

# Quantitative convergence of the vectorial Allen–Cahn equation towards multiphase mean curvature flow

Julian Fischer and Alice Marveggio

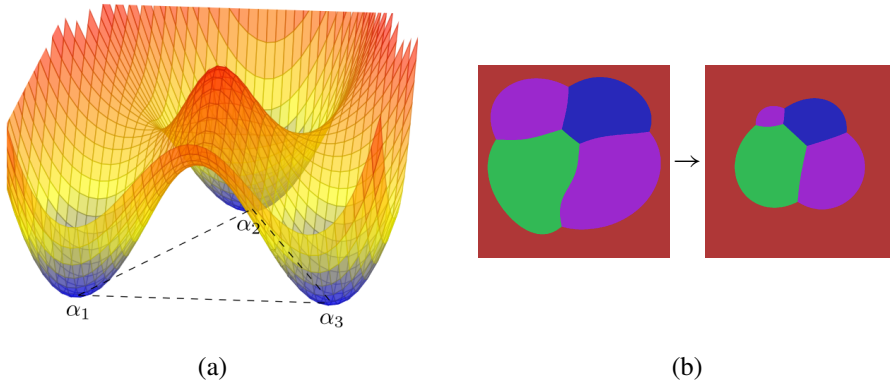
**Abstract.** Phase-field models such as the Allen–Cahn equation may give rise to the formation and evolution of geometric shapes, a phenomenon that may be analyzed rigorously in suitable scaling regimes. In its sharp-interface limit, the vectorial Allen–Cahn equation with a potential with  $N \geq 3$  distinct minima has been conjectured to describe the evolution of branched interfaces by multiphase mean curvature flow. In the present work, we give a rigorous proof for this statement in two and three ambient dimensions and for a suitable class of potentials: as long as a strong solution to multiphase mean curvature flow exists, solutions to the vectorial Allen–Cahn equation with well-prepared initial data converge towards multiphase mean curvature flow in the limit of vanishing interface width parameter  $\varepsilon \searrow 0$ . We even establish the rate of convergence  $O(\varepsilon^{1/2})$ . Our approach is based on the gradient-flow structure of the Allen–Cahn equation and its limiting motion: building on the recent concept of “gradient-flow calibrations” for multiphase mean curvature flow, we introduce a notion of relative entropy for the vectorial Allen–Cahn equation with multi-well potential. This enables us to overcome the limitations of other approaches, e.g. avoiding the need for a stability analysis of the Allen–Cahn operator or additional convergence hypotheses for the energy at positive times.

## 1. Introduction

In the present work, we study the behavior of solutions to the vector-valued Allen–Cahn equation

$$\partial_t u_\varepsilon = \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \quad (1.1)$$

(with  $W$  being an  $N$ -well potential, see e.g. Figure 1a, and  $u_\varepsilon: \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}^{N-1}$ ) in the limit of vanishing interface width  $\varepsilon \rightarrow 0$ . We prove that for a suitable class of  $N$ -well potentials  $W$ , in the limit  $\varepsilon \rightarrow 0$  the solutions  $u_\varepsilon$  describe a branched interface evolving by multiphase mean curvature flow (see Figure 1b), provided that a classical solution to the latter exists and provided that one starts with a sequence of well-prepared initial data  $u_\varepsilon(\cdot, 0)$ . For quantitatively well-prepared initial data  $u_\varepsilon(\cdot, 0)$ , we even establish a rate of convergence  $O(\varepsilon^{1/2})$  towards the multiphase mean curvature flow limit.



**Figure 1.** (a) A triple-well potential that attains its minimum at the three points  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ . (b) A partition of  $\mathbb{R}^2$  evolving by multiphase mean curvature flow, corresponding to the sharp-interface limit  $\varepsilon \rightarrow 0$  of the vectorial Allen–Cahn equation (1.1) with  $N$ -well potential  $W$ .

The Allen–Cahn equation (1.1) with  $N$ -well potential is an important example of a phase-field model, an evolution equation for an order parameter  $u_\varepsilon$  that may vary in space and time. Phase-field models may give rise to the formation and evolution of geometric shapes, a phenomenon that becomes amenable to a rigorous mathematical analysis in suitable scaling regimes. For several important structural classes of potentials  $W$ , such a rigorous analysis has long been available for the Allen–Cahn equation: For instance, for the scalar Allen–Cahn equation with two-well potential  $W$  – that is, for (1.1) with  $N = 2$  – the convergence towards (two-phase) mean curvature flow in the limit  $\varepsilon \rightarrow 0$  has been established by De Mottoni and Schatzman [9], Bronsard and Kohn [5], Chen [7], Ilmanen [18], and Evans, Soner, and Souganidis [10] in the context of three different notions of solutions to mean curvature flow (namely, strong solutions, Brakke solutions, resp. viscosity solutions). In such two-phase situations, sharp-interface limits have also been established for more complex phase-field models [1–3, 8, 11], typically based on an approach that relies on matched asymptotic expansions and a stability analysis of the PDE linearized around a transition profile. Beyond the case of two-well potentials, results have been much more scarce. One of the few well-understood settings is the case of the Ginzburg–Landau equation, which corresponds to the Allen–Cahn equation (1.1) with a Sombbrero-type potential  $W(u) = (1 - |u|^2)^2$  and  $N = 3$ , i.e. with a potential that features a continuum of minima at  $\{u \in \mathbb{R}^2 : |u| = 1\}$ . In this case, the convergence of solutions to (codimension two) vortex filaments evolving by mean curvature has been shown in dimensions  $d \geq 3$  by Jerrard and Soner [20], Lin [23], and Bethuel, Orlandi, and Smets [4].

In contrast, for the (vectorial) Allen–Cahn equation (1.1) with a potential  $W$  with  $N \geq 3$  distinct minima, the only previous results on the sharp-interface limit have been a formal expansion analysis by Bronsard and Reitich [6] and a convergence result that is

conditional on the convergence of the Allen–Cahn energy

$$E[u_\varepsilon] := \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{W(u_\varepsilon)}{\varepsilon} dx$$

at positive times (more precisely, in  $L^1([0, T])$ ) by Laux and Simon [22]. In particular, to the best of our knowledge not even an unconditional proof of qualitative convergence for well-prepared initial data has been available so far. One of the main challenges that has prevented a full analysis is the emergence of “branching” interfaces in the (conjectured) limit of multiphase mean curvature flow (see Figure 1b), corresponding to a geometric singularity in the limiting motion.

In the present work, we introduce a relative energy approach for the problem of the sharp-interface limit of the vectorial Allen–Cahn equation in a multiphase setting: Building on the concept of “gradient-flow calibrations” that was introduced by Hensel, Laux, Simon, and the first author [13] precisely for the purpose of handling these branching singularities in multiphase mean curvature flow and combining it with ideas from [14], we introduce a notion of relative energy for the Allen–Cahn equation

$$E[u_\varepsilon|\xi] := \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{W(u_\varepsilon)}{\varepsilon} + \sum_{i=1}^N \xi_i \cdot \nabla \psi_i(u_\varepsilon) dx.$$

Here, the  $\xi_i$  denote a “gradient-flow calibration” for the strong solution to multiphase mean curvature flow; in particular,  $\xi_{i,j}(x, t) := \xi_i - \xi_j$  is an extension of the unit normal vector field of the interface between phases  $i$  and  $j$  in the strong solution to mean curvature flow at time  $t$ . The  $\psi_i: \mathbb{R}^{N-1} \rightarrow [0, 1]$  are suitable  $C^{1,1}$  functions that serve as phase indicator functions; in particular, denoting the  $N$  minima of the  $N$ -well potential  $W$  by  $\alpha_j$  ( $1 \leq j \leq N$ ), the functions  $\psi_i$  satisfy  $\psi_i(\alpha_j) = \delta_{ij}$ . Note that the functions  $\psi_j - \psi_i$  will play a role that is somewhat similar to the role of the functions  $\psi(u) = \int_0^u \sqrt{2W(s)} ds$  in the Modica–Mortola trick for a two-well potential  $W: \mathbb{R} \rightarrow \mathbb{R}_0^+$  like  $W(u) = \frac{9}{8}(1 - u^2)^2$ .

The properties of the gradient-flow calibration  $\xi_i$  and the assumptions on the functions  $\psi_i: \mathbb{R}^{N-1} \rightarrow [0, 1]$  will ensure that the estimate  $|\sum_{i=1}^N \xi_i \cdot \nabla \psi_i(u_\varepsilon)| \leq \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon)$  holds, thereby guaranteeing coercivity of the relative energy  $E[u_\varepsilon|\xi]$ . In our main result, we prove that for suitable initial data  $u_\varepsilon(\cdot, 0)$  we have  $\|\psi_i(u_\varepsilon(\cdot, t)) - \bar{\chi}_i(\cdot, t)\|_{L^1(\mathbb{R}^d)} \leq C\varepsilon^{1/2}$  for all  $t \leq T$ , where the  $\bar{\chi}_i$  denote the phase indicator functions from the strong solution to multiphase mean curvature flow.

Rigorous results on sharp-interface limits for phase-field models – such as our result – are also of particular interest from a numerical perspective: in evolution equations for interfaces like e.g. mean curvature flow, the occurrence of topological changes typically poses a challenge for numerical simulations. One approach to the simulation of evolving interfaces is to construct a mesh that discretizes the initial interface and to numerically evolve the resulting mesh over time; however, it is then a highly nontrivial (and still widely open) question how to continue the numerical mesh beyond a topology change in a numerically consistent way. An alternative approach to the simulation of evolving interfaces that

avoids this issue is phase-field models, in which the geometric evolution equation for the interface is replaced by an evolution equation for an order parameter posed on the entire space, allowing also for “mixtures” of the phases at the transition regions. The natural diffuse-interface approximation for multiphase mean curvature flow is given by the vector-valued Allen–Cahn equation with  $N$ -well potential (1.1). The advantage of phase-field approximations for geometric motions such as (1.1) is that one may solve them numerically using standard discretization schemes for parabolic PDEs; however, to establish convergence of the overall scheme towards the original interface evolution problem, it is necessary to rigorously justify the sharp-interface limit for the diffuse-interface model.

**Notation.** Throughout the paper, we use standard notation for parabolic PDEs. By  $\dot{H}^1(\mathbb{R}^d)$  we denote the space of functions that have a weak derivative  $\nabla u \in L^2(\mathbb{R}^d)$  and (in the case  $d \geq 3$ ) decay at infinity. In particular, for a function  $u \in L^2([0, T]; \dot{H}^1(\mathbb{R}^d))$  we denote by  $\nabla u$  its (weak) spatial gradient and by  $\partial_t u$  its (weak) time derivative. For functions defined on phase space, like our potential  $W: \mathbb{R}^{N-1} \rightarrow [0, \infty)$  or the approximate phase indicator functions  $\psi_i: \mathbb{R}^{N-1} \rightarrow [0, \infty)$ , we denote their gradients by  $\partial_u W$  resp.  $\partial_u \psi_i$ . For a smooth interface  $I_{i,j}$ , we denote its mean curvature vector by  $\bar{H}_{i,j}$ .

## 2. Main result

Our main result identifies the sharp-interface limit  $\varepsilon \rightarrow 0$  for the vectorial Allen–Cahn equation (1.1) for a sufficiently broad class of  $N$ -well potentials  $W$  characterized by the following conditions:

- (A1) Let  $W: \mathbb{R}^{N-1} \rightarrow [0, \infty)$  be an  $N$ -well potential of class  $C_{\text{loc}}^{1,1}(\mathbb{R}^{N-1})$  that attains its minimum  $W(u) = 0$  precisely in  $N$  distinct points  $\alpha_1, \dots, \alpha_N \in \mathbb{R}^{N-1}$ . Assume that there exists an integer  $q \geq 2$  and constants  $C, c > 0$  such that in a neighborhood of each  $\alpha_i$  we have

$$c|u - \alpha_i|^q \leq W(u) \leq C|u - \alpha_i|^q.$$

- (A2) Let  $U \subset \mathbb{R}^{N-1}$  be a bounded convex open set with piecewise  $C^1$  boundary and  $\{\alpha_1, \dots, \alpha_N\} \subset \bar{U}$ . Suppose that  $\partial_u W(u)$  points towards  $U$  for any  $u \in \partial U$ .
- (A3) Suppose that for any two distinct  $i, j \in \{1, \dots, N\}$ , there exists a unique minimizing path  $\gamma_{i,j}$  connecting  $\alpha_i$  to  $\alpha_j$  in the sense  $\int_{\gamma_{i,j}} \sqrt{2W(\gamma)} \, d\gamma_{i,j} = \inf_{\gamma} \int_{\gamma} \sqrt{2W(\gamma)} \, d\gamma = 1$ , where the infimum is taken over all continuously differentiable paths  $\gamma$  connecting  $\alpha_i$  to  $\alpha_j$ .
- (A4) Suppose that there exist continuously differentiable functions  $\psi_i: \bar{U} \rightarrow [0, 1]$ ,  $1 \leq i \leq N$ , and a disjoint partition of  $\bar{U}$  into sets  $\bar{\mathcal{T}}_{i,j}$ ,  $i < j \in \{1, \dots, N\}$ , subject to the following properties:
- For any  $i < j \in \{1, \dots, N\}$ , we have  $\alpha_i, \alpha_j \in \bar{\mathcal{T}}_{i,j}$  and  $\gamma_{i,j} \in \bar{\mathcal{T}}_{i,j}$ .
  - For any  $i \in \{1, \dots, N\}$ , we have  $\psi_i(\alpha_i) = 1$  and  $\psi_i(u) < 1$  for  $u \neq \alpha_i$ .

- Suppose that on  $\mathcal{T}_{i,j}$ , all  $\psi_k$  with  $k \notin \{i, j\}$  vanish.
- Set  $\psi_0 := 1 - \sum_{i=1}^N \psi_i$  to achieve  $\sum_{i=0}^N \psi_i \equiv 1$  and define  $\psi_{i,j} := \psi_j - \psi_i$ . Suppose that there exists  $\delta > 0$  such that for any  $i < j \in \{1, \dots, N\}$  and any  $u \in \mathcal{T}_{i,j}$  we have

$$\begin{aligned} & \left| \frac{1}{2} \partial_u \psi_{i,j}(u) \right|^2 + \left( \frac{5}{4} + \delta \right) \left| \frac{1}{2} \partial_u \psi_0(u) \right|^2 + \delta |\partial_u \psi_{i,j}(u) \cdot \partial_u \psi_0(u)| \\ & \leq 2W(u). \end{aligned}$$

Additionally, suppose there exists a constant  $C > 0$  such that for any distinct  $i, j \in \{1, \dots, N\}$  and any  $u \in \mathcal{T}_{i,j}$  it holds that  $|\partial_u \psi_i(u)| \leq C \sqrt{2W(u)}$ .

