist}(x, \mathcal{I}_{i,j}(t))\}),$$

$$(118) \quad \xi \cdot (\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j}) = O(\min\{1, \text{dist}^2(x, \mathcal{I}_{i,j}(t))\}),$$

$$(119) \quad B \cdot \xi_{i,j} + \nabla \cdot \xi_{i,j} = O(\min\{1, \text{dist}(x, \mathcal{I}(t))\}),$$

for all $(x, t) \in \mathbb{R}^d \times [0, T_{\text{cal}}]$ and some $c \in (0, 1]$. Moreover, we assume that there exists a family of sufficiently regular and sufficiently integrable weights $(\vartheta^1, \dots, \vartheta^P)$ such that (cf. [24, Definition 4])

$$(120) \quad \vartheta^i(\cdot, t) < 0 \text{ in } \mathcal{A}^i(t), \quad \vartheta^i(\cdot, t) > 0 \text{ in } \mathbb{R}^d \setminus \overline{\mathcal{A}^i(t)},$$

$$(121) \quad (\partial_t \vartheta^i + (B \cdot \nabla) \vartheta^i)(x, t) = O(\min\{1, \text{dist}(x, \partial \mathcal{A}^i(t))\}),$$

$$(122) \quad c \min\{1, \text{dist}(x, \partial \mathcal{A}^i(t))\} \leq |\vartheta^i(x, t)| \leq C \min\{1, \text{dist}(x, \partial \mathcal{A}^i(t))\}$$

for some $c \in (0, 1]$, $C \geq 1$ and all $(x, t) \in \mathbb{R}^d \times [0, T_{\text{cal}}]$.

It turns out that sufficiently smooth solutions to multi-phase mean curvature flow are indeed calibrated in this sense, see [24, Proposition 6 and Lemma 7] for the regime $d = 2$ and $P \geq 2$, or [26, Theorem 1] for the case $d = 3$ and $P = 3$. For the purposes of the present work, it turns out that one needs one additional requirement in form of

$$(123) \quad |(\xi_{i,j} \otimes (I_d - \xi_{i,j} \otimes \xi_{i,j})v + (I_d - \xi_{i,j} \otimes \xi_{i,j})v \otimes \xi_{i,j}) : \nabla B| = O(|v| \min\{1, \text{dist}(x, \mathcal{I}_{i,j}(t))\})$$

for all $(x, t) \in \mathbb{R}^d \times [0, T_{\text{cal}}]$ and all $v \in \mathbb{R}^d$. The above existence proofs only require a minor extension to include this additional condition.

We now define the multi-phase analogues of the relative entropy and the bulk error from (81) and (82), respectively, measuring the difference between a De Giorgi type varifold solution (μ, χ) of multi-phase mean curvature flow and a calibrated evolution \mathcal{A} . These are defined by

$$(124) \quad E[\mu, \chi | \mathcal{A}](t) := \int_{\mathbb{R}^d} 1 \, d\omega_t - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \cdot \xi_{i,j}(\cdot, t) \, d\mu_t^{i,j} \geq 0, \quad t \in [0, T_{\text{cal}}],$$

$$(125) \quad E_{\text{bulk}}[\chi | \mathcal{A}](t) := \sum_{i=1}^P \int_{\mathbb{R}^d} |\chi_{A^i(t)} - \chi_{\mathcal{A}^i(t)}| |\vartheta^i(\cdot, t)| \, dx, \quad t \in [0, T_{\text{cal}}].$$

Note that by (111a)–(111b) we may equivalently write

$$(126) \quad E[\mu, \chi | \mathcal{A}](t) = \sum_{i=1}^P \int_{\mathbb{R}^d} 1 \, d\omega_t^{i,i} + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} 1 - p \cdot \xi_{i,j}(\cdot, t) \, d\mu_t^{i,j}.$$

By the calibration property (114), the anti-symmetry condition (111g), the decomposition (111a), and the compatibility condition (111j) we have

$$\begin{aligned} & -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \cdot \xi_{i,j}(\cdot, t) \, d\mu_t^{i,j} \\ &= \sum_{i,j=1}^P \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \cdot \xi_i(\cdot, t) \, d\mu_t^{j,i} = \sum_{i=1}^P \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \cdot \xi_i(\cdot, t) \, d\mu_t^i = \sum_{i=1}^P \int_{\mathbb{R}^d} (\xi_i \cdot d\nabla \chi^i)(\cdot, t) \end{aligned}$$

so that by an integration by parts

$$(127) \quad -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \cdot \xi_{i,j}(\cdot, t) d\mu_t^{i,j} = -\sum_{i=1}^P \int_{\mathbb{R}^d} \chi^i(\cdot, t) (\nabla \cdot \xi_i)(\cdot, t) dx.$$