The assumption that our potential  $W$  has a finite set of minima as stated in (A1) is fundamental for the scaling limit we consider, as a different structure of the potential would give rise to a different limiting motion – recall that for instance a Sombbrero-type potential would lead to (codimension two) vortex filament structures [4, 20]. Assumption (A2) is rather mild, ensuring the existence of bounded weak solutions to the vectorial Allen–Cahn equation by a maximum principle (see Remark 4). Condition (A3) ensures that for each pair of minima, there is a unique optimal profile connecting the two phases; furthermore, it fixes the normalization for  $W$  as well as the surface energy density for an interface between any pair of phases  $i$  and  $j$  to be 1. We expect that it would be possible to generalize our results to more general classes of surface tensions as considered in [13]; to avoid even more complex notation, we refrain from doing so in the present manuscript.

Assumption (A4) is the only truly restrictive condition in our assumptions; in fact, it does not include potentials which at the same time feature quadratic growth at the minima  $\alpha_i$  (i.e. with  $q = 2$  in (A1)) and regularity of class  $C^2$ . The idea behind (A4) is to penalize the mixing of more than two phases in each  $\mathcal{T}_{i,j}$ . As a drawback, it restricts our result to potentials whose second radial derivatives at each  $\alpha_i$  – when viewed as a function of the outgoing angle – have strict local minima in the direction of the minimizing paths  $\gamma_{i,j}$ . Roughly speaking, this corresponds to the presence of a “bump” of the potential along any curve connecting two points belonging to two distinct minimizing paths  $\gamma_{i,j}$ . Nevertheless, as we shall see in Proposition 8 below, there exists a broad class of  $N$ -well potentials – including in particular potentials of class  $C^{1,1}$  with quadratic growth at the minima  $\alpha_i$  – that satisfy all of our assumptions.

Our main result on the quantitative convergence of the vectorial Allen–Cahn equation towards multiphase mean curvature flow reads as follows.

**Theorem 1.** *Let  $d \in \{2, 3\}$ . In the case  $d = 2$ , let  $(\bar{\chi}_1, \dots, \bar{\chi}_N)$  be a classical solution to multiphase mean curvature flow on  $\mathbb{R}^d$  on a time interval  $[0, T]$  in the sense of Definition 5 below; in the case  $d = 3$ , let  $(\bar{\chi}_1, \dots, \bar{\chi}_N)$  be a classical solution to multiphase mean curvature flow of double-bubble type in the sense of [15, Definition 10]. Let  $\xi$  be a corresponding gradient-flow calibration in the sense of Definition 6 below. Suppose that  $W$  is a potential satisfying assumptions (A1)–(A4). For every  $\varepsilon > 0$ , let*

$u_\varepsilon \in L^\infty([0, T]; \dot{H}^1(\mathbb{R}^d; \bar{U}))$  be a bounded weak solution to the vectorial Allen–Cahn equation (1.1).

Assume furthermore that the initial data  $u_\varepsilon(\cdot, 0)$  are well prepared in the sense that

$$E[u_\varepsilon|\xi](0) \leq C\varepsilon,$$

$$\max_{i \in \{1, \dots, N\}} \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon(\cdot, 0)) - \bar{\chi}_i(\cdot, 0)| \min\{\text{dist}(x, \partial \text{supp } \bar{\chi}_i(\cdot, 0)), 1\} dx \leq C\varepsilon,$$

where  $E[u_\varepsilon|\xi]$  denotes the relative entropy given as

$$E[u_\varepsilon|\xi] := \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) + \sum_{i=1}^N \xi_i \cdot \nabla(\psi_i \circ u_\varepsilon) dx. \quad (2.1)$$

Then the solutions  $u_\varepsilon$  to the vectorial Allen–Cahn equation converge towards multiphase mean curvature flow with the rate  $O(\varepsilon^{1/2})$  in the sense that

$$\begin{aligned} \sup_{t \in [0, T]} E[u_\varepsilon|\xi] &\leq C\varepsilon, \\ \sup_{t \in [0, T]} \max_{i \in \{1, \dots, N\}} \|\psi_i(u_\varepsilon(\cdot, t)) - \bar{\chi}_i(\cdot, t)\|_{L^1(\mathbb{R}^d)} &\leq C\varepsilon^{1/2}. \end{aligned}$$

First, let us remark that in the planar case, strong solutions to multiphase mean curvature flow are known to exist prior to the first topology change for quite general initial data [6, 24]. Beyond topology changes, in general the evolution by multiphase mean curvature flow may become unstable and uniqueness of solutions may fail; see e.g. the discussion in [24] or [13]. Thus, quantitative approximation results for multiphase mean curvature flow of the form of our Theorem 1 should not be expected to hold beyond the first topology change. In this sense, our result is optimal.

Second, let us emphasize that by [13] and [15] the existence of a gradient-flow calibration is ensured in the following situations:

- In the planar case  $d = 2$ , gradient-flow calibrations exist as long as a strong solution exists.
- In the three-dimensional case  $d = 3$ , gradient-flow calibrations exist as long as a strong solution of double-bubble type (i.e. in particular with at most three phases meeting at each point) exists.

Note that more generally we expect gradient-flow calibrations to exist as long as a classical solution to multiphase mean curvature flow exists. Since the construction becomes increasingly technical when the geometrical features become more complex, the construction has not yet been carried out in these more general situations. Nevertheless, as soon as gradient-flow calibration becomes available, our results below apply and yield the convergence of the vectorial Allen–Cahn equation to multiphase mean curvature flow in the corresponding setting.

Next, let us remark that we may weaken the assumptions on the sequence of initial data if we are content with lower rates of convergence or merely qualitative convergence statements.

**Remark 2.** As an inspection of the proof of Theorem 1 readily reveals, the assumption of quantitative well-preparedness of the initial data in our theorem can be relaxed, even to a qualitative one. For instance, by merely assuming the qualitative convergences

$$\lim_{\varepsilon \rightarrow 0} E[u_\varepsilon | \xi](0) = 0 \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} \max_{i \in \{1, \dots, N\}} \|\psi_i(u_\varepsilon(\cdot, 0)) - \bar{\chi}_i(\cdot, 0)\|_{L^1(\mathbb{R}^d)} = 0$$

at initial time, from Theorem 12 and Proposition 13 we are able to obtain the qualitative convergence statement

$$\lim_{\varepsilon \rightarrow 0} \sup_{t \in [0, T]} E[u_\varepsilon | \xi] = 0 = \lim_{\varepsilon \rightarrow 0} \sup_{t \in [0, T]} \max_{i \in \{1, \dots, N\}} \|\psi_i(u_\varepsilon(\cdot, t)) - \bar{\chi}_i(\cdot, t)\|_{L^1(\mathbb{R}^d)}.$$

Furthermore, by the definition of the relative entropy, the convergence  $\lim_{\varepsilon \rightarrow 0} E[u_\varepsilon | \xi](0) = 0$  is in fact implied by the convergence of the initial energies  $E[u_\varepsilon](0) \rightarrow E[\bar{\chi}](0) = \frac{1}{2} \sum_i |\nabla \bar{\chi}_i(\cdot, 0)|(\mathbb{R}^d)$  and the convergence of the initial data  $u_\varepsilon(\cdot, 0) \rightarrow \sum_{i=1}^N \alpha_i \bar{\chi}_i(\cdot, 0)$  in  $L^1(\mathbb{R}^d)$ .

To summarize, under the assumptions of Theorem 1 but given now a sequence of solutions  $(u_\varepsilon)_\varepsilon$  to the Allen–Cahn equation (1.1) satisfying only the qualitative convergence properties at initial time

$$\begin{aligned} u_\varepsilon(\cdot, 0) &\xrightarrow{\varepsilon \rightarrow 0} \sum_{i=1}^N \alpha_i \bar{\chi}_i(\cdot, 0) \quad \text{in } L^1(\mathbb{R}^d), \\ E[u_\varepsilon](0) &\xrightarrow{\varepsilon \rightarrow 0} E[\bar{\chi}](0), \end{aligned}$$

the solutions  $u_\varepsilon$  converge to multiphase mean curvature flow in the sense that

$$u_\varepsilon(\cdot, t) \xrightarrow{\varepsilon \rightarrow 0} \sum_{i=1}^N \alpha_i \bar{\chi}_i(\cdot, t) \quad \text{in } L^1(\mathbb{R}^d) \text{ for all } t \in [0, T].$$

As the next proposition (and its rather straightforward proof, proceeding by gluing together one-dimensional Modica–Mortola profiles) shows, well-prepared initial data satisfying the upper bound  $O(\varepsilon)$  for the relative energy actually exist.

**Proposition 3.** *Let assumptions (A1)–(A4) be in place. Let  $d = 2$  and  $(\bar{\chi}_1(\cdot, 0), \dots, \bar{\chi}_N(\cdot, 0))$  be any initial data whose interfaces consist of finitely many  $C^1$  curves that meet at finitely many triple junctions at angles of  $120^\circ$ . Alternatively, let  $d = 3$  and  $(\bar{\chi}_1(\cdot, 0), \dots, \bar{\chi}_N(\cdot, 0))$  be any initial data whose interfaces consist of finitely many  $C^1$  interfaces that meet at finitely many triple lines of class  $C^1$  at angles of  $120^\circ$ .*

Then for any  $\varepsilon > 0$  there exist initial data  $u_\varepsilon(\cdot, 0)$  that are well prepared in the sense that

$$E_\varepsilon[u_\varepsilon|\xi](0) \leq C\varepsilon,$$

$$\max_{i \in \{1, \dots, N\}} \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon(\cdot, 0)) - \bar{\chi}_i(\cdot, 0)| \operatorname{dist}(x, \partial \operatorname{supp} \bar{\chi}_i(\cdot, 0)) \, dx \leq C\varepsilon,$$

where the constant  $C$  depends on the initial data  $(\bar{\chi}_1(\cdot, 0), \dots, \bar{\chi}_N(\cdot, 0))$  and on the potential  $W$ .

Nevertheless, note that in the presence of triple junctions this rate of convergence  $O(\varepsilon)$  for the relative entropy cannot be improved without modifying either the definition of the relative entropy (3.1) or our assumptions (A1)–(A4), as it may be impossible to construct initial data  $u_\varepsilon(\cdot, 0)$  with  $E_\varepsilon[u_\varepsilon|\xi] \ll \varepsilon$ . Let us illustrate the reason for this limitation in the case  $d = 2$ : Suppose that the initial data  $\bar{\chi}(\cdot, 0)$  for the strong solution contain at least one triple junction. By virtue of the term  $\int \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 \, dx$  in the energy and the pointwise nonnegativity of the integrand in the relative entropy, if we were to have  $E[u_\varepsilon|\xi](0) \ll \varepsilon$ , the approximating initial data  $u_\varepsilon(\cdot, 0)$  would have to contain a true mixture of three phases in an  $\varepsilon$ -ball  $B_\varepsilon(y)$  somewhere. At the same time, our assumptions (A1)–(A4) allow the potential  $W$  to be arbitrarily large for a true mixture of three phases (i.e. away from the boundary of the triangle in Figure 1a for a three-well potential as in Definition 17), independently of the functions  $\psi_i$ . If  $W$  is large enough, on  $B_\varepsilon(y)$  the energy density  $\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon)$  then cannot be compensated by the term involving  $\nabla \psi_i(u_\varepsilon)$  in the relative entropy, resulting in a lower bound for the relative entropy of the order of  $\int_{B_\varepsilon(y)} \frac{1}{2\varepsilon} W(u_\varepsilon) \, dx \geq c\varepsilon^{-1} \times \varepsilon^2 = c\varepsilon$ . This limits the overall convergence rate for our method to  $O(\varepsilon^{1/2})$  when measured e.g. in the  $L^1$  norm. We expect this to be a limitation of our method, caused by an insufficient control of the precise dynamics of the diffuse-interface model at triple junctions by the relative entropy  $E[u_\varepsilon|\xi]$ ; for suitably prepared initial data, we would anticipate a convergence rate  $O(\varepsilon)$ . Whether such an improved convergence rate can be deduced by a more refined relative entropy approach is an open question.

Observe that assumptions (A1) and (A2) are indeed sufficient to deduce global existence of bounded solutions to the Allen–Cahn equation (1.1), starting from any measurable initial data taking values in  $\bar{U}$ .

**Remark 4.** Let  $W$  be any potential of class  $C_{\text{loc}}^{1,1}$  satisfying our assumption (A2). Given any measurable initial data  $u_\varepsilon(\cdot, 0)$  taking values in  $\bar{U}$ , for any  $T > 0$  there exists a unique bounded weak solution  $u_\varepsilon$  to the Allen–Cahn equation (1.1) on the time interval  $[0, T]$ . To see this, one may first show existence of a weak solution for a slightly modified PDE obtained by replacing  $\partial_u W$  outside  $\bar{U}$  by a Lipschitz extension. For this modified PDE, existence of a weak solution can be shown in a standard way. A comparison argument (using (A2) and in particular the convexity of  $\bar{U}$ ) then ensures that the weak solution to this modified equation may only take values in  $\bar{U}$ , proving both that it is bounded and that it actually solves the original equation. Uniqueness is shown via the standard argument



of a Grönwall-type estimate for the squared  $L^2(\mathbb{R}^d)$  norm of the difference between two solutions.

We next recall the definition of strong solutions to multiphase mean curvature flow in the case of two dimensions. For intuitive but technical-to-state geometric notions, we shall refer to the precise definitions in [13] for the planar case. In the three-dimensional case, a similar notion of strong solutions for double bubbles can be found in [15].

**Definition 5** (Strong solution for multiphase mean curvature flow). Let  $d = 2$ , let  $P \geq 2$  be an integer, and let  $T > 0$  be a finite time horizon. Let  $\bar{\chi}_0 = (\bar{\chi}_1^0, \dots, \bar{\chi}_P^0)$  be an initial regular partition of  $\mathbb{R}^2$  with finite interface energy in the sense of [13, Definition 14].

A measurable map

$$\bar{\chi} = (\bar{\chi}_1, \dots, \bar{\chi}_P): \mathbb{R}^d \times [0, T] \rightarrow \{0, 1\}^P$$

is called a *strong solution for multiphase mean curvature flow with initial data  $\bar{\chi}_0$*  if it satisfies the following conditions:

- (i) (Smoothly evolving regular partition with finite interface energy). Denote by  $I_{i,j} := \text{supp } \bar{\chi}_i \cap \text{supp } \bar{\chi}_j$  for  $i \neq j$  the interface between phases  $i$  and  $j$ . The map  $\bar{\chi}$  is a smoothly evolving regular partition of  $\mathbb{R}^d \times [0, T]$  and  $\mathcal{I} := \bigcup_{i,j \in \{1, \dots, P\}, i \neq j} I_{i,j}$  is a smoothly evolving regular network of interfaces in  $\mathbb{R}^d \times [0, T]$  in the sense of [13, Definition 15]. In particular, for every  $t \in [0, T]$ ,  $\bar{\chi}(\cdot, t)$  is a regular partition of  $\mathbb{R}^d$  and  $\bigcup_{i \neq j} I_{i,j}(t)$  is a regular network of interfaces in  $\mathbb{R}^d$  in the sense of [13, Definition 14] such that

$$\sup_{t \in [0, T]} E[\bar{\chi}(\cdot, t)] = \sup_{t \in [0, T]} \sum_{\substack{i,j=1 \\ i < j}}^P \int_{I_{i,j}(t)} 1 \, dS < \infty.$$

- (ii) (Evolution by mean curvature). For  $i, j = 1, \dots, P$  with  $i \neq j$  and  $(x, t) \in I_{i,j}$ , let  $\bar{V}_{i,j}(x, t)$  denote the normal speed of the interface at the point  $x \in I_{i,j}(t)$ . Denoting by  $\bar{H}_{i,j}(x, t)$  and  $\bar{n}_{i,j}(x, t)$  the mean curvature vector and the normal vector of  $I_{i,j}(t)$  at  $x \in I_{i,j}(t)$ , the interfaces  $I_{i,j}$  evolve by mean curvature in the sense

$$\bar{V}_{i,j}(x, t) \bar{n}_{i,j}(x, t) = \bar{H}_{i,j}(x, t) \quad \text{for all } t \in [0, T], x \in I_{i,j}(t). \quad (2.2)$$

- (iii) (Initial conditions). We have  $\bar{\chi}_i(x, 0) = \bar{\chi}_i^0(x)$  for all points  $x \in \mathbb{R}^d$  and each phase  $i \in \{1, \dots, P\}$ .

Our main results centrally rely on the concept of gradient-flow calibrations introduced in [13], whose definition we next recall.

**Definition 6.** Let  $d \geq 2$ . Let  $(\bar{\chi}_1, \dots, \bar{\chi}_N)$  be a smoothly evolving partition of  $\mathbb{R}^d$  on a time interval  $[0, T]$ . Denote by  $I_{i,j} := \text{supp } \bar{\chi}_i \cap \text{supp } \bar{\chi}_j$ ,  $1 \leq i, j \leq N$ ,  $i \neq j$ , the corresponding interfaces. We say that a collection of  $C^{1,1}$  vector fields  $\xi_i: \mathbb{R}^d \times [0, T] \rightarrow$

$\mathbb{R}^d$ ,  $1 \leq i \leq N$ , and  $B: \mathbb{R}^d \times [0, T) \rightarrow \mathbb{R}^d$  is a gradient-flow calibration if the following conditions are satisfied:

$$\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j} + (\nabla B)^\top \xi_{i,j} = O(\text{dist}(\cdot, I_{i,j})), \quad (2.3a)$$

$$\frac{1}{2} \xi_{i,j} \cdot (\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j}) = O(\text{dist}^2(\cdot, I_{i,j})), \quad (2.3b)$$

$$(B \cdot \xi_{i,j}) \xi_{i,j} + (\nabla \cdot \xi_{i,j}) \xi_{i,j} = O(\text{dist}(\cdot, I_{i,j})), \quad (2.3c)$$

$$\nabla B : \xi_{i,j} \otimes \xi_{i,j} = O(\text{dist}(\cdot, I_{i,j})), \quad (2.3d)$$

$$\nabla B : (\xi_{i,j}^\perp \otimes \xi_{i,j} + \xi_{i,j} \otimes \xi_{i,j}^\perp) = O(\text{dist}(\cdot, I_{i,j})), \quad (2.3e)$$

$$1 - C_{\text{len}} \text{dist}^2(\cdot, I_{i,j}) \leq |\xi_{i,j}|^2 \leq 1 - c_{\text{len}} \min\{\text{dist}^2(\cdot, I_{i,j}), 1\}, \quad (2.3f)$$

$$\xi_{i,j} = \vec{n}_{i,j} \quad \text{on } I_{i,j}, \quad (2.3g)$$

$$|\sqrt{3} \xi_i| \leq 1 \quad \text{and} \quad \sum_{i=1}^N \xi_i = 0, \quad (2.3h)$$

$$|\xi_{i,j}|^2 + (4 - \delta_{\text{cal}}) \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N |\sqrt{3} \xi_{i,j} \cdot \xi_k|^2 \leq 1, \quad (2.3i)$$

for some constants  $C_{\text{len}} > 0$ ,  $c_{\text{len}} \in (0, 1)$ , an arbitrarily small  $\delta_{\text{cal}} > 0$  and any distinct  $i, j \in \{1, \dots, N\}$ .

Moreover, we call a family of  $C^{1,1}$  functions  $\vartheta_i$  a family of evolving distance weights if they satisfy

$$\vartheta_i(\cdot, t) \leq -c \min\{\text{dist}(\cdot, I_{i,j}(t)), 1\} \quad \text{in } \{\bar{\chi}_i(\cdot, t) = 1\}, \quad (2.4a)$$

$$\vartheta_i(\cdot, t) \geq c \min\{\text{dist}(\cdot, I_{i,j}(t)), 1\} \quad \text{outside } \{\bar{\chi}_i(\cdot, t) = 1\}, \quad (2.4b)$$

$$|\vartheta_i(\cdot, t)| \leq C \min\{\text{dist}(\cdot, I_{i,j}(t)), 1\} \quad \text{globally}, \quad (2.4c)$$

and

$$|\partial_t \vartheta_i + B \cdot \nabla \vartheta_i| \leq C |\vartheta_i|. \quad (2.5)$$

Note that the existence of a calibration for a given smoothly evolving partition entails that the partition must evolve by multiphase mean curvature flow (i.e. the partition must be a strong solution to multiphase mean curvature flow). In fact, conditions (2.3a), (2.3c), (2.3g), and (2.3f) are sufficient to deduce property (2.2). Observe that condition (2.3i) is not stated in [13]; however, it follows from the construction of the gradient-flow calibration provided in [13] (for more details see Appendix A). For the three-dimensional case of double bubbles, we refer the reader to [15].

For many geometries,  $(\bar{\chi}_1, \dots, \bar{\chi}_N)$  being a strong solution to multiphase mean curvature flow is also sufficient to construct a gradient-flow calibration.

**Theorem 7** (Existence of gradient-flow calibrations, [13, Theorem 6] and [15, Theorem 1]). *Let  $d \in \{2, 3\}$  and let  $\bar{\chi}_0$  be a regular partition of  $\mathbb{R}^d$  with finite surface energy;*

for  $d = 3$ , assume furthermore that the partition corresponds to a double-bubble-type geometry. Let  $\bar{\chi}$  be a strong solution to multiphase mean curvature flow on the time interval  $[0, T]$  in the sense of Definition 5 (for  $d = 2$ ) resp. in the sense of [15, Definition 10] (for  $d = 3$ ). Then for any  $\delta_{\text{cal}} > 0$  and any  $c_{\text{len}} \geq 1$  there exists a gradient-flow calibration in the sense of Definition 6 up to time  $T$ . Furthermore, there also exists a family of evolving distance weights.

Finally, we conclude this section by showing that the class of potentials  $W$  satisfying assumptions (A1)–(A4) is indeed sufficiently broad. In fact, given

- a prescribed set of  $N$  minima  $\alpha_i \in \mathbb{R}^{N-1}$ ,  $1 \leq i \leq N$ ,
- a prescribed set of nonintersecting minimal paths  $\gamma_{i,j}$ ,  $1 \leq i < j \leq N$ , that meet at the  $\alpha_i$  at positive angles, and
- a potential  $\tilde{W}: \bigcup_{i,j:i < j} \gamma_{i,j} \rightarrow [0, \infty)$  defined on the minimal paths  $\gamma_{i,j}$  and subject to (A1) and (A3), i.e. in particular with  $\int_{\gamma_{i,j}} \sqrt{2\tilde{W}(u)} \, d\gamma(u) = 1$ ,

it is always possible to extend the potential  $\tilde{W}$  to a potential  $W: \mathbb{R}^{N-1} \rightarrow [0, \infty)$  that satisfies condition (A4). More precisely, to satisfy (A4) it is sufficient to require  $W(u) \geq (1 + M|u - \alpha_i|^{-4} \text{dist}(u, \gamma)^4)W(P_\gamma u)$  in some neighborhood  $\mathcal{U}_i$  of  $\alpha_i$  (with  $P_\gamma$  denoting the projection onto the nearest point among all paths  $\gamma := \bigcup_{j,k} \gamma_{j,k}$ ) as well as  $W(u) \geq M \text{dist}(u, \bigcup_{i < j} \gamma_{i,j})^2$  in  $\mathbb{R}^{N-1} \setminus \bigcup_i \mathcal{U}_i$ . Here,  $M$  is a constant depending only on  $\tilde{W}$ , the paths  $\gamma_{i,j}$ , and the neighborhoods  $\mathcal{U}_i$ .

For the sake of simplicity, we limit ourselves in our rigorous statement to the study of potentials defined on a simplex  $\Delta^{N-1}$ ; however, it is not too difficult to see that our construction would generalize to the aforementioned situation.

**Proposition 8.** *Let  $N \geq 3$ . Let  $\Delta^{N-1}$  be an  $(N - 1)$ -simplex with edges of unit length in  $\mathbb{R}^{N-1}$ . Let  $W: \Delta^{N-1} \rightarrow [0, \infty)$  be a strongly coercive symmetric  $N$ -well potential on the simplex  $\Delta^{N-1}$  in the sense of Definition 17 below. Then assumptions (A1)–(A4) (see Section 2) are satisfied. In particular, (A4) holds true for the set of functions  $\psi_i: \Delta \rightarrow [0, 1]$ ,  $1 \leq i \leq N$ , provided by Construction 19 below.*

### 3. Strategy of the proof

The key idea for our proof is the notion of relative entropy (or, more accurately, relative energy) given by

$$\begin{aligned} E[u_\varepsilon | \xi] &:= E[u_\varepsilon] + \sum_{i=1}^N \int_{\mathbb{R}^d} \xi_i \cdot \nabla(\psi_i \circ u_\varepsilon) \, dx \\ &= \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) + \sum_{i=1}^N \xi_i \cdot \nabla(\psi_i \circ u_\varepsilon) \, dx. \end{aligned} \quad (3.1)$$

The form of the ansatz for the relative entropy is inspired by two earlier approaches:

- The concept of gradient-flow calibrations introduced in [13] by the first author, Hensel, Laux, and Simon to derive weak–strong uniqueness and stability results for distributional solutions to multiphase mean curvature flow. Gradient-flow calibrations provide a lower bound of the form  $-\sum_i \int \xi_i \cdot d\nabla\chi_i$  on the interface energy functional, i.e.  $\frac{1}{2} \sum_i \int 1 \, d|\nabla\chi_i|$ , thereby facilitating a relative entropy approach to weak–strong uniqueness principles for multiphase mean curvature flow. We emphasize that gradient-flow calibrations are specifically designed to handle the (singular) geometries at triple junctions in the strong solution. We refer to [12, 19] for earlier uses of relative entropy techniques for weak–strong uniqueness for geometric evolution problems with smooth geometries (in the strong solution) (see also [16] for a further development of the relative entropy argument in order to incorporate the constant  $90^\circ$  contact angle condition).
- The relative entropy approach to the sharp-interface limit of the scalar Allen–Cahn equation by the first author, Laux, and Simon [14], relying on the Modica–Mortola trick to obtain a lower bound of the form  $\int \xi \cdot \nabla\psi(u_\varepsilon) \, dx$  for the Ginzburg–Landau energy  $E[u_\varepsilon] = \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \, dx$  (see also [17] for an adaptation of this approach in order to encode the constant contact angle condition and [21] for a subsequent application of the relative energy method to a problem in the context of liquid crystals).

The two key steps towards establishing our main results are as follows:

- Establishing a number of coercivity properties of the relative entropy  $E[u_\varepsilon|\xi]$ , including for example

$$E[u_\varepsilon|\xi] \geq c \int \min\left\{\text{dist}^2\left(\cdot, \bigcup_{i \neq j} I_{i,j}\right), 1\right\} \left(\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon)\right) \, dx. \quad (3.2)$$

- Deriving a Grönwall-type estimate for the time evolution of the relative energy of the type

$$\partial_t E[u_\varepsilon|\xi] \leq CE[u_\varepsilon|\xi].$$

We shall illustrate this strategy by stating the main intermediate results in the present section below.

As it is central for our strategy, let us first give the main argument for the coercivity of the relative entropy (3.1) (despite it being slightly technical). It makes use of the following elementary lemma.

**Lemma 9.** *Let  $\xi_i$ ,  $1 \leq i \leq N$ , be vector fields of class  $C^1$  satisfying  $\sum_{i=1}^N \xi_i = 0$ ; suppose that at any point  $(x, t) \in \mathbb{R}^d \times [0, T]$  at most three of the  $\xi_i$  do not vanish. Let  $\psi_i: \mathbb{R}^{N-1} \rightarrow [0, 1]$ ,  $1 \leq i \leq N$ , be functions as in assumption (A4). In particular, set  $\psi_0 := 1 - \sum_{i=1}^N \psi_i$ . Let  $u_\varepsilon \in L^\infty([0, T]; \dot{H}^1(\mathbb{R}^d))$ . Defining  $\psi_{i,j} := \psi_j - \psi_i$  and  $\xi_{i,j} := \xi_i - \xi_j$ , we have for*

any distinct  $i, j, k \in \{1, \dots, N\}$ ,

$$\sum_{\ell=1}^N \xi_{\ell} \otimes \nabla(\psi_{\ell} \circ u_{\varepsilon}) = -\frac{1}{2} \xi_{i,j} \otimes \nabla(\psi_{i,j} \circ u_{\varepsilon}) + \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2} \xi_k \otimes \nabla(\psi_0 \circ u_{\varepsilon}) \quad (3.3)$$

almost everywhere in  $\{u_{\varepsilon} \in \mathcal{T}_{i,j}\}$ , as well as

$$\sum_{\ell=1}^N \nabla \xi_{\ell} \otimes \nabla(\psi_{\ell} \circ u_{\varepsilon}) = -\frac{1}{2} \nabla \xi_{i,j} \otimes \nabla(\psi_{i,j} \circ u_{\varepsilon}) + \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2} \nabla \xi_k \otimes \nabla(\psi_0 \circ u_{\varepsilon}), \quad (3.4)$$

$$\sum_{\ell=1}^N \nabla \xi_{\ell} \otimes \partial_u \psi_{\ell}(u_{\varepsilon}) = -\frac{1}{2} \nabla \xi_{i,j} \otimes \partial_u \psi_{i,j}(u_{\varepsilon}) + \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2} \nabla \xi_k \otimes \partial_u \psi_0(u_{\varepsilon}) \quad (3.5)$$

almost everywhere in  $\{u_{\varepsilon} \in \mathcal{T}_{i,j}\}$ .