Alternatively, writing $n_{i,j}(\cdot, t) := \frac{\nabla \chi_j(\cdot, t)}{|\nabla \chi_j(\cdot, t)|} = -\frac{\nabla \chi_i(\cdot, t)}{|\nabla \chi_i(\cdot, t)|}$ along $\partial^* A^i(t) \cap \partial^* A^j(t)$ and capitalizing on the anti-symmetry $n_{i,j} = -n_{j,i}$ as well as the calibration property (114), we also have

$$(128) \quad -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \cdot \xi_{i,j}(\cdot, t) d\mu_t^{i,j} = -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_{\partial^* A^i(t) \cap \partial^* A^j(t)} (n_{i,j} \cdot \xi_{i,j})(\cdot, t) d\mathcal{H}^{d-1}.$$

Since also by (111b) and (113)

$$(129) \quad \int_{\mathbb{R}^d} 1 d\omega_t = \frac{1}{2} \sum_{i=1}^P \int_{\mathbb{R}^d} 1 - \rho^i(\cdot, t) d\omega_t^i + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_{\partial^* A^i(t) \cap \partial^* A^j(t)} 1 d\mathcal{H}^{d-1},$$

a third representation of the relative entropy is therefore given by

$$(130) \quad E[\mu, \chi | \mathcal{A}](t) = \frac{1}{2} \sum_{i=1}^P \int_{\mathbb{R}^d} 1 - \rho^i(\cdot, t) d\omega_t^i + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_{\partial^* A^i(t) \cap \partial^* A^j(t)} (1 - n_{i,j} \cdot \xi_{i,j})(\cdot, t) d\mathcal{H}^{d-1}.$$

The three equivalent expressions (124), (126) and (130) combined with the length constraint (116) again imply a number of useful coercivity properties for the relative entropy:

$$(131) \quad \int_{\mathbb{R}^d} 1 d\omega_t^{i,i} + \int_{\mathbb{R}^d} 1 - \rho^i(\cdot, t) d\omega_t^i \leq CE[\mu, \chi | \mathcal{A}](t),$$

$$(132) \quad \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} 1 - |\langle \lambda_{x,t}^{i,j} \rangle| d\omega_t^{i,j} \leq CE[\mu, \chi | \mathcal{A}](t),$$

$$(133) \quad \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} |p - \xi_{i,j}(\cdot, t)|^2 d\mu_t^{i,j} \leq CE[\mu, \chi | \mathcal{A}](t),$$

$$(134) \quad \int_{\mathbb{R}^d} \min\{1, \text{dist}^2(\cdot, \mathcal{I}_{i,j}(t))\} d\omega_t^{i,j} \leq CE[\mu, \chi | \mathcal{A}](t),$$

$$(135) \quad \int_{\partial^* A^i(t) \cap \partial^* A^j(t)} |(n_{i,j} - \xi_{i,j})(\cdot, t)|^2 d\mathcal{H}^{d-1} \leq CE[\mu, \chi | \mathcal{A}](t),$$

$$(136) \quad \int_{\partial^* A^i(t) \cap \partial^* A^j(t)} \min\{1, \text{dist}^2(\cdot, \mathcal{I}_{i,j}(t))\} d\mathcal{H}^{d-1} \leq CE[\mu, \chi | \mathcal{A}](t)$$

for all $i \neq j \in \{1, \dots, P\}$ and all $t \in [0, T_{\text{cal}}]$. Here $C = C(\mathcal{A})$ depends on the strong solution, but not on the weak solution. Finally, by (120) and (122) we also have in terms of the multi-phase bulk error

$$(137) \quad \sum_{i=1}^P \int_{\mathbb{R}^d} (\chi_{A^i(t)} - \chi_{\mathcal{A}^i(t)}) \vartheta^i(\cdot, t) dx = E_{\text{bulk}}[\chi | \mathcal{A}](t),$$

$$(138) \quad \int_{\mathbb{R}^d} |\chi_{A^i(t)} - \chi_{\mathcal{A}^i(t)}| \min\{1, \text{dist}(\cdot, \partial \mathcal{A}^i(t))\} dx \leq CE_{\text{bulk}}[\chi | \mathcal{A}](t)$$

for all $i \in \{1, \dots, P\}$ and all $t \in [0, T_{\text{cal}}]$.

We have everything in place to formulate and prove the main result of this section.

Theorem 3 (Weak-strong uniqueness and quantitative stability for multi-phase De Giorgi type varifold solutions). *Let $T_{\text{cal}} \in (0, \infty)$ be a finite time horizon, and let \mathcal{A} be a calibrated evolution on $[0, T_{\text{cal}}]$ in the above sense. Furthermore, let (μ, χ) be a De Giorgi type varifold solution for multi-phase mean curvature flow with initial data (μ_0, χ_0) in the sense of Definition 2.*

Defining the relative entropy $E[\mu, \chi|\mathcal{A}]$ and the bulk error $E_{\text{bulk}}[\chi|\mathcal{A}]$ by (124) and (125), respectively, we then have stability estimates for these quantities in form of

$$(139) \quad E[\mu, \chi|\mathcal{A}](T) \leq E[\mu_0, \chi_0|\mathcal{A}(0)] + C \int_0^T E[\mu, \chi|\mathcal{A}](t) dt,$$

$$(140) \quad E_{\text{bulk}}[\chi|\mathcal{A}](T) \leq E_{\text{bulk}}[\chi_0|\mathcal{A}(0)] + E[\mu_0, \chi_0|\mathcal{A}(0)] + \int_0^T E_{\text{bulk}}[\chi|\mathcal{A}](t) + E[\mu, \chi|\mathcal{A}](t) dt$$

for some constant $C = C(\mathcal{A}) > 0$ and almost every $T \in [0, T_{\text{cal}}]$.