*Proof.* By adding zeros, using the definitions  $\xi_{i,j} := \xi_i - \xi_j$  and  $\psi_{i,j} := \psi_j - \psi_i$ , we obtain

$$\begin{aligned} \sum_{\ell=1}^N \xi_{\ell} \cdot \nabla(\psi_{\ell} \circ u_{\varepsilon}) &= -\frac{1}{2} \xi_{i,j} \cdot \nabla(\psi_{i,j} \circ u_{\varepsilon}) + \frac{1}{2} \xi_i \cdot \nabla((\psi_i + \psi_j) \circ u_{\varepsilon}) \\ &\quad + \frac{1}{2} \xi_j \cdot \nabla((\psi_i + \psi_j) \circ u_{\varepsilon}) + \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \xi_k \cdot \nabla(\psi_k \circ u_{\varepsilon}), \end{aligned}$$

almost everywhere in  $\mathbb{R}^d \times (0, T)$ . Equation (3.3) now follows by exploiting that  $\partial_u \psi_k = 0$  on  $\mathcal{T}_{i,j}$  for  $k \notin \{i, j\}$ , inserting the definition of  $\psi_0$ , and using  $\sum_{\ell=1}^N \xi_{\ell} = 0$ . The proof of the other properties is analogous. ■

With the previous lemma and our assumptions (A1)–(A4), it becomes rather straightforward to establish coercivity of our relative energy: Observe that we may compute for  $(x, t)$  with  $u_{\varepsilon}(x, t) \in \mathcal{T}_{i,j}$ ,

$$\begin{aligned} &\frac{\varepsilon}{2} |\nabla u_{\varepsilon}|^2 + \frac{1}{\varepsilon} W(u_{\varepsilon}) + \sum_{\ell=1}^N \xi_{\ell} \cdot \nabla(\psi_{\ell} \circ u_{\varepsilon}) \\ &= \frac{\varepsilon}{2} |\nabla u_{\varepsilon}|^2 + \frac{1}{\varepsilon} W(u_{\varepsilon}) \\ &\quad - \left( \frac{1}{2} \partial_u \psi_{i,j}(u_{\varepsilon}) \otimes \xi_{i,j} - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_{\varepsilon}) \otimes \sqrt{3} \xi_k \right) : \nabla u_{\varepsilon} \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \left| \sqrt{\varepsilon} \nabla u_\varepsilon - \frac{1}{\sqrt{\varepsilon}} \left( \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \xi_{i,j} - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \otimes \sqrt{3} \xi_k \right) \right|^2 \\
 &+ \frac{1}{2\varepsilon} \left[ 2W(u_\varepsilon) - \left| \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \xi_{i,j} - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \otimes \sqrt{3} \xi_k \right|^2 \right] \quad (3.6)
 \end{aligned}$$

due to the fact that  $\psi_k \equiv 0$  on  $\mathcal{T}_{i,j}$  for any  $k \in \{1, \dots, N\} \setminus \{i, j\}$ . This will be the starting point to prove the coercivity properties satisfied by the relative energy functional (3.1); note in particular that

$$\begin{aligned}
 &\left| \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \xi_{i,j} - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \otimes \sqrt{3} \xi_k \right|^2 \\
 &= |\xi_{i,j}|^2 \left| \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \right|^2 + \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N |\sqrt{3} \xi_k|^2 \left| \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \right|^2 \\
 &- \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2} (\xi_{i,j} \cdot \xi_k) \partial_u \psi_{i,j}(u_\varepsilon) \cdot \partial_u \psi_0(u_\varepsilon).
 \end{aligned}$$

Using our assumption (A4) and the properties of the gradient-flow calibration  $|\xi_i| \leq \frac{1}{\sqrt{3}}$ ,  $|\xi_{i,j}| \leq 1$ , and (2.3i), this establishes a first coercivity bound like (3.2). Going substantially beyond this simple estimate, we shall see that in fact we have the following coercivity properties.

**Proposition 10.** *Let  $W$  and  $\psi_i$  be functions subject to assumption (A4). Let  $\xi_i$ ,  $1 \leq i \leq N$ , be any collection of  $C^1$  vector fields satisfying  $\sum_{i=1}^N \xi_i = 0$ ,  $|\sqrt{3} \xi_i| \leq 1$  for all  $i$ , as well as, with the notation  $\xi_{i,j} := \xi_i - \xi_j$ ,*

$$|\xi_{i,j}|^2 + (4 - \delta_{\text{cal}}) \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N |\sqrt{3} \xi_{i,j} \cdot \xi_k|^2 \leq 1 \quad (3.7)$$

for some arbitrarily small  $\delta_{\text{cal}} > 0$ . Furthermore, suppose that at each point at most three of the vector fields  $\xi_i$  do not vanish. For any function  $u_\varepsilon \in \dot{H}^1(\mathbb{R}^d; \bar{U})$  with  $E[u_\varepsilon] < \infty$ , we then have the estimates

$$\int_{\mathbb{R}^d} \left( \sqrt{\varepsilon} |\nabla u_\varepsilon| - \frac{1}{\sqrt{\varepsilon}} \sqrt{2W(u_\varepsilon)} \right)^2 dx \leq CE[u_\varepsilon] |\xi|, \quad (3.8a)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left| \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right|^2 |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \leq CE[u_\varepsilon] |\xi|, \quad (3.8b)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \min\{\text{dist}^2(x, I_{i,j}), 1\} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \leq CE[u_\varepsilon|\xi], \quad (3.8c)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \min\{\text{dist}^2(x, I_{i,j}), 1\} \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \leq CE[u_\varepsilon|\xi], \quad (3.8d)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \varepsilon |(\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \nabla u_\varepsilon^\top|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \leq CE[u_\varepsilon|\xi]. \quad (3.8e)$$

These coercivity estimates will be derived as a consequence of the computation (3.6) and the following coercivity properties.

**Proposition 11.** *Let  $W$  and  $\psi_i$  be functions subject to assumption (A4). Let  $\xi_i$ ,  $1 \leq i \leq N$ , be as in Proposition 10. For any function  $u_\varepsilon \in \dot{H}^1(\mathbb{R}^d; \bar{U})$  with  $E[u_\varepsilon] < \infty$ , we then have the estimates*

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_0(u_\varepsilon)|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \leq CE[u_\varepsilon|\xi], \quad (3.9a)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_{i,j}(u_\varepsilon) \cdot \partial_u \psi_0(u_\varepsilon)| \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \leq CE[u_\varepsilon|\xi], \quad (3.9b)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} |\nabla(\psi_0 \circ u_\varepsilon)| \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \leq CE[u_\varepsilon|\xi]. \quad (3.9c)$$

To introduce a proxy at the level of the Allen–Cahn equation for the limiting mean curvature (or, more precisely, a quantity  $H_\varepsilon$  such that  $|H_\varepsilon|^2$  is a proxy for the dissipation in mean curvature flow), we introduce the abbreviation

$$H_\varepsilon := -\varepsilon \left( \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \right) \cdot \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|}. \quad (3.10)$$

The key step in our proof is to establish the following estimate for the relative energy using a Grönwall-type argument.

**Theorem 12** (Relative energy inequality). *Let  $\bar{\chi} = (\bar{\chi}_1, \dots, \bar{\chi}_N)$  be a smoothly evolving partition of  $\mathbb{R}^d$ ; let  $((\xi_i)_i, B)$  be an associated gradient-flow calibration in the sense of Definition 6. Let  $W$  be a potential, subject to assumptions (A1)–(A4). Let  $u_\varepsilon$  be a bounded solution to the vector-valued Allen–Cahn equation (1.1) with initial data  $u_\varepsilon(\cdot, 0) \in$*

$\dot{H}^1(\mathbb{R}^d; \bar{U})$  with finite energy  $E[u_\varepsilon(\cdot, 0)] < \infty$ . Then for any  $t \in [0, T]$  the estimate

$$\begin{aligned} \frac{d}{dt} E[u_\varepsilon|\xi] &+ \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} |\mathbf{H}_\varepsilon - \varepsilon(B \cdot \xi_{i,j}) \xi_{i,j} |\nabla u_\varepsilon||^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &+ \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left( \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 - |\mathbf{H}_\varepsilon|^2 \right) dx \\ &+ \int_{\mathbb{R}^d} \frac{1}{4\varepsilon} \left| \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) + \sum_{i=1}^N (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \right|^2 dx \\ &\leq C(d, \bar{\chi}) E[u_\varepsilon|\xi] \end{aligned}$$

holds true, with  $\mathbf{H}_\varepsilon$  as defined in (3.10) and  $E[u_\varepsilon|\xi]$  as defined in (2.1).

Building on the previous estimate and the coercivity properties of the relative entropy, we shall show the following error estimate at the level of the indicator functions.

**Proposition 13.** *Let the assumptions of Theorem 12 be in place. In addition, let  $\vartheta_i$  be a family of evolving distance weights as defined in Definition 6. We then have for all  $i \in \{1, \dots, N\}$ ,*

$$\begin{aligned} &\sup_{t \in [0, T]} \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon) - \bar{\chi}_i| \min\{\text{dist}(\cdot, \partial \text{supp } \bar{\chi}_i(\cdot, t)), 1\} dx \\ &\leq C(d, T, (\bar{\chi}(t))_{t \in [0, T]}) E[u_\varepsilon|\xi](0) \\ &\quad + C(d, T, (\bar{\chi}(t))_{t \in [0, T]}) \\ &\quad \times \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon(\cdot, 0)) - \bar{\chi}_i(\cdot, 0)| \min\{\text{dist}(\cdot, \partial \text{supp } \bar{\chi}_i(\cdot, 0)), 1\} dx. \end{aligned}$$

The proof of Theorem 12 crucially relies on the coercivity properties of Propositions 11 and 10 and the following simplification of the evolution equation for the relative entropy.

**Lemma 14.** *Let  $W$  be a potential of class  $C_{\text{loc}}^{1,1}(\mathbb{R}^{N-1})$  subject to assumptions (A1)–(A4). Let  $u_\varepsilon$  be a solution to the vector-valued Allen–Cahn equation (1.1) with initial data  $u_\varepsilon(\cdot, 0) \in \dot{H}^1(\mathbb{R}^d; \bar{U})$  with finite energy  $E[u_\varepsilon(\cdot, 0)] < \infty$ . Let  $(\xi_i, B)$  be a gradient-flow calibration in the sense of Definition 6. The time evolution of the relative energy (3.1) is then given by*

$$\begin{aligned} \frac{d}{dt} E[u_\varepsilon|\xi] &= - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} |\mathbf{H}_\varepsilon - \varepsilon(B \cdot \xi_{i,j}) \xi_{i,j} |\nabla u_\varepsilon||^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &\quad - \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) + \sum_{i=1}^N (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \right|^2 dx \end{aligned}$$



$$\begin{aligned}
 & - \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left( \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 - |\mathbf{H}_\varepsilon|^2 \right) dx \\
 & + \text{Err}_{\text{AllenCahn}} + \text{Err}_{\text{instab}} + \text{Err}_{\text{dt}\xi} + \text{Err}_{\text{MC}\xi} + \text{Err}_{\text{OtherPhases}}, \quad (3.11)
 \end{aligned}$$

where we have abbreviated

$$\begin{aligned}
 \text{Err}_{\text{instab}} := & \int_{\mathbb{R}^d} (\nabla \cdot \mathbf{B}) \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) + \sum_{i=1}^N \xi_i \cdot \nabla(\psi_i \circ u_\varepsilon) \right) dx \\
 & - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \left( \frac{\xi_{i,j}}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) \otimes \left( \frac{\xi_{i,j}}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) \\
 & \quad \times |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \quad (3.12a)
 \end{aligned}$$

and

$$\begin{aligned}
 \text{Err}_{\text{AllenCahn}} := & \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \nabla B : \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \frac{1}{2} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \right. \\
 & \quad \left. - \varepsilon \nabla u_\varepsilon^\top \nabla u_\varepsilon \right) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \quad (3.12b)
 \end{aligned}$$

and

$$\begin{aligned}
 \text{Err}_{\text{dt}\xi} := & \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\partial_t \xi_{i,j} + (\mathbf{B} \cdot \nabla) \xi_{i,j} + (\nabla B)^\top \xi_{i,j}) \\
 & \quad \cdot \left( \frac{\xi_{i,j}}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\
 & - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \xi_{i,j} \cdot (\partial_t \xi_{i,j} + (\mathbf{B} \cdot \nabla) \xi_{i,j}) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx, \quad (3.12c)
 \end{aligned}$$

as well as

$$\begin{aligned}
 \text{Err}_{\text{MC}\xi} := & \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \sum_{i=1}^N (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \right|^2 dx \\
 & - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) \mathbf{B} \cdot \nabla(\psi_i \circ u_\varepsilon) dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |\mathbf{B} \cdot \xi_{i,j}|^2 |\xi_{i,j}|^2 |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : \mathbf{H}_\varepsilon \otimes \mathbf{B} |\nabla u_\varepsilon| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \quad (3.12d)
 \end{aligned}$$

and

$$\begin{aligned}
 \text{Err}_{\text{OtherPhases}} &:= \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\partial_t \xi_k + (B \cdot \nabla) \xi_k + (\nabla B)^\top \xi_k) \cdot \nabla (\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 &- \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \nabla (\psi_0 \circ u_\varepsilon) \otimes \xi_k \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 &- \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \xi_k \otimes \nabla (\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx. \tag{3.12e}
 \end{aligned}$$

## 4. The relative energy argument

### 4.1. Derivation of the Grönwall inequality for the relative entropy

We first show how the evolution estimate for the relative entropy from Lemma 14 and the coercivity properties of our relative entropy together imply a Grönwall-type estimate for the evolution of the relative entropy.

*Proof of Theorem 12.* We proceed by estimating the terms on the right-hand side of equation (3.11) for the time evolution of the relative energy. Note that it will be sufficient to prove

$$\begin{aligned}
 &\text{Err}_{\text{AllenCahn}} + \text{Err}_{\text{instab}} + \text{Err}_{\text{dt}\xi} + \text{Err}_{\text{MC}\xi} + \text{Err}_{\text{OtherPhases}} \\
 &\leq C(\xi(t), B(t), \bar{\delta}) E[u_\varepsilon | \xi] \\
 &\quad + \bar{\delta} \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} |H_\varepsilon - \varepsilon(B \cdot \xi_{i,j}) \xi_{i,j} | \nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 &\quad + \bar{\delta} \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) + \sum_{i=1}^N (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \right|^2 \, dx
 \end{aligned}$$

for any  $\bar{\delta} > 0$ , as then an absorption argument applied to (3.11) (for  $\bar{\delta} < \frac{1}{4}$ ) yields

$$\frac{d}{dt} E[u_\varepsilon | \xi] \leq C(\xi(t), B(t), \bar{\delta}) E[u_\varepsilon | \xi].$$

The Grönwall inequality then implies our conclusion.

*Step 1: Estimates for  $\text{Err}_{\text{OtherPhases}}$ ,  $\text{Err}_{\text{dt}\xi}$ , and  $\text{Err}_{\text{instab}}$ .* We first show that

$$\text{Err}_{\text{instab}} + \text{Err}_{\text{dt}\xi} + \text{Err}_{\text{OtherPhases}} \leq C(\xi(t), B(t)) E[u_\varepsilon | \xi].$$

Indeed, it is immediate by definition (3.12e) and the coercivity property (3.9c) of our relative energy that the inequality

$$\begin{aligned} \text{Err}_{\text{OtherPhases}} &\leq \sum_{\substack{i,j=1 \\ i < j}}^N C(\xi(t), B(t)) \int_{\mathbb{R}^d} |\nabla \psi_0| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\ &\leq C(\xi(t), B(t)) E[u_\varepsilon | \xi] \end{aligned}$$

holds. Using the defining properties (2.3a) and (2.3b) of the calibration  $\xi$  and the coercivity properties (3.8b) and (3.8c) of our relative energy, we likewise deduce from definition (3.12c) that  $\text{Err}_{\text{dt}\xi} \leq C(\xi(t), B(t)) E[u_\varepsilon | \xi]$ , using for instance the estimate

$$\begin{aligned} &\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j} + (\nabla B)^\top \xi_{i,j}) \\ &\quad \cdot \left( \xi_{i,j} - \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\ &\leq \sum_{\substack{i,j=1 \\ i < j}}^N \left( \int_{\mathbb{R}^d} |\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j} + (\nabla B)^\top \xi_{i,j}|^2 |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \right)^{1/2} \\ &\quad \times \left( \int_{\mathbb{R}^d} \left| \xi_{i,j} - \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right|^2 |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \right)^{1/2} \\ &\stackrel{(2.3a), (3.8b), (3.8c)}{\leq} C(\xi(t), B(t)) E[u_\varepsilon | \xi]. \end{aligned}$$

Similarly, recalling definitions (3.12a) and (3.8b), as well as (2.1), we immediately get  $\text{Err}_{\text{instab}} \leq C(\xi(t), B(t)) E[u_\varepsilon | \xi]$ . It therefore only remains to estimate  $\text{Err}_{\text{AllenCahn}}$  and  $\text{Err}_{\text{MC}\xi}$ .

*Step 2: Estimate for  $\text{Err}_{\text{AllenCahn}}$ .* By adding zeros, we may rewrite

$$\begin{aligned} \text{Err}_{\text{AllenCahn}} &= \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \nabla B : \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \otimes \xi_{i,j} \right) \\ &\quad \times \left( \frac{1}{2} |\nabla(\psi_{i,j} \circ u_\varepsilon)| - \varepsilon |\nabla u_\varepsilon|^2 \right) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\ &\quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \nabla B : \xi_{i,j} \otimes \xi_{i,j} \left( \frac{1}{2} |\nabla(\psi_{i,j} \circ u_\varepsilon)| - \varepsilon |\nabla u_\varepsilon|^2 \right) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\ &\quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \nabla B : \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \frac{\nabla u_\varepsilon^\top \nabla u_\varepsilon}{|\nabla u_\varepsilon|^2} \right) \\ &\quad \times \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx. \end{aligned}$$

The first term on the right-hand side can be bounded by  $C(\xi(t), B(t))E[u_\varepsilon|\xi]$  by writing

$$\begin{aligned} & \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \otimes \xi_{i,j} \\ &= \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right) \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} + \xi_{i,j} \otimes \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right) \end{aligned}$$

and using Young's inequality, together with the coercivity estimates (4.3) and (4.4) for our relative energy. The second term on the right-hand side in the above formula can be estimated similarly by exploiting Young's inequality, as well as the gradient-flow calibration property (2.3d) and the coercivity estimates (3.8d) and (4.4).

It remains to bound the third term on the right-hand side. To this aim, we note that for any symmetric matrix  $A$  we have

$$\begin{aligned} \nabla B : A &= \nabla B : (\text{Id} - \xi_{i,j} \otimes \xi_{i,j})A(\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \\ &\quad + (\text{Id} - \xi_{i,j} \otimes \xi_{i,j})(\nabla B + (\nabla B)^\top)\xi_{i,j} \cdot (\xi_{i,j} \cdot A(\text{Id} - \xi_{i,j} \otimes \xi_{i,j})) \\ &\quad + \xi_{i,j} \cdot (\text{Id} - \xi_{i,j} \otimes \xi_{i,j})(\nabla B + (\nabla B)^\top)\xi_{i,j}(\xi_{i,j} \cdot A\xi_{i,j}) \\ &\quad + \xi_{i,j} \cdot \nabla B \xi_{i,j}(\xi_{i,j} \cdot A\xi_{i,j}). \end{aligned}$$

This entails by (2.3e) and  $|\xi_{i,j}(\text{Id} - \xi_{i,j} \otimes \xi_{i,j})| \leq C \text{dist}^2(\cdot, I_{i,j})$  (the latter being a consequence of (2.3f)),

$$\begin{aligned} & \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \nabla B : \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \frac{\nabla u_\varepsilon^\top \nabla u_\varepsilon}{|\nabla u_\varepsilon|^2} \right) \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\ & \leq C(\xi(t), B(t)) \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left( \left| (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right|^2 \varepsilon |\nabla u_\varepsilon|^2 \right. \\ & \quad \left. + \varepsilon |(\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \nabla u_\varepsilon^\top|^2 \right) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\ & \quad + C(\xi(t), B(t)) \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \min\{\text{dist}(x, I_{i,j}), 1\} \\ & \quad \times \left( \left| (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right| \varepsilon |\nabla u_\varepsilon|^2 \right. \\ & \quad \left. + \varepsilon |(\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \nabla u_\varepsilon^\top| |\nabla u_\varepsilon| \right) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\ & \quad + C(\xi(t), B(t)) \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \min\{\text{dist}^2(x, I_{i,j}), 1\} \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\ & \quad + C(\xi(t), B(t)) \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left| \xi_{i,j} \cdot \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right|^2 \varepsilon |\nabla u_\varepsilon|^2 \\ & \quad - \varepsilon |(\xi_{i,j} \cdot \nabla) u_\varepsilon|^2 \Big| \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \end{aligned}$$

$$\begin{aligned}
 &\leq C(\xi(t), B(t)) \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left( \left| (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right|^2 \varepsilon |\nabla u_\varepsilon|^2 \right. \\
 &\quad \left. + \varepsilon |(\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \nabla u_\varepsilon|^2 \right) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &+ C(\xi(t), B(t)) \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left( \left( 1 - \left| \xi_{i,j} \cdot \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right|^2 \right) \varepsilon |\nabla u_\varepsilon|^2 \right. \\
 &\quad \left. + \varepsilon (|\nabla u_\varepsilon|^2 - |(\xi_{i,j} \cdot \nabla) u_\varepsilon|^2) \right) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &+ C(\xi(t), B(t)) \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \min\{\text{dist}^2(x, I_{i,j}), 1\} \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx,
 \end{aligned}$$

where in the last step we have used Young’s inequality. By the coercivity properties (3.8d), (4.3), (4.6), and (3.8e), we conclude that

$$\text{Err}_{\text{AllenCahn}} \leq C(\xi(t), B(t)) E[u_\varepsilon | \xi].$$

*Step 3: Estimate for  $\text{Err}_{\text{MC}\xi}$ .* For the estimate on  $\text{Err}_{\text{MC}\xi}$ , we have to work a bit more. We begin by adding zeros and using (3.5) to obtain

$$\begin{aligned}
 \text{Err}_{\text{MC}\xi} &\leq \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| -\frac{1}{2} (\nabla \cdot \xi_{i,j}) \partial_u \psi_{i,j}(u_\varepsilon) + \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2} (\nabla \cdot \xi_k) \partial_u \psi_0(u_\varepsilon) \right|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &+ \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\nabla \cdot \xi_{i,j}) B \cdot \nabla(\psi_{i,j} \circ u_\varepsilon) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &+ \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |B \cdot \xi_{i,j}|^2 |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &+ \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : H_\varepsilon \otimes B |\nabla u_\varepsilon| \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &- \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\nabla \cdot \xi_k) B \cdot \nabla(\psi_0 \circ u_\varepsilon) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &= \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \left| \frac{1}{\sqrt{\varepsilon}} (\nabla \cdot \xi_{i,j}) \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right. \\
 &\quad \left. + \sqrt{\varepsilon} (B \cdot \xi_{i,j}) \xi_{i,j} \otimes \nabla u_\varepsilon \right|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx
 \end{aligned}$$

$$\begin{aligned}
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2} (\nabla \cdot \xi_k) \partial_u \psi_0(u_\varepsilon) \right|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & - \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{4\varepsilon} (\nabla \cdot \xi_{i,j}) (\nabla \cdot \xi_k) \partial_u \psi_{i,j}(u_\varepsilon) \cdot \partial_u \psi_0(u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\nabla \cdot \xi_{i,j}) (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : B \otimes \nabla(\psi_{i,j} \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{\varepsilon}{2} (1 - |\xi_{i,j}|^2) |B \cdot \xi_{i,j}|^2 |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : H_\varepsilon \otimes B |\nabla u_\varepsilon| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & - \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\nabla \cdot \xi_k) B \cdot \nabla(\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx, \tag{4.1}
 \end{aligned}$$

where in the second step we have also used  $\partial_u \psi_{i,j}(u_\varepsilon) \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} : \nabla u_\varepsilon = |\nabla(\psi_{i,j} \circ u_\varepsilon)|$ .

Now note that the three terms on the right-hand side of (4.1) that involve a  $\partial_u \psi_0(u_\varepsilon)$  or  $\nabla(\psi_0 \circ u_\varepsilon)$  can be directly estimated by  $CE[u_\varepsilon|\xi]$  by relying on the coercivity properties (3.9a), (3.9b), and (3.9c). Similarly, the third-to-last term on the right-hand side is estimated by  $CE[u_\varepsilon|\xi]$  using (2.3f) and (3.8d). This shows

$$\begin{aligned}
 \text{Err}_{\text{MC}\xi} & \leq \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \left| \frac{1}{\sqrt{\varepsilon}} (\nabla \cdot \xi_{i,j}) \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right. \\
 & \quad \left. + \sqrt{\varepsilon} (B \cdot \xi_{i,j}) \xi_{i,j} \otimes \nabla u_\varepsilon \right|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\nabla \cdot \xi_{i,j}) (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : B \otimes \nabla(\psi_{i,j} \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : H_\varepsilon \otimes B |\nabla u_\varepsilon| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & + CE[u_\varepsilon|\xi]. \tag{4.2}
 \end{aligned}$$

By adding zeros, the first term on the right-hand side of (4.2) can be rewritten as

$$\begin{aligned}
 & \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \left| [(B \cdot \xi_{i,j}) \xi_{i,j} + (\nabla \cdot \xi_{i,j}) \xi_{i,j}] \otimes \sqrt{\varepsilon} \nabla u_\varepsilon \right. \\
 & \quad \left. + (\nabla \cdot \xi_{i,j}) \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right. \\
 & \quad \left. \otimes \left( \frac{1}{2\sqrt{\varepsilon}} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \sqrt{\varepsilon} \nabla u_\varepsilon \right) \right. \\
 & \quad \left. + (\nabla \cdot \xi_{i,j}) \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right) \otimes \sqrt{\varepsilon} \nabla u_\varepsilon \right|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \leq \sum_{\substack{i,j=1 \\ i < j}}^N \frac{3}{2} \int_{\mathbb{R}^d} |(B \cdot \xi_{i,j}) \xi_{i,j} + (\nabla \cdot \xi_{i,j}) \xi_{i,j}|^2 \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{i,j=1 \\ i < j}}^N \frac{3}{2} \|\nabla \cdot \xi_{i,j}\|_{L^\infty}^2 \int_{\mathbb{R}^d} \left| \frac{1}{2\sqrt{\varepsilon}} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \sqrt{\varepsilon} \nabla u_\varepsilon \right|^2 \\
 & \quad \quad \quad \times \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{i,j=1 \\ i < j}}^N \frac{3}{2} \|\nabla \cdot \xi_{i,j}\|_{L^\infty}^2 \int_{\mathbb{R}^d} \left| \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right|^2 \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx,
 \end{aligned}$$

where in the second step we have used Young's inequality. Using property (2.3c) of the calibration  $(\xi, B)$  and the coercivity properties (3.8d), (4.5), and (4.3) of the relative energy, we see that the right-hand side is bounded by  $CE[u_\varepsilon|\xi]$ .

It remains to estimate the second and third terms on the right-hand side of (4.2). Adding zero, these terms are seen to be equal to

$$\begin{aligned}
 & \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : B \otimes \nabla u_\varepsilon^\top \cdot \left( \frac{1}{2} (\nabla \cdot \xi_{i,j}) \partial_u \psi_{i,j}(u_\varepsilon) \right. \\
 & \quad \left. - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2} (\nabla \cdot \xi_k) \partial_u \psi_0(u_\varepsilon) \right. \\
 & \quad \left. - \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \right) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\nabla \cdot \xi_k) (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : B \otimes \nabla u_\varepsilon^\top \cdot \partial_u \psi_0(u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \stackrel{(3.5)}{\leq} \sum_{\substack{i,j=1 \\ i < j}}^N C(\bar{\delta}) \|B\|_{L^\infty}^2 \int_{\mathbb{R}^d} \varepsilon |(\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \nabla u_\varepsilon^\top|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx
 \end{aligned}$$

$$\begin{aligned}
 & + \bar{\delta} \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) + \sum_{i=1}^N (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \right|^2 dx \\
 & + \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \frac{1}{2} \|\nabla \cdot \xi_k\|_{L_x^\infty} \|B\|_{L_x^\infty} \int_{\mathbb{R}^d} |\nabla(\psi_0 \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx.
 \end{aligned}$$

Here, in the last step we have used Young's inequality for  $\bar{\delta} > 0$  small enough. Using the coercivity properties (3.8e), (3.8d), and (3.9c) of the relative entropy, we see that the first and last terms on the right-hand side are bounded by  $CE[u_\varepsilon|\xi]$ .

Overall, we have shown

$$\text{Err}_{\text{MC}\xi} \leq C(\bar{\delta})E[u_\varepsilon|\xi] + \bar{\delta} \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) + \sum_{i=1}^N (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \right|^2 dx,$$

which was the only missing ingredient for the proof of the theorem.  $\blacksquare$

In the above estimates, we have used the following additional coercivity properties of the relative entropy. We shall defer their proof to that of the other coercivity properties from Propositions 11 and 10.

**Lemma 15.** *Let  $W$ ,  $\psi_i$ ,  $\psi_{i,j}$ ,  $\xi_i$ , and  $\xi_{i,j}$  be as in Proposition 11. We then have*

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left| \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right|^2 \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \leq CE[u_\varepsilon|\xi], \quad (4.3)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left| \frac{1}{2\sqrt{\varepsilon}} \frac{|\nabla(\psi_{i,j} \circ u_\varepsilon)|}{|\nabla u_\varepsilon|} - \sqrt{\varepsilon} |\nabla u_\varepsilon| \right|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \leq CE[u_\varepsilon|\xi], \quad (4.4)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left| \frac{1}{2\sqrt{\varepsilon}} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \sqrt{\varepsilon} \nabla u_\varepsilon \right|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \leq CE[u_\varepsilon|\xi], \quad (4.5)$$

$$\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left| \xi_{i,j} \otimes \xi_{i,j} - \frac{\nabla u_\varepsilon^\top \nabla u_\varepsilon}{|\nabla u_\varepsilon|^2} \right|^2 \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \leq CE[u_\varepsilon|\xi]. \quad (4.6)$$

## 4.2. Time evolution of the relative energy

We next give the technical computation that provides the estimate for the evolution of the relative entropy stated in Lemma 14. Although in parts technical, it is at the very heart of the proof of our results.