In particular, if $\chi_0^i = \chi_{\mathcal{A}^i(0)}$ and $\mu_0^i = |\nabla \chi_0^i| \otimes (\delta_{n_0^i(x)})_{x \in \mathbb{R}^d}$ hold true with $n_0^i = \frac{\nabla \chi_0^i}{|\nabla \chi_0^i|}$ for all $i \in \{1, \dots, P\}$, then the De Giorgi type varifold solution (μ, χ) reduces to the calibrated evolution \mathcal{A} in the sense that for almost every $t \in [0, T_{\text{cal}}]$ and all $i \in \{1, \dots, P\}$

$$(141) \quad \chi_{\mathcal{A}^i}(t) = \chi_{\mathcal{A}^i(t)} \text{ almost everywhere in } \mathbb{R}^d,$$

$$(142) \quad \mu_t^i = |\nabla \chi_{\mathcal{A}^i}(t)| \otimes (\delta_{n^i(x,t)})_{x \in \mathbb{R}^d},$$

where $n^i(\cdot, t) = \frac{\nabla \chi_{\mathcal{A}^i}(t)}{|\nabla \chi_{\mathcal{A}^i}(t)|}$.

Proof. We again split the proof into three steps. Naturally, the argument is of similar structure as the one for the proof of Theorem 2. We thus focus in the following on the key differences.

Step 1: Proof of (139). The first major goal is to arrive at the multi-phase analogue of the estimate (97). Starting point for this is to rewrite the relative entropy thanks to (127) as

$$E[\mu, \chi|\mathcal{A}](T) = \int_{\mathbb{R}^d} 1 d\omega_T - \sum_{i=1}^P \int_{\mathbb{R}^d} \chi^i(\cdot, T) (\nabla \cdot \xi_i)(\cdot, T) dx,$$

so that the De Giorgi type inequality (111e), the evolution equations (111c), and an integration by parts imply

$$(143) \quad E[\mu, \chi|\mathcal{A}](T) \leq E[\mu_0, \chi_0|\mathcal{A}(0)] - \frac{1}{2} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^i dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t dt \\ + \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} V^i (\nabla \cdot \xi_i) d\omega_t^i dt + \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{\nabla \chi_i}{|\nabla \chi_i|} \cdot \partial_t \xi_i d|\nabla \chi_i| dt.$$

Splitting first $|\nabla \chi_i(\cdot, t)| = \sum_{j=1, j \neq i}^P \mathcal{H}^{d-1} \llcorner S_{i,j}(t)$, where for the purposes of the whole proof we abbreviate $S_{i,j}(t) := \partial^* A^i(t) \cap \partial^* A^j(t)$, and recalling second that $n_{i,j}(\cdot, t) = \frac{\nabla \chi_j(\cdot, t)}{|\nabla \chi_j(\cdot, t)|} = -\frac{\nabla \chi_i(\cdot, t)}{|\nabla \chi_i(\cdot, t)|} = -n_{j,i}(\cdot, t)$ along $\partial^* A^i(t) \cap \partial^* A^j(t)$, the calibration property (114) allows to compute

$$(144) \quad \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{\nabla \chi_i}{|\nabla \chi_i|} \cdot \partial_t \xi_i d|\nabla \chi_i| dt = -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} n_{i,j} \cdot \partial_t \xi_{i,j} d\mathcal{H}^{d-1} dt.$$

In view of the estimates (117)–(118) and the coercivity property (136), the same argument as for (93) entails for all $i, j \in \{1, \dots, P\}$ with $i \neq j$

$$(145) \quad - \int_0^T \int_{S_{i,j}(t)} n_{i,j} \cdot \partial_t \xi_{i,j} d\mathcal{H}^{d-1} dt \\ \leq \int_0^T \int_{S_{i,j}(t)} \xi_{i,j} \otimes (n_{i,j} - \xi_{i,j}) : \nabla B d\mathcal{H}^{d-1} dt + \int_0^T \int_{S_{i,j}(t)} n_{i,j} \cdot (B \cdot \nabla) \xi_{i,j} d\mathcal{H}^{d-1} dt \\ + C \int_0^T E[\mu, \chi|\mathcal{A}](t) dt.$$

Observe next that the structurally analogous argument enabling (144) further guarantees

$$\begin{aligned} & \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} n_{i,j} \cdot (\nabla \cdot (\xi_{i,j} \otimes B)) d\mathcal{H}^{d-1} dt \\ &= - \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{\nabla \chi_i}{|\nabla \chi_i|} \cdot (\nabla \cdot (\xi_i \otimes B)) d|\nabla \chi_i| dt. \end{aligned}$$