*Proof of Lemma 14.* By direct computations, using definitions (2.1) and (1.1), as well as (an analogue for  $\partial_t \xi_i$  of) relation (3.3), we obtain

$$\begin{aligned}
 \frac{d}{dt} E[u_\varepsilon | \xi] &= - \int_{\mathbb{R}^d} \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \partial_t u_\varepsilon \, dx \\
 &\quad - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \cdot \partial_t u_\varepsilon \, dx \\
 &\quad + \sum_{i=1}^N \int_{\mathbb{R}^d} \partial_t \xi_i \cdot \nabla (\psi_i \circ u_\varepsilon) \, dx \\
 &= - \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 \, dx \\
 &\quad - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \cdot \left( \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \right) \, dx \\
 &\quad - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \partial_t \xi_{i,j} \cdot \nabla (\psi_{i,j} \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 &\quad + \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \partial_t \xi_k \cdot \nabla (\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx.
 \end{aligned}$$

By adding zeros and using (3.3) again, as well as (3.4), we get

$$\begin{aligned}
 \frac{d}{dt} E[u_\varepsilon | \xi] &= - \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 \, dx \\
 &\quad - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \cdot \left( \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \right) \, dx \\
 &\quad - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j} + (\nabla B)^\top \xi_{i,j}) \cdot \nabla (\psi_{i,j} \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 &\quad + \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\partial_t \xi_k + (B \cdot \nabla) \xi_k + (\nabla B)^\top \xi_k) \cdot \nabla (\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 &\quad - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla \xi_i : \nabla (\psi_i \circ u_\varepsilon) \otimes B \, dx - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla B : \xi_i \otimes \nabla (\psi_i \circ u_\varepsilon) \, dx. \quad (4.7)
 \end{aligned}$$

Integrating by parts several times and making use of an approximation argument for  $((\xi_i)_i, B)$ , the last two terms in the equation above can be rewritten as

$$\begin{aligned}
 & - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla \xi_i : \nabla(\psi_i \circ u_\varepsilon) \otimes B \, dx - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla B : \xi_i \otimes \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 &= \sum_{i=1}^N \int_{\mathbb{R}^d} \psi_i(u_\varepsilon) (B \cdot \nabla)(\nabla \cdot \xi_i) \, dx + \sum_{i=1}^N \int_{\mathbb{R}^d} \psi_i(u_\varepsilon) (\nabla B)^\top : \nabla \xi_i \, dx \\
 &\quad - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla B : \xi_i \otimes \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 &= - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) B \cdot \nabla(\psi_i \circ u_\varepsilon) \, dx - \sum_{i=1}^N \int_{\mathbb{R}^d} \psi_i(u_\varepsilon) (\nabla \cdot B) (\nabla \cdot \xi_i) \, dx \\
 &\quad - \sum_{i=1}^N \int_{\mathbb{R}^d} \psi_i(u_\varepsilon) (\xi_i \cdot \nabla) (\nabla \cdot B) \, dx - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla B)^\top : \xi_i \otimes \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 &\quad - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla B : \xi_i \otimes \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 &= \sum_{i=1}^N \int_{\mathbb{R}^d} ((\nabla \cdot B) \xi_i - (\nabla \cdot \xi_i) B) \cdot \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 &\quad - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla B : \nabla(\psi_i \circ u_\varepsilon) \otimes \xi_i \, dx - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla B : \xi_i \otimes \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 &\stackrel{(3.3)}{=} \sum_{i=1}^N \int_{\mathbb{R}^d} ((\nabla \cdot B) \xi_i - (\nabla \cdot \xi_i) B) \cdot \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 &\quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \nabla(\psi_{i,j} \circ u_\varepsilon) \otimes \xi_{i,j} \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &\quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \xi_{i,j} \otimes \nabla(\psi_{i,j} \circ u_\varepsilon) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &\quad - \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \nabla(\psi_0 \circ u_\varepsilon) \otimes \xi_k \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 &\quad - \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \xi_k \otimes \nabla(\psi_0 \circ u_\varepsilon) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx.
 \end{aligned}$$

By adding zero, we obtain

$$\begin{aligned}
 & - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla \xi_i : \nabla(\psi_i \circ u_\varepsilon) \otimes B \, dx - \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla B : \xi_i \otimes \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 & = \sum_{i=1}^N \int_{\mathbb{R}^d} ((\nabla \cdot B) \xi_i - (\nabla \cdot \xi_i) B) \cdot \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 & \quad - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \left( \xi_{i,j} - \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) \otimes \left( \xi_{i,j} - \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) \\
 & \quad \quad \quad \times |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \xi_{i,j} \otimes \xi_{i,j} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad - \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \nabla(\psi_0 \circ u_\varepsilon) \otimes \xi_k \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad - \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \xi_k \otimes \nabla(\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx. \tag{4.8}
 \end{aligned}$$

Using the relation

$$\begin{aligned}
 & - \int_{\mathbb{R}^d} H_\varepsilon \cdot B |\nabla u_\varepsilon| \, dx \stackrel{(3.10)}{=} - \int_{\mathbb{R}^d} \varepsilon \nabla B : \nabla u_\varepsilon^\top \nabla u_\varepsilon \, dx \\
 & \quad \quad \quad + \int_{\mathbb{R}^d} (\nabla \cdot B) \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \, dx,
 \end{aligned}$$

in view of  $\sum_{i,j=1:i < j}^N \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) = 1$  we can rewrite the third term on the right-hand side as

$$\begin{aligned}
 & \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 & = \int_{\mathbb{R}^d} H_\varepsilon \cdot B |\nabla u_\varepsilon| \, dx + \int_{\mathbb{R}^d} (\nabla \cdot B) \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \, dx
 \end{aligned}$$

$$+ \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \nabla B : \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \frac{1}{2} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \right. \\ \left. - \varepsilon \nabla u_\varepsilon^\top \nabla u_\varepsilon \right) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx.$$

Inserting this relation into (4.8) and inserting the resulting equation into (4.7), we obtain by collecting terms and adding and subtracting  $\sum_{i,j=1:i < j}^N \int_{\mathbb{R}^d} \frac{1}{2} \xi_{i,j} \cdot (\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j}) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx$ ,

$$\begin{aligned} \frac{d}{dt} E[u_\varepsilon | \xi] &= - \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 dx + \int_{\mathbb{R}^d} H_\varepsilon \cdot B |\nabla u_\varepsilon| dx \\ &- \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \cdot \left( \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \right) dx \\ &- \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) B \cdot \nabla(\psi_i \circ u_\varepsilon) dx \\ &+ \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j} + (\nabla B)^\top \xi_{i,j}) \\ &\quad \cdot \left( \xi_{i,j} - \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &- \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \xi_{i,j} \cdot (\partial_t \xi_{i,j} + (B \cdot \nabla) \xi_{i,j}) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &+ \int_{\mathbb{R}^d} (\nabla \cdot B) \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) + \sum_{i=1}^N \xi_i \cdot \nabla(\psi_i \circ u_\varepsilon) \right) dx \\ &+ \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} (\partial_t \xi_k + (B \cdot \nabla) \xi_k + (\nabla B)^\top \xi_k) \cdot \nabla(\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &- \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \nabla(\psi_0 \circ u_\varepsilon) \otimes \xi_k \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &- \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \xi_k \otimes \nabla(\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &- \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \nabla B : \left( \xi_{i,j} - \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) \otimes \left( \xi_{i,j} - \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) \\ &\quad \times |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \end{aligned}$$

$$\begin{aligned}
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \nabla B : \left( \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \frac{1}{2} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \right. \\
 & \quad \left. - \varepsilon \nabla u_\varepsilon^\top \nabla u_\varepsilon \right) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx.
 \end{aligned}$$

Lemma 14 follows from this equation using the definitions of the errors (3.12a)–(3.12e) and the next formula (whose derivation relies on  $\sum_{i,j=1:i < j}^N \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) = 1$  and repeated addition of zero):

$$\begin{aligned}
 & - \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 + \int_{\mathbb{R}^d} \mathbf{H}_\varepsilon \cdot B |\nabla u_\varepsilon| \, dx \\
 & \quad - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \cdot \left( \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \right) \, dx \\
 & \quad - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) B \cdot \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 & = - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} |\mathbf{H}_\varepsilon - \varepsilon(B \cdot \xi_{i,j}) \xi_{i,j}|^2 |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad - \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left( \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 - |\mathbf{H}_\varepsilon|^2 \right) \, dx \\
 & \quad - \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 \, dx \\
 & \quad - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \cdot \left( \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \right) \, dx \\
 & \quad - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) B \cdot \nabla(\psi_i \circ u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |B \cdot \xi_{i,j}|^2 |\xi_{i,j}|^2 |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : \mathbf{H}_\varepsilon \otimes B |\nabla u_\varepsilon| \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 & = - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} |\mathbf{H}_\varepsilon - \varepsilon(B \cdot \xi_{i,j}) \xi_{i,j}|^2 |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad - \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) + \sum_{i=1}^N (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \right|^2 \, dx
 \end{aligned}$$

$$\begin{aligned}
 & - \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left( \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 - |H_\varepsilon|^2 \right) dx \\
 & + \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left| \sum_{i=1}^N (\nabla \cdot \xi_i) \partial_u \psi_i(u_\varepsilon) \right|^2 dx \\
 & - \sum_{i=1}^N \int_{\mathbb{R}^d} (\nabla \cdot \xi_i) B \cdot \nabla (\psi_i \circ u_\varepsilon) dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{\varepsilon}{2} |B \cdot \xi_{i,j}|^2 |\xi_{i,j}|^2 |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\
 & + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) : H_\varepsilon \otimes B |\nabla u_\varepsilon| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx. \quad \blacksquare
 \end{aligned}$$

### 4.3. Derivation of the coercivity properties

We next show how our assumption (A4) implies the coercivity properties of our relative entropy.

*Proof of Proposition 11.* To prove (3.9a)–(3.9c), let  $i, j \in \{1, \dots, N\}$ ,  $i \neq j$ ; suppose that  $(x, t)$  is such that  $u_\varepsilon(x, t) \in \mathcal{T}_{i,j}$ . In particular, we then have  $\chi_{\mathcal{T}_{i,j}}(u_\varepsilon) = 1$ .

*Proof of (3.9a) and (3.9b).* Starting from (3.6), expanding the second square, and making use of Young's inequality, the fact that for each  $(x, t)$  there exist at most three indices  $k \in \{1, \dots, N\} \setminus \{i, j\}$  with  $\xi_k(x, t) \neq 0$ , and (3.7), we obtain

$$\begin{aligned}
 & \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) + \sum_{\ell=1}^N \xi_\ell \cdot \nabla (\psi_\ell \circ u_\varepsilon) \\
 & \geq \frac{1}{2} \left| \sqrt{\varepsilon} \nabla u_\varepsilon - \frac{1}{\sqrt{\varepsilon}} \left( \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \xi_{i,j} - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \otimes \sqrt{3} \xi_k \right) \right|^2 \\
 & + \frac{1}{2\varepsilon} \left[ 2W(u_\varepsilon) - \left( |\xi_{i,j}|^2 + (4 - \delta_{\text{cal}}) \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N |\sqrt{3} \xi_{i,j} \cdot \xi_k|^2 \right) \left| \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \right|^2 \right. \\
 & \quad \left. - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \left( |\sqrt{3} \xi_k|^2 + \frac{1}{4 - \delta_{\text{cal}}} \right) \left| \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \right|^2 \right].
 \end{aligned}$$

Then, by adding zeros and using  $|\sqrt{3}\xi_k| \leq 1$ , we obtain

$$\begin{aligned}
 & \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) + \sum_{\ell=1}^N \xi_\ell \cdot \nabla(\psi_\ell \circ u_\varepsilon) \\
 & \geq \frac{1}{2} \left| \sqrt{\varepsilon} \nabla u_\varepsilon - \frac{1}{\sqrt{\varepsilon}} \left( \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \xi_{i,j} - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \otimes \sqrt{3}\xi_k \right) \right|^2 \\
 & \quad + \frac{1}{2\varepsilon} \left[ 2W(u_\varepsilon) - \left| \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \right|^2 - \left( \frac{5}{4} + \delta_{\text{cal}} + \delta_{\text{coer},1} \right) \left| \frac{1}{2} \partial_u \psi_0(u_\varepsilon) \right|^2 \right. \\
 & \quad \quad \left. - \delta_{\text{coer},2} |\partial_u \psi_{i,j}(u_\varepsilon) \cdot \partial_u \psi_0(u_\varepsilon)| \right] \\
 & \quad + \frac{\delta_{\text{coer},1}}{2\varepsilon} \left| \frac{1}{2} \partial_u \psi_0(u_\varepsilon) \right|^2 + \frac{\delta_{\text{coer},2}}{2\varepsilon} |\partial_u \psi_{i,j}(u_\varepsilon) \cdot \partial_u \psi_0(u_\varepsilon)|, \tag{4.9}
 \end{aligned}$$

where  $\delta_{\text{cal}}, \delta_{\text{coer},1}, \delta_{\text{coer},2}, \delta_{\text{coer},3} > 0$  are arbitrarily small constants. Finally, using (A4) and integrating over the set  $\{x : u_\varepsilon(x, t) \in \mathcal{T}_{i,j}\}$ , we can conclude about the validity of (3.9a) and (3.9b).

*Proof of (3.9c).* By adding zero, we can write

$$\begin{aligned}
 & \nabla(\psi_0 \circ u_\varepsilon) = \partial_u \psi_0(u_\varepsilon) \cdot \nabla u_\varepsilon \\
 & = \frac{1}{\sqrt{\varepsilon}} \partial_u \psi_0(u_\varepsilon) \cdot \left[ \sqrt{\varepsilon} \nabla u_\varepsilon - \frac{1}{\sqrt{\varepsilon}} \left( \frac{1}{2} \partial_u \psi_{i,j} \otimes \xi_{i,j} - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \otimes \sqrt{3}\xi_k \right) \right] \\
 & \quad + \frac{1}{2\varepsilon} \partial_u \psi_0(u_\varepsilon) \cdot \partial_u \psi_{i,j}(u_\varepsilon) \xi_{i,j} - \frac{1}{\varepsilon} \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} |\partial_u \psi_0(u_\varepsilon)|^2 \sqrt{3}\xi_k.
 \end{aligned}$$

Then Young's inequality yields

$$\begin{aligned}
 & |\nabla(\psi_0 \circ u_\varepsilon)| \\
 & \leq \frac{1}{2\varepsilon} |\partial_u \psi_0(u_\varepsilon)|^2 \\
 & \quad + \frac{1}{2} \left| \sqrt{\varepsilon} \nabla u_\varepsilon - \frac{1}{\sqrt{\varepsilon}} \left( \frac{1}{2} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \xi_{i,j} - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} \partial_u \psi_0(u_\varepsilon) \otimes \sqrt{3}\xi_k \right) \right|^2 \\
 & \quad + \frac{1}{2\varepsilon} |\partial_u \psi_0(u_\varepsilon) \cdot \partial_u \psi_{i,j}(u_\varepsilon)| |\xi_{i,j}| \\
 & \quad + \frac{1}{\varepsilon} \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \frac{1}{2\sqrt{3}} |\partial_u \psi_0(u_\varepsilon)|^2 |\sqrt{3}\xi_k|,
 \end{aligned}$$

Using Young's inequality and estimate (4.9), we can conclude about the validity of (3.9c).  $\blacksquare$

*Proof of Proposition 10.*

*Proof of (3.8a), (3.8b), and (3.8c).* Using (3.3) and adding zero, we obtain

$$\begin{aligned} E[u_\varepsilon|\xi] &= E[u_\varepsilon] - \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &\quad + \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \xi_k \cdot \nabla(\psi_0 \circ u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &\quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \left( 1 - \xi_{i,j} \cdot \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx. \end{aligned}$$

From assumption (A4), we can deduce

$$\frac{1}{2} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \leq \frac{1}{2} |\partial_u \psi_{i,j}(u_\varepsilon)| |\nabla u_\varepsilon| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \leq \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon).$$

Hence, using the definition of  $E[u_\varepsilon]$ , we have

$$\begin{aligned} E[u_\varepsilon|\xi] &+ \sum_{\substack{i,j,k=1 \\ i < j, k \notin \{i,j\}}}^N \int_{\mathbb{R}^d} \frac{1}{2} \|\xi_k\|_{L^\infty} |\nabla(\psi_0 \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &\geq \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \left( \sqrt{\varepsilon} |\nabla u_\varepsilon| - \frac{\sqrt{2W(u_\varepsilon)}}{\sqrt{\varepsilon}} \right)^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx \\ &\quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \frac{1}{2} \left( 1 - \xi_{i,j} \cdot \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) dx. \end{aligned}$$

Then, noting that

$$\begin{aligned} &\left| \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right|^2 |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\ &\quad \leq 2 \left( 1 - \xi_{i,j} \cdot \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right) |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon), \end{aligned}$$

together with the fact that  $\sum_{i,j=1:i < j}^N \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) = 1$ , we see that both (3.8a) and (3.8b) follow from the preceding two formulas and (3.9c). Furthermore, using

$$\min\{\text{dist}^2(x, I_{i,j}), 1\} \leq C(1 - |\xi_{i,j}|) \leq C \left( 1 - \xi_{i,j} \cdot \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} \right),$$

we obtain (3.8c).



*Proof of (3.8d).* By exploiting (3.3) and by adding zeros, we obtain

$$\begin{aligned}
 & \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \\
 & \leq \left[ \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) + \sum_{\ell=1}^N \xi_\ell \cdot \nabla(\psi_\ell \circ u_\varepsilon) \right] \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \\
 & \quad + \left[ \frac{1}{2} |\nabla(\psi_{i,j} \circ u_\varepsilon)| + \frac{1}{2} |\nabla(\psi_0 \circ u_\varepsilon)| \right] \chi_{\mathcal{I}_{i,j}}(u_\varepsilon). \tag{4.10}
 \end{aligned}$$

As a consequence, we have

$$\begin{aligned}
 & \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \min\{\text{dist}^2(x, I_{i,j}), 1\} \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 & \leq E[u_\varepsilon | \xi] + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \min\{\text{dist}^2(x, I_{i,j}), 1\} |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} |\nabla(\psi_0 \circ u_\varepsilon)| \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \, dx,
 \end{aligned}$$

whence we deduce (3.8d) from (3.9c) and (3.8c).

*Proof of (3.8e).* Expanding the square and using  $1 - c \, \text{dist}^2(\cdot, I_{i,j}) \leq |\xi_{i,j}| \leq \max\{1 - C \, \text{dist}^2(\cdot, I_{i,j}), 0\}$ , we obtain

$$\begin{aligned}
 \varepsilon |(\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \nabla u_\varepsilon^\top|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) & \leq [\varepsilon |\nabla u_\varepsilon|^2 - \varepsilon |(\xi_{i,j} \cdot \nabla) u_\varepsilon|^2] \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \\
 & \quad + C \min\{\text{dist}^2(\cdot, I_{i,j}), 1\} \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon).
 \end{aligned}$$

Then, by adding zeros, we obtain

$$\begin{aligned}
 & \varepsilon |(\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \nabla u_\varepsilon^\top|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \\
 & = \left[ \varepsilon |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} 2W(u_\varepsilon) - \xi_{i,j} \cdot \nabla(\psi_{i,j} \circ u_\varepsilon) + \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \xi_k \cdot \nabla(\psi_0 \circ u_\varepsilon) \right] \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \\
 & \quad - \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^N \xi_k \cdot \nabla(\psi_0 \circ u_\varepsilon) \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) + \left[ \frac{1}{4\varepsilon} |\partial_u \psi_{i,j}(u_\varepsilon)|^2 - \frac{1}{\varepsilon} 2W(u_\varepsilon) \right] \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \\
 & \quad - \left[ \frac{1}{4\varepsilon} |\partial_u \psi_{i,j}(u_\varepsilon)|^2 - \xi_{i,j} \cdot \nabla(\psi_{i,j} \circ u_\varepsilon) + \varepsilon |(\xi_{i,j} \cdot \nabla) u_\varepsilon|^2 \right] \chi_{\mathcal{I}_{i,j}}(u_\varepsilon) \\
 & \quad + C \min\{\text{dist}^2(\cdot, I_{i,j}), 1\} \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{I}_{i,j}}(u_\varepsilon)
 \end{aligned}$$

$$\begin{aligned}
 &\leq 2 \left[ \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) + \sum_{\ell=1}^N \xi_\ell \cdot \nabla(\psi_\ell \circ u_\varepsilon) \right] \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\
 &\quad + \sum_{\substack{k=1 \\ k \neq \{i,j\}}}^N \|\xi_k\|_{L^\infty} |\nabla(\psi_0 \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\
 &\quad + \left[ \frac{1}{4\varepsilon} |\partial_u \psi_{i,j}(u_\varepsilon)|^2 - \frac{1}{\varepsilon} 2W(u_\varepsilon) \right] \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\
 &\quad - \left| \frac{1}{2\sqrt{\varepsilon}} \partial_u \psi_{i,j}(u_\varepsilon) - \sqrt{\varepsilon}(\xi_{i,j} \cdot \nabla) u_\varepsilon \right|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\
 &\quad + C \min\{\text{dist}^2(\cdot, I_{i,j}), 1\} \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon),
 \end{aligned}$$

due to (3.3). Noting that assumption (A4) implies

$$\frac{1}{4} |\partial_u \psi_{i,j}(u_\varepsilon)|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \leq 2W(u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon),$$

the validity of (3.8e) follows from (3.9c) and (3.8d).  $\blacksquare$

We next prove the additional coercivity properties stated in Lemma 15.

*Proof of Lemma 15.*

*Proof of (4.3).* Note that (4.10) yields

$$\begin{aligned}
 &\sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left| \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right|^2 \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 &\leq 8E[u_\varepsilon|\xi] + \sum_{\substack{i,j=1 \\ i < j}}^N \int_{\mathbb{R}^d} \left| \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \xi_{i,j} \right|^2 |\nabla(\psi_{i,j} \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\
 &\quad + \sum_{\substack{i,j=1 \\ i < j}}^N 4 \int_{\mathbb{R}^d} |\nabla(\psi_0 \circ u_\varepsilon)| \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx.
 \end{aligned}$$

Hence, using (3.9c) and (3.8b), we obtain (4.3).

*Proof of (4.4) and (4.5).* First, we compute

$$\begin{aligned}
 &\left| \frac{1}{2\sqrt{\varepsilon}} \frac{|\nabla(\psi_{i,j} \circ u_\varepsilon)|}{|\nabla u_\varepsilon|} - \sqrt{\varepsilon} |\nabla u_\varepsilon| \right|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\
 &= \left[ \frac{1}{\varepsilon} \frac{|\nabla(\psi_{i,j} \circ u_\varepsilon)|^2}{4|\nabla u_\varepsilon|^2} + \varepsilon |\nabla u_\varepsilon|^2 - |\nabla(\psi_{i,j} \circ u_\varepsilon)| \right] \chi_{\mathcal{T}_{i,j}}(u_\varepsilon)
 \end{aligned}$$

and

$$\begin{aligned} & \left| \frac{1}{2\sqrt{\varepsilon}} \partial_u \psi_{i,j}(u_\varepsilon) \otimes \frac{\nabla(\psi_{i,j} \circ u_\varepsilon)}{|\nabla(\psi_{i,j} \circ u_\varepsilon)|} - \sqrt{\varepsilon} \nabla u_\varepsilon \right|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \, dx \\ &= \left[ \frac{1}{4\varepsilon} |\partial_u \psi_{i,j}(u_\varepsilon)|^2 + \varepsilon |\nabla u_\varepsilon|^2 - |\nabla(\psi_{i,j} \circ u_\varepsilon)|^2 \right] \chi_{\mathcal{T}_{i,j}}(u_\varepsilon). \end{aligned}$$

Then, from assumption (A4) we deduce

$$\frac{|\nabla(\psi_{i,j} \circ u_\varepsilon)|^2}{|\nabla u_\varepsilon|^2} \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \leq |\partial_u \psi_{i,j}(u_\varepsilon)|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \leq 8W(u_\varepsilon) \chi_{\mathcal{T}_{i,j}}(u_\varepsilon).$$

Finally, by exploiting (3.3) and by adding zero, one can conclude about the validity of both (4.5) and (4.4), due to (3.9a).

*Proof of (4.6).* Since we have

$$\begin{aligned} & \left| \xi_{i,j} \otimes \xi_{i,j} - \frac{\nabla u_\varepsilon^\top \nabla u_\varepsilon}{|\nabla u_\varepsilon|^2} \right|^2 \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\ & \leq \left[ 2 - 2 \frac{\xi_{i,j} \cdot \nabla u_\varepsilon^\top \nabla u_\varepsilon \cdot \xi_{i,j}}{|\nabla u_\varepsilon|^2} \right] \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\ & \leq 2[\varepsilon |\nabla u_\varepsilon|^2 - \varepsilon |(\xi_{i,j} \cdot \nabla) u_\varepsilon|^2] \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \\ & \leq 2\varepsilon (\text{Id} - \xi_{i,j} \otimes \xi_{i,j}) \cdot \nabla u_\varepsilon^\top \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) + C \min\{\text{dist}^2(\cdot, I_{i,j}), 1\} \varepsilon |\nabla u_\varepsilon|^2 \chi_{\mathcal{T}_{i,j}}(u_\varepsilon) \end{aligned}$$

(where in the last step we have used the estimate  $1 - C \text{dist}^2(\cdot, I_{i,j}) \leq |\xi_{i,j}| \leq \max\{1 - c \text{dist}^2(\cdot, I_{i,j}), 0\}$ ), the bound (4.6) follows from (3.8e) and (3.8d). ■

#### 4.4. Convergence of the phase indicator functions

We now show how to obtain the error estimate at the level of the indicator functions.

*Proof of Proposition 13.* Using (1.1) and the fact that  $\text{supp } \partial_t \bar{\chi}_i \subset \partial \text{supp } \bar{\chi}_i$ , as well as  $\vartheta_i = 0$  on  $\partial \text{supp } \bar{\chi}_i$ , we compute

$$\begin{aligned} & \frac{d}{dt} \int_{\mathbb{R}^d} (\psi_i(u_\varepsilon) - \bar{\chi}_i) \vartheta_i \, dx \\ &= \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \partial_u \psi_i(u_\varepsilon) \cdot \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \vartheta_i \, dx + \int_{\mathbb{R}^d} (\psi_i(u_\varepsilon) - \bar{\chi}_i) \partial_t \vartheta_i \, dx \\ &= \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \partial_u \psi_i(u_\varepsilon) \cdot \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \vartheta_i \, dx + \int_{\mathbb{R}^d} B \cdot \nabla(\psi_i \circ u_\varepsilon) \vartheta_i \, dx \\ & \quad + \int_{\mathbb{R}^d} (\nabla \cdot B)(\psi_i(u_\varepsilon) - \bar{\chi}_i) \vartheta_i \, dx + \int_{\mathbb{R}^d} (\psi_i(u_\varepsilon) - \bar{\chi}_i) (\partial_t \vartheta_i + B \cdot \nabla \vartheta_i) \, dx, \end{aligned} \quad (4.11)$$

where we added a zero and then integrated by parts. Note that we used the fact that  $\vartheta_i = 0$  on  $\partial\{\bar{\chi}_i = 1\}$ .

By (2.5), the last two terms on the right-hand side of (4.11) can be bounded by

$$(\|\nabla \cdot B\|_{L_x^\infty} + C) \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon) - \bar{\chi}_i| |\vartheta_i| dx.$$

As for the second term on the right-hand side of (4.11), we perform the following decomposition

$$\begin{aligned} & \int_{\mathbb{R}^d} B \cdot \nabla(\psi_i \circ u_\varepsilon) \vartheta_i dx \\ &= \sum_{\substack{j,k=1 \\ j < k}}^N \int_{\mathbb{R}^d} (B \cdot \xi_{j,k}) \xi_{j,k} \cdot \nabla(\psi_i \circ u_\varepsilon) \vartheta_i \chi_{\mathcal{T}_{j,k}}(u_\varepsilon) dx \\ &+ \sum_{\substack{j,k=1 \\ j < k}}^N \int_{\mathbb{R}^d} [(\text{Id} - \xi_{j,k} \otimes \xi_{j,k}) B] \cdot \nabla(\psi_i \circ u_\varepsilon) \vartheta_i \chi_{\mathcal{T}_{j,k}}(u_\varepsilon) dx, \end{aligned}$$

whence, by adding a zero and using  $\sum_{j,k=1: j \neq k}^N \chi_{\mathcal{T}_{j,k}} = 1$ ,

$$\begin{aligned} & \int_{\mathbb{R}^d} B \cdot \nabla(\psi_i \circ u_\varepsilon) \vartheta_i dx \\ &= \int_{\mathbb{R}^d} \frac{1}{\varepsilon} H_\varepsilon \cdot \frac{\nabla u_\varepsilon^\top}{|\nabla u_\varepsilon|} \cdot \partial_u \psi_i(u_\varepsilon) dx \\ &+ \sum_{\substack{j,k=1 \\ j < k}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} [\varepsilon (B \cdot \xi_{j,k}) \xi_{j,k} |\nabla u_\varepsilon| - H_\varepsilon] \cdot \frac{\nabla u_\varepsilon^\top}{|\nabla u_\varepsilon|} \cdot \partial_u \psi_i(u_\varepsilon) \vartheta_i \chi_{\mathcal{T}_{j,k}}(u_\varepsilon) dx \\ &+ \sum_{\substack{j,k=1 \\ j < k}}^N \int_{\mathbb{R}^d} [(\text{Id} - \xi_{j,k} \otimes \xi_{j,k}) B] \cdot \nabla(\psi_i \circ u_\varepsilon) \vartheta_i \chi_{\mathcal{T}_{j,k}}(u_\varepsilon) dx. \end{aligned}$$

Note that the last two terms are nonzero only if  $j = i$  or  $k = i$ . Hence, using Young's inequality, the second term can be estimated by

$$\begin{aligned} & \frac{1}{2} \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\varepsilon (B \cdot \xi_{i,k}) \xi_{i,k} |\nabla u_\varepsilon| - H_\varepsilon|^2 \chi_{\mathcal{T}_{i,k}}(u_\varepsilon) dx \\ &+ \frac{1}{2} \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i(u_\varepsilon)|^2 |\vartheta_i|^2 \chi_{\mathcal{T}_{i,k}}(u_\varepsilon) dx. \end{aligned}$$