Of course, this identity is also valid with the roles of B and $\xi_{i,j}$ as well as B and ξ_i reversed, respectively, and the same arguments moreover show

$$\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} (n_{i,j} \cdot B)(\nabla \cdot \xi_{i,j}) d\mathcal{H}^{d-1} dt = - \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \left(\frac{\nabla \chi_i}{|\nabla \chi_i|} \cdot B \right) (\nabla \cdot \xi_i) d|\nabla \chi_i| dt.$$

Due to the compatibility condition (111j) and the calibration property (114), as well as by replacing the anti-symmetry condition $n_{i,j} = -n_{j,i}$ and the splitting $|\nabla \chi_i(\cdot, t)| = \sum_{j=1, j \neq i}^P \mathcal{H}^{d-1} \llcorner S_{i,j}(t)$ by the anti-symmetry condition of (111g) and the decomposition (111a), respectively, the identity of the previous display may in turn be updated by the previous arguments to

$$\begin{aligned} (146) \quad & \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} (n_{i,j} \cdot B)(\nabla \cdot \xi_{i,j}) d\mathcal{H}^{d-1} dt \\ &= - \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B)(\nabla \cdot \xi_i) d\mu_t^i dt = \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B)(\nabla \cdot \xi_{i,j}) d\mu_t^{i,j} dt. \end{aligned}$$

In particular, the argument in favor of (95) and (96) now implies because of the previous three displays, (130) and (113)

$$\begin{aligned} (147) \quad & \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} n_{i,j} \cdot (B \cdot \nabla) \xi_{i,j} d\mathcal{H}^{d-1} dt \\ &\leq \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B)(\nabla \cdot \xi_{i,j}) d\mu_t^{i,j} dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B d\mu_t dt \\ &\quad + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} n_{i,j} \otimes \xi_{i,j} : \nabla B d\mathcal{H}^{d-1} dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \otimes p : \nabla B d\mu_t dt. \end{aligned}$$

Hence, combining the estimates (145) and (147) ensures

$$\begin{aligned} (148) \quad & - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} n_{i,j} \cdot \partial_t \xi_{i,j} d\mathcal{H}^{d-1} dt \\ &\leq \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B)(\nabla \cdot \xi_{i,j}) d\mu_t^{i,j} dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B d\mu_t dt \\ &\quad - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} \xi_{i,j} \otimes \xi_{i,j} : \nabla B d\mathcal{H}^{d-1} dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \otimes p : \nabla B d\mu_t dt \\ &\quad + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} (n_{i,j} \otimes \xi_{i,j} + \xi_{i,j} \otimes n_{i,j}) : \nabla B d\mathcal{H}^{d-1} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt. \end{aligned}$$

We further post-process this estimate by noting first that the argument of (146) also yields

$$(149) \quad \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} n_{i,j} \otimes \xi_{i,j} : \nabla B d\mathcal{H}^{d-1} dt = \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} p \otimes \xi_{i,j} : \nabla B d\mu_t^{i,j} dt,$$

where we remark that the same identity also holds true with the roles of $n_{i,j}$ and $\xi_{i,j}$ as well as p and $\xi_{i,j}$ reversed, respectively. For the next step, we claim that

$$(150) \quad \begin{aligned} & -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} \xi_{i,j} \otimes \xi_{i,j} : \nabla B d\mathcal{H}^{d-1} dt \\ & \leq -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \xi_{i,j} \otimes \xi_{i,j} : \nabla B d\mu_t^{i,j} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt. \end{aligned}$$

Before we prove (150), let us first observe that inserting (149)–(150) back into (148) and then estimating the resulting quadratic term in the difference between the calibration normals $\xi_{i,j}$ and the varifold normal p from above by (133), we consequently obtain from (143) and (144)

$$(151) \quad \begin{aligned} E[\mu, \chi | \mathcal{A}](T) & \leq E[\mu_0, \chi_0 | \mathcal{A}](0) - \frac{1}{2} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^i dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t dt \\ & \quad + \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} V^i (\nabla \cdot \xi_i) d\omega_t^i dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B d\mu_t dt \\ & \quad + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B) (\nabla \cdot \xi_{i,j}) d\mu_t^{i,j} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt. \end{aligned}$$

For a proof of (150), a purely algebraic manipulation and $|\xi_{i,j} \cdot (\xi_{i,j} - n_{i,j})| \leq 3(1 - n_{i,j} \cdot \xi_{i,j})$ gives

$$(152) \quad \begin{aligned} 2\xi_{i,j} \otimes \xi_{i,j} & = ((\xi_{i,j} - n_{i,j}) \otimes \xi_{i,j} + \xi_{i,j} \otimes (\xi_{i,j} - n_{i,j})) + n_{i,j} \otimes \xi_{i,j} + \xi_{i,j} \otimes n_{i,j} \\ & = (I_d - \xi_{i,j} \otimes \xi_{i,j})(\xi_{i,j} - n_{i,j}) \otimes \xi_{i,j} + \xi_{i,j} \otimes (I_d - \xi_{i,j} \otimes \xi_{i,j})(\xi_{i,j} - n_{i,j}) \\ & \quad + n_{i,j} \otimes \xi_{i,j} + \xi_{i,j} \otimes n_{i,j} + O(1 - n_{i,j} \cdot \xi_{i,j}). \end{aligned}$$

Hence, plugging in (149) and estimating based on (123), (130), (135) and (136) we get

$$\begin{aligned} & -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{S_{i,j}(t)} \xi_{i,j} \otimes \xi_{i,j} : \nabla B d\mathcal{H}^{d-1} dt \\ & \leq -\frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \otimes \xi_{i,j} + \xi_{i,j} \otimes p) : \nabla B d\mu_t^{i,j} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt. \end{aligned}$$

Reading (152) with $n_{i,j}$ replaced by p and estimating the resulting error terms this time by means of (123), (126), (133) and (134), we deduce that the previous display implies the asserted estimate (150). In particular, we may continue with (151).

To this end, relying on the conditions of (111f) and (111h) and yet again the calibration property (114), we have

$$\sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} V^i (\nabla \cdot \xi_i) d\omega_t^i dt = \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} V^i (\nabla \cdot \xi_{i,j}) d\omega_t^{i,j} dt.$$

We then insert the decomposition (111a), complete squares twice for off-diagonal terms, and estimate the diagonal terms with the help of Young's inequality and the coercivity property (131)

to obtain together with the previous display and finally (119) as well as (134) that

$$\begin{aligned}
& -\frac{1}{2} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^i dt + \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} V^i (\nabla \cdot \xi_i) d\omega_t^i dt \\
& + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B) (\nabla \cdot \xi_{i,j}) d\mu_t^{i,j} dt \\
& \leq -\frac{1}{4} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^{i,i} dt - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i - (\nabla \cdot \xi_{i,j})|^2 d\omega_t^{i,j} dt \\
& + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B) (\nabla \cdot \xi_{i,j}) d\mu_t^{i,j} dt \\
& + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} (\nabla \cdot \xi_{i,j})^2 d\omega_t^{i,j} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt \\
& \leq -\frac{1}{4} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^{i,i} dt - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i - (\nabla \cdot \xi_{i,j})|^2 d\omega_t^{i,j} dt \\
& + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |(B \cdot \xi_{i,j}) + (\nabla \cdot \xi_{i,j})|^2 d\omega_t^{i,j} dt - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} (B \cdot \xi_{i,j})^2 d\omega_t^{i,j} dt \\
& + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} B \cdot (p - \xi_{i,j}) (\nabla \cdot \xi_{i,j}) d\omega_t^{i,j} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt \\
(153) \quad & \leq -\frac{1}{4} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^{i,i} dt - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i - (\nabla \cdot \xi_{i,j})|^2 d\omega_t^{i,j} dt \\
& + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} B \cdot (p - \xi_{i,j}) (\nabla \cdot \xi_{i,j}) d\omega_t^{i,j} dt - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} (B \cdot \xi_{i,j})^2 d\omega_t^{i,j} dt \\
& + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt.
\end{aligned}$$

Moreover, by (111a) and (111b) as well as (111i) in form of the identity $|\mathbf{H}(x, t)|^2 |\langle \lambda_{x,t}^{i,j} \rangle|^2 = |\mathbf{H}(x, t) \cdot \langle \lambda_{x,t}^{i,j} \rangle|^2$, we rewrite one of the helpful dissipation-terms as

$$\begin{aligned}
& -\frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t dt \\
(154) \quad & = -\frac{1}{2} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t^{i,i} dt - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{H}|^2 (1 - |\langle \lambda_{x,t}^{i,j} \rangle|^2) d\omega_t^{i,j} dt \\
& - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{H}(x, t) \cdot \langle \lambda_{x,t}^{i,j} \rangle|^2 d\omega_t^{i,j} dt.
\end{aligned}$$

Since we may exploit (111d) and then decompose in a similar fashion using (111i) this time directly, we also get

$$-\int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B d\mu_t dt$$

$$\begin{aligned}
(155) \quad &= \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \mathbf{H} \cdot B d\omega_t^{i,i} dt + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} (1 - |\langle \lambda_{x,t}^{i,j} \rangle|^2) \mathbf{H} \cdot B d\omega_t^{i,j} dt \\
&\quad + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} (\mathbf{H} \cdot \langle \lambda_{x,t}^{i,j} \rangle) B \cdot p d\mu_t^{i,j} dt
\end{aligned}$$

Thanks to the control provided by (131) and (132), the bounds $1 - |\langle \lambda_{x,t}^{i,j} \rangle| \leq 1 - |\langle \lambda_{x,t}^{i,j} \rangle|^2 \leq 2(1 - |\langle \lambda_{x,t}^{i,j} \rangle|)$, and an absorption argument, adding (154) to (155) thus results in the estimate

$$\begin{aligned}
& - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B d\mu_t dt \\
& \leq - \frac{1}{4} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t^{i,i} dt - \frac{1}{4} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{H}|^2 (1 - |\langle \lambda_{x,t}^{i,j} \rangle|^2) d\omega_t^{i,j} dt \\
& \quad - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{H}(x,t) \cdot \langle \lambda_{x,t}^{i,j} \rangle|^2 d\omega_t^{i,j} dt + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} (\mathbf{H} \cdot \langle \lambda_{x,t}^{i,j} \rangle) B \cdot p d\mu_t^{i,j} dt \\
& \quad + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt.
\end{aligned}$$

Hence, as a consequence of the previous display, completing the square, adding zero, recalling the control provided by (133), and finally performing yet another absorption argument, it holds

$$\begin{aligned}
& - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\mathbf{H}|^2 d\omega_t dt - \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (I_d - p \otimes p) : \nabla B d\mu_t dt \\
& \leq - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |(\mathbf{H}(x,t) \cdot \langle \lambda_{x,t}^{i,j} \rangle) - (B \cdot \xi_{i,j})|^2 d\omega_t^{i,j} dt \\
& \quad + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} (B \cdot \xi_{i,j})^2 d\omega_t^{i,j} dt \\
& \quad - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} ((\mathbf{H} \cdot \langle \lambda_{x,t}^{i,j} \rangle) - (B \cdot \xi_{i,j})) B \cdot (\xi_{i,j} - p) d\mu_t^{i,j} dt \\
& \quad - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} (B \cdot \xi_{i,j}) B \cdot (\xi_{i,j} - p) d\mu_t^{i,j} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt \\
(156) \quad & \leq \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} (B \cdot \xi_{i,j})^2 d\omega_t^{i,j} dt \\
& \quad - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} (B \cdot \xi_{i,j}) B \cdot (\xi_{i,j} - p) d\mu_t^{i,j} dt + C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt.
\end{aligned}$$

Adding (153) to (156), observing that as a result of (119), (133) and (134)

$$- \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} ((B \cdot \xi_{i,j}) + (\nabla \cdot \xi_{i,j})) B \cdot (\xi_{i,j} - p) d\mu_t^{i,j} dt \leq C \int_0^T E[\mu, \chi | \mathcal{A}](t) dt,$$

we may thus upgrade the preliminary estimate (151) to

$$(157) \quad E[\mu, \chi|\mathcal{A}](T) \leq E[\mu_0, \chi_0|\mathcal{A}](0) - \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i - (\nabla \cdot \xi_{i,j})|^2 d\omega_t^{i,j} dt \\ - \frac{1}{4} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^{i,i} dt + C \int_0^T E[\mu, \chi|\mathcal{A}](t) dt.$$

This, however, is already a stronger form of what is claimed by (139).

Step 2: Proof of (140). Analogous to (104) and (105), relying this time on (121), the lower bound from (122), the coercivity property (138), and the compatibility condition (111j), we obtain

$$E_{\text{bulk}}[\chi|\mathcal{A}](T) \leq E_{\text{bulk}}[\chi_0|\mathcal{A}](0) - \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} V^i \vartheta^i d\omega_t^i dt \\ + \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B) \vartheta^i d\mu_t^i dt + C \int_0^T E_{\text{bulk}}[\chi|\mathcal{A}](t) dt.$$

Due to adding zero, the upper bound from (122), the decomposition (111a), and the coercivity properties (131), (133) and (134), it further holds

$$\sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B) \vartheta^i d\mu_t^i dt \\ \leq - \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (p \cdot B) \vartheta^i d\mu_t^{i,j} dt + C \int_0^T E[\mu, \chi|\mathcal{A}](t) dt \\ \leq - \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (\xi_{i,j} \cdot B) \vartheta^i d\omega_t^{i,j} dt + C \int_0^T E[\mu, \chi|\mathcal{A}](t) dt \\ \leq \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (\nabla \cdot \xi_{i,j}) \vartheta^i d\omega_t^{i,j} dt + C \int_0^T E[\mu, \chi|\mathcal{A}](t) dt$$

where in the last step, we also exploited (119). Moreover, inserting first the decomposition (111a), rewriting off-diagonal terms based on the first condition of (111f) and (111h), and finally estimating diagonal terms by Young's inequality and (131), we have

$$- \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} V^i \vartheta^i d\omega_t^i dt \leq \frac{1}{4} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^{i,i} dt - \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} V^i \vartheta^i d\omega_t^{i,j} dt \\ + C \int_0^T E[\mu, \chi|\mathcal{A}](t) dt.$$



Adding the previous two displays and then estimating by Young's inequality, the upper bound from (122) as well as the control provided by (134), we infer

$$E_{\text{bulk}}[\chi|\mathcal{A}](T) \leq E_{\text{bulk}}[\chi_0|\mathcal{A}](0) + \frac{1}{4} \sum_{i=1}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i|^2 d\omega_t^{i,i} dt \\ + \frac{1}{2} \sum_{i,j=1, i \neq j}^P \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |V^i - (\nabla \cdot \xi_{i,j})|^2 d\omega_t^{i,j} dt \\ + C \int_0^T E_{\text{bulk}}[\chi|\mathcal{A}](t) + E[\mu, \chi|\mathcal{A}](t) dt,$$

so that adding the previous display to (157) finally yields (140) as desired.

Step 3: Proof of (141)–(142). This follows from straightforward arguments based on the coercivity properties of the relative entropy and the bulk error. \square

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 948819)  , and from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2047/1 – 390685813. The content of this paper was developed and parts of it were written during a visit of the first author to the Hausdorff Center of Mathematics (HCM), University of Bonn. The hospitality and the support of HCM are gratefully acknowledged.

REFERENCES

- [1] H. Abels. On generalized solutions of two-phase flows for viscous incompressible fluids. *Interfaces Free Bound.*, 9(1):31–65, 2007.
- [2] W. K. Allard. On the first variation of a varifold. *Ann. of Math. (2)*, 95(3):417–491, 1972.
- [3] F. Almgren, J. E. Taylor, and L. Wang. Curvature-driven flows: A variational approach. *SIAM J. Control Optim.*, 31(2):387–438, 1993.
- [4] L. Ambrosio. Movimenti minimizzanti. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl.(5)*, 19:191–246, 1995.
- [5] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Birkhäuser, 2008.
- [6] G. Bellettini and L. Mugnai. Some aspects of the variational nature of mean curvature flow. *J. Eur. Math. Soc. (JEMS)*, 10(4):1013–1036, 2008.
- [7] E. B. Bogomol’nyi. The stability of classical solutions. *Soviet J. Nuclear Phys.*, 24(4):861–870, 1976.
- [8] K. A. Brakke. *The motion of a surface by its mean curvature*. Princeton University Press, Princeton, 1978.
- [9] L. Bronsard and R. V. Kohn. Motion by mean curvature as the singular limit of Ginzburg–Landau dynamics. *J. Differential Equations*, 90(2):211–237, 1991.
- [10] A. Chambolle and T. Laux. Mullins–Sekerka as the Wasserstein flow of the perimeter. *Proc. Amer. Math. Soc.*, 149(7):2943–2956, 2021.
- [11] Y. G. Chen, Y. Giga, and S. Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differential Geom.*, 33(3):749–786, 1991.
- [12] O. Chodosh and C. Mantoulidis. Minimal surfaces and the Allen–Cahn equation on 3-manifolds: index, multiplicity, and curvature estimates. *Ann. of Math. (2)*, 191(1):213–328, 2020.
- [13] C. M. Dafermos. The second law of thermodynamics and stability. *Arch. Ration. Mech. Anal.*, 70(2):167–179, 1979.
- [14] E. De Giorgi. New problems on minimizing movements. In *Boundary value problems for partial differential equations and applications*, volume 29 of *RMA Res. Notes Appl. Math.*, pages 81–98. Masson, Paris, 1993.
- [15] G. De Philippis and T. Laux. Implicit time discretization for the mean curvature flow of mean convex sets. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, (5):911–930, 2020.
- [16] M. del Pino, M. Kowalczyk, and J. Wei. On De Giorgi’s conjecture in dimension $N \geq 9$. *Ann. of Math. (2)*, 174(3):1485–1569, 2011.
- [17] R. J. DiPerna. Uniqueness of solutions to hyperbolic conservation laws. *Indiana Univ. Math. J.*, 28(1):137–188, 1979.
- [18] L. C. Evans, H. M. Soner, and P. E. Souganidis. Phase transitions and generalized motion by mean curvature. *Comm. Pure Appl. Math.*, 45(9):1097–1123, 1992.
- [19] L. C. Evans and J. Spruck. Motion of level sets by mean curvature I. *J. Differential Geom.*, 33(3):635–681, 1991.
- [20] L. C. Evans and J. Spruck. Motion of level sets by mean curvature IV. *J. Geom. Anal.*, 5(1):77–114, 1995.
- [21] E. Feireisl, B. J. Jin, and A. Novotný. Relative entropies, suitable weak solutions, and weak-strong uniqueness for the compressible Navier–Stokes system. *J. Math. Fluid Mech.*, 14(4):717–730, 2012.
- [22] E. Feireisl and A. Novotný. Weak-strong uniqueness property for the full Navier–Stokes–Fourier system. *Arch. Ration. Mech. Anal.*, 204(2):683–706, 2012.
- [23] J. Fischer and S. Hensel. Weak-strong uniqueness for the Navier–Stokes equation for two fluids with surface tension. *Arch. Ration. Mech. Anal.*, 236(2):967–1087, 2020.
- [24] J. Fischer, S. Hensel, T. Laux, and T. M. Simon. The local structure of the energy landscape in multi-phase mean curvature flow: Weak-strong uniqueness and stability of evolutions. *arXiv preprint (v2)*, 2020. arXiv:2003.05478v2.
- [25] M. Gage and R. S. Hamilton. The heat equation shrinking convex plane curves. *J. Differential Geom.*, 23(1):69–96, 1986.

- [26] S. Hensel and T. Laux. Weak-strong uniqueness for the mean curvature flow of double bubbles. *arXiv preprint*, 2021. arXiv:2108.01733.
- [27] G. Huisken. Asymptotic behavior for singularities of the mean curvature flow. *J. Differential Geom.*, 31(1):285–299, 1990.
- [28] T. Ilmanen. Convergence of the Allen–Cahn equation to Brakke’s motion by mean curvature. *J. Differential Geom.*, 38(2):417–461, 1993.
- [29] T. Ilmanen. Elliptic regularization and partial regularity for motion by mean curvature. *Mem. Amer. Math. Soc.*, 108(520):x+90, 1994.
- [30] R. Jerrard and D. Smets. On the motion of a curve by its binormal curvature. *J. Eur. Math. Soc.*, 17(6):1487–1515, 2015. doi:10.4171/JEMS/536.
- [31] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [32] L. Kim and Y. Tonegawa. On the mean curvature flow of grain boundaries. *Ann. Inst. Fourier*, 67(1):43–142, 2017.
- [33] T. Laux. Distributional solutions to mean curvature flow. *arXiv preprint*, 2021. arXiv:2108.08347.
- [34] T. Laux and J. Lelmi. De Giorgi’s inequality for the thresholding scheme with arbitrary mobilities and surface tensions. *arXiv preprint*, 2021.
- [35] T. Laux and F. Otto. Convergence of the thresholding scheme for multi-phase mean-curvature flow. *Calc. Var. Partial Differential Equations*, 55(5), 2016.
- [36] T. Laux and F. Otto. The thresholding scheme for mean curvature flow and De Giorgi’s ideas for minimizing movements. *Adv. Stud. Pure Math.*, 85:63–93, 2020. The Role of Metrics in the Theory of Partial Differential Equations, Mathematical Society of Japan.
- [37] T. Laux and T. M. Simon. Convergence of the Allen–Cahn equation to multiphase mean curvature flow. *Comm. Pure Appl. Math.*, 71(8):1597–1647, 2018.
- [38] S. Luckhaus and T. Sturzenhecker. Implicit time discretization for the mean curvature flow equation. *Calc. Var. Partial Differential Equations*, 3(2):253–271, 1995.
- [39] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc.*, 8:1–48, 2006.
- [40] A. Mielke. On evolutionary Γ -convergence for gradient systems. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 187–249. Springer, [Cham], 2016.
- [41] M. Mizuno and Y. Tonegawa. Convergence of the Allen-Cahn equation with Neumann boundary conditions. *SIAM J. Math. Anal.*, 47(3):1906–1932, 2015.
- [42] M. Mizuno and Y. Tonegawa. Erratum to “Convergence of the Allen-Cahn equation with Neumann boundary conditions” [MR3348119]. *SIAM J. Math. Anal.*, 48(4):3035–3036, 2016.
- [43] L. Modica and S. Mortola. Un esempio di Gamma-convergenza. *Bollettino della Unione Matematica Italiana B (5)*, 14(1):285–299, 1977.
- [44] L. Mugnai and M. Röger. The Allen-Cahn action functional in higher dimensions. *Interfaces Free Bound.*, 10(1):45–78, 2008.
- [45] G. D. Philippis, A. D. Rosa, and F. Ghiraldin. Rectifiability of varifolds with locally bounded first variation with respect to anisotropic surface energies. *Comm. Pure Appl. Math.*, 71(6):1123–1148, 2017.
- [46] A. Pisante and F. Punzo. Allen–Cahn approximation of mean curvature flow in Riemannian manifolds II: Brakke’s flows. *Commun. Contemp. Math.*, 17(5):1450041, 35, 2015.
- [47] A. Pisante and F. Punzo. Allen–Cahn approximation of mean curvature flow in Riemannian manifolds I: uniform estimates. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 15:309–341, 2016.
- [48] M. Röger and R. Schätzle. On a modified conjecture of De Giorgi. *Math. Z.*, 254(4):675–714, 2006.
- [49] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg–Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.
- [50] S. Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.*, 31(4):1427–1451, 2011.
- [51] J. Serrin. The initial value problem for the Navier-Stokes equations. In *Nonlinear Problems (Proc. Sympos., Madison, Wis., 1962)*, pages 69–98. Univ. of Wisconsin Press, Madison, Wis., 1963.
- [52] Y. Tonegawa. Integrality of varifolds in the singular limit of reaction-diffusion equations. *Hiroshima Math. J.*, 33(3):323–341, 2003.
- [53] E. Wiedemann. Weak-strong uniqueness in fluid dynamics. In *Partial Differential Equations in Fluid Mechanics*, pages 289–326. Cambridge University Press, 2018.

INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA (IST AUSTRIA), AM CAMPUS 1, 3400 KLOSTERNEUBURG, AUSTRIA

E-mail address: `sebastian.hensel@ist.ac.at`

Current address: Hausdorff Center for Mathematics, Universität Bonn, Endenicher Allee 62, 53115 Bonn, Germany (`sebastian.hensel@hcm.uni-bonn.de`)

HAUSDORFF CENTER FOR MATHEMATICS, UNIVERSITÄT BONN, ENDENICHER ALLEE 62, 53115 BONN, GERMANY

E-mail address: `tim.laux@hcm.uni-bonn.de`