As for the third one, we obtain using Young’s inequality and exploiting the coercivity property (3.8e),

$$\begin{aligned}
 & \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} [(\text{Id} - \xi_{i,k} \otimes \xi_{i,k})B] \cdot \nabla(\psi_i \circ u_\varepsilon) \vartheta_i \chi_{\mathcal{T}_{i,k}}(u_\varepsilon) \, dx \\
 & \leq \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \varepsilon |[(\text{Id} - \xi_{i,k} \otimes \xi_{i,k})B] \nabla u_\varepsilon^\top|^2 \chi_{\mathcal{T}_{i,k}}(u_\varepsilon) \, dx \\
 & \quad + \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i|^2 |\vartheta_i|^2 \chi_{\mathcal{T}_{i,k}}(u_\varepsilon) \, dx \\
 & \leq CE[u_\varepsilon | \xi] + \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i|^2 |\vartheta_i|^2 \chi_{\mathcal{T}_{i,k}}(u_\varepsilon) \, dx.
 \end{aligned}$$

In summary, we have shown

$$\begin{aligned}
 & \frac{d}{dt} \int_{\mathbb{R}^d} (\psi_i(u_\varepsilon) - \bar{\chi}_i) \vartheta_i \, dx \\
 & \leq C \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon) - \bar{\chi}_i| |\vartheta_i| \, dx + \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i(u_\varepsilon)|^2 |\vartheta_i|^2 \chi_{\mathcal{T}_{i,k}}(u_\varepsilon) \, dx \\
 & \quad + \frac{1}{2} \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\varepsilon(B \cdot \xi_{i,k}) \xi_{i,k} |\nabla u_\varepsilon| - H_\varepsilon|^2 \chi_{\mathcal{T}_{i,k}}(u_\varepsilon) \, dx \\
 & \quad + \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \partial_u \psi_i(u_\varepsilon) \cdot \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \vartheta_i \, dx + \int_{\mathbb{R}^d} \frac{1}{\varepsilon} H_\varepsilon \cdot \frac{\nabla u_\varepsilon^\top}{|\nabla u_\varepsilon|} \cdot \partial_u \psi_i(u_\varepsilon) \, dx.
 \end{aligned}$$

We estimate the two terms in the last line:

$$\begin{aligned}
 & \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \partial_u \psi_i(u_\varepsilon) \cdot \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \vartheta_i \, dx + \int_{\mathbb{R}^d} \frac{1}{\varepsilon} H_\varepsilon \cdot \frac{\nabla u_\varepsilon^\top}{|\nabla u_\varepsilon|} \cdot \partial_u \psi_i(u_\varepsilon) \vartheta_i \, dx \\
 & = \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \left[ \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) + H_\varepsilon \cdot \frac{\nabla u_\varepsilon^\top}{|\nabla u_\varepsilon|} \right] \cdot \partial_u \psi_i(u_\varepsilon) \vartheta_i \, dx \\
 & \leq \frac{1}{2} \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \left[ \left( \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) + H_\varepsilon \cdot \frac{\nabla u_\varepsilon^\top}{|\nabla u_\varepsilon|} \right]^2 \, dx \\
 & \quad + \frac{1}{2} \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i(u_\varepsilon)|^2 |\vartheta_i|^2 \, dx \\
 & \leq \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left( \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 - |H_\varepsilon|^2 \right) \, dx
 \end{aligned}$$

$$\begin{aligned}
 & - \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left[ \text{Id} - \frac{\nabla u_\varepsilon^\top \nabla u_\varepsilon}{|\nabla u_\varepsilon|^2} \right] : \mathbf{H}_\varepsilon \otimes \mathbf{H}_\varepsilon \, dx \\
 & + \frac{1}{2} \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i(u_\varepsilon)|^2 |\vartheta_i|^2 \, dx \\
 & \leq \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left( \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 - |\mathbf{H}_\varepsilon|^2 \right) dx + \frac{1}{2} \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i(u_\varepsilon)|^2 |\vartheta_i|^2 \, dx,
 \end{aligned}$$

where we used the fact that  $[\text{Id} - \frac{\nabla u_\varepsilon^\top \nabla u_\varepsilon}{|\nabla u_\varepsilon|^2}]$  is a positive semidefinite matrix. Since  $|\partial_u \psi_i(u_\varepsilon)| \leq C \sqrt{2W(u_\varepsilon)}$  and  $|\vartheta_i| \leq \min\{\text{dist}^2(\cdot, \partial \text{supp } \bar{\chi}_i), 1\} \leq C \min\{\text{dist}(\cdot, I_{i,k}), 1\}$  (see (2.4)), from (3.8d) it follows that

$$\begin{aligned}
 \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i(u_\varepsilon)|^2 |\vartheta_i|^2 \, dx & = \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\partial_u \psi_i(u_\varepsilon)|^2 |\vartheta_i|^2 \chi_{\mathcal{I}_{i,k}}(u_\varepsilon) \, dx \\
 & \leq C \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} 2W(u_\varepsilon) |\vartheta_i|^2 \chi_{\mathcal{I}_{i,k}}(u_\varepsilon) \, dx \\
 & \leq CE[u_\varepsilon | \xi].
 \end{aligned}$$

Summarizing the previous estimates, we get

$$\begin{aligned}
 \frac{d}{dt} \int_{\mathbb{R}^d} (\psi_i(u_\varepsilon) - \bar{\chi}_i) \vartheta_i \, dx & \leq CE[u_\varepsilon | \xi] + C \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon) - \bar{\chi}_i| |\vartheta_i| \, dx \\
 & + \frac{1}{2} \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\varepsilon(B \cdot \xi_{i,k}) \xi_{i,k} |\nabla u_\varepsilon| - \mathbf{H}_\varepsilon|^2 \chi_{\mathcal{I}_{i,k}}(u_\varepsilon) \, dx \\
 & + \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left( \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 - |\mathbf{H}_\varepsilon|^2 \right) dx.
 \end{aligned}$$

An application of the Grönwall inequality to Theorem 12 yields

$$\begin{aligned}
 \sup_{t \in [0, T]} E[u_\varepsilon | \xi](t) & + \sum_{i=1}^N \sum_{\substack{k=1 \\ k \neq i}}^N \int_{\mathbb{R}^d} \frac{1}{\varepsilon} |\varepsilon(B \cdot \xi_{i,k}) \xi_{i,k} |\nabla u_\varepsilon| - \mathbf{H}_\varepsilon|^2 \chi_{\mathcal{I}_{i,k}}(u_\varepsilon) \, dx \\
 & + \int_{\mathbb{R}^d} \frac{1}{2\varepsilon} \left( \left| \varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right|^2 - |\mathbf{H}_\varepsilon|^2 \right) dx \leq C(d, T, (\bar{\chi}(t))_{t \in [0, T]}) E[u_\varepsilon | \xi](0).
 \end{aligned}$$

Integrating the previous formula in time and inserting this estimate, we deduce by condition (2.4) on the weight  $\vartheta_i$ ,

$$\begin{aligned}
 & \int_{\mathbb{R}^d} (\psi_i(u_\varepsilon(\cdot, T)) - \bar{\chi}_i(\cdot, T)) \vartheta_i(\cdot, T) \, dx \\
 & \leq C(d, T, (\bar{\chi}(t))_{t \in [0, T]}) E[u_\varepsilon | \xi](0) + \int_0^T \int_{\mathbb{R}^d} (\psi_i(u_\varepsilon) - \bar{\chi}_i) \vartheta_i \, dx dt.
 \end{aligned}$$

The Grönwall inequality now implies our result, using (2.4) again.  $\blacksquare$

#### 4.5. Proof of the main theorem

Our main theorem (Theorem 1) is a simple consequence of Theorem 12 and Proposition 13.

*Proof.* Combining Theorem 12 and Proposition 13, we obtain the desired bounds

$$\sup_{t \in [0, T]} E_\varepsilon[u_\varepsilon | \xi](t) \leq C\varepsilon,$$

$$\sup_{t \in [0, T]} \max_{i \in \{1, \dots, N\}} \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon(\cdot, t)) - \bar{\chi}_i(\cdot, t)| \operatorname{dist}(x, \partial \operatorname{supp} \bar{\chi}_i(\cdot, t)) \, dx \leq C\varepsilon.$$

Finally, proceeding as in [14, Section 3], one can conclude about the error estimate of order  $\varepsilon^{1/2}$  in terms of the  $L^1$ -norm.  $\blacksquare$

### 5. Construction of well-prepared initial data

In this section we construct an initial datum  $u_\varepsilon(\cdot, 0)$  complying with the following relative energy estimate:

$$E[u_\varepsilon | \xi](0) \leq C\varepsilon, \quad (5.1)$$

where the constant  $C > 0$  depends on the initial data  $(\bar{\chi}_1(\cdot, 0), \dots, \bar{\chi}_N(\cdot, 0))$  and the potential  $W$  satisfies assumptions (A1)–(A4) (see Section 2). In particular, we provide an explicit construction of  $u_\varepsilon(\cdot, 0)$  for a network of interfaces meeting at two-dimensional triple junctions ( $d = 2$ ) satisfying the  $120^\circ$  angle condition. To this aim, we adopt a geometric setting for the initial network which was introduced in [13, Sections 5–6] in the general time-dependent case. A similar construction can be provided for three-dimensional double bubbles ( $d = 3$ ) satisfying the correct angle condition along the triple line, this time by exploiting the corresponding geometric setting given by [15, Sections 3–4].

Note that from our construction and the fast decay of the Modica–Mortola profiles towards the pure phases  $\alpha_i$ ,  $1 \leq i \leq N$ , it will also be apparent that our initial data  $u_\varepsilon(\cdot, 0)$  also satisfy the estimate

$$\max_{i \in \{1, \dots, N\}} \int_{\mathbb{R}^d} |\psi_i(u_\varepsilon(\cdot, 0)) - \bar{\chi}_i(\cdot, 0)| \operatorname{dist}(x, \partial \operatorname{supp} \bar{\chi}_i(\cdot, 0)) \, dx \leq C\varepsilon.$$

In fact, for this lower-order quantity one may even show the stronger bound  $O(\varepsilon^2)$ . In summary, the considerations in the present section will establish Proposition 3.

#### 5.1. Rescaled one-dimensional equilibrium profiles

For any distinct  $i, j \in \{1, \dots, N\}$ , let  $\gamma_{i,j}: [-1, 1] \rightarrow \mathbb{R}^2$  be the unique constant-speed  $C^1$  path connecting  $\alpha_i$  to  $\alpha_j$  such that  $\gamma_{i,j}(-1) = \alpha_i$ ,  $\gamma_{i,j}(1) = \alpha_j$ , and

$$\int_{-1}^1 \sqrt{2W(\gamma_{i,j}(r))} |\gamma'_{i,j}(r)| \, dr = 1.$$

Let  $\tilde{\theta}_{i,j}: \mathbb{R} \rightarrow [-1, +1]$  be the unique solution of the ODE

$$\tilde{\theta}'_{i,j}(s) = |\gamma'_{i,j}(\tilde{\theta}_{i,j}(s))|^{-1} \sqrt{2W(\gamma_{i,j}(\tilde{\theta}_{i,j}(s)))}$$

with boundary conditions  $\tilde{\theta}_{i,j}(\pm\infty) = \pm 1$ . Due to the growth properties of  $W$  in the neighborhoods of  $\alpha_i$  and  $\alpha_j$  (see condition (A1) in Section 2), the profile  $\tilde{\theta}_{i,j}$  approaches its boundary values  $\pm 1$  at  $\pm\infty$  with a power law of order  $\frac{2}{q-2}$  for  $q > 2$  and an exponential rate for  $q = 2$  [25].

Let  $s_{i,j}^0 \in \mathbb{R}$  such that  $\tilde{\theta}_{i,j}(s_{i,j}^0) = 0$ . Let  $\rho > 0$  be such that  $\tilde{\theta}_{i,j}(\rho + s_{i,j}^0) = \bar{\theta}_{i,j}^+$  and  $\tilde{\theta}_{i,j}(-\rho + s_{i,j}^0) = -\bar{\theta}_{i,j}^-$  for  $\bar{\theta}_{i,j}^\pm \in (0, 1)$ . We define the rescaled one-dimensional equilibrium profiles  $\theta_{i,j}: \mathbb{R} \rightarrow \gamma_{i,j}$  as

$$\theta_{i,j}(s) := \begin{cases} \gamma_{i,j} \left( \frac{1}{\bar{\theta}_{i,j}^+} \tilde{\theta}_{i,j}(s_{i,j}^0 \vee (s + s_{i,j}^0) \wedge (\rho + s_{i,j}^0)) \right) & \text{for } s \in [0, \infty), \\ \gamma_{i,j} \left( \frac{1}{\bar{\theta}_{i,j}^-} \tilde{\theta}_{i,j}((-\rho + s_{i,j}^0) \vee (s + s_{i,j}^0) \wedge s_{i,j}^0) \right) & \text{for } s \in (-\infty, 0), \end{cases} \quad (5.2)$$

so that  $\theta_{i,j}(s) = \alpha_i$  for any  $s \leq \rho$  and  $\theta_{i,j}(s) = \alpha_j$  for any  $s \geq \rho$ . Furthermore, we have

$$(\theta_{i,j})'(s) = \begin{cases} \frac{1}{\bar{\theta}_{i,j}^+} \sqrt{2W(\theta_{i,j}(s))} \frac{\gamma'_{i,j}}{|\gamma'_{i,j}|} \\ \quad \times \left( \frac{1}{\bar{\theta}_{i,j}^+} \tilde{\theta}_{i,j}(s_{i,j}^0 \vee (s + s_{i,j}^0) \wedge (\rho + s_{i,j}^0)) \right) & \text{for } s \in [0, \rho), \\ \frac{1}{\bar{\theta}_{i,j}^-} \sqrt{2W(\theta_{i,j}(s))} \frac{\gamma'_{i,j}}{|\gamma'_{i,j}|} \\ \quad \times \left( \frac{1}{\bar{\theta}_{i,j}^-} \tilde{\theta}_{i,j}((-\rho + s_{i,j}^0) \vee (s + s_{i,j}^0) \wedge s_{i,j}^0) \right) & \text{for } s \in (-\rho, 0), \\ 0 & \text{for } s \in (-\infty, -\rho] \\ & \cup [\rho, +\infty). \end{cases}$$

Note that, if  $W$  satisfies additional symmetry properties along the path  $\gamma_{i,j}$ , then  $\tilde{\theta}_{i,j}$  is odd, thus  $s_{i,j}^0 = 0$  and  $\bar{\theta}_{i,j}^- = \bar{\theta}_{i,j}^+$ . Moreover, if  $W$  satisfies additional symmetry properties with respect to all the paths  $\gamma_{i,j}$ , then all  $\bar{\theta}_{i,j}^\pm$  coincide.

## 5.2. Geometry of the initial network

For simplicity of notation, we shall omit the evaluation at initial time throughout this chapter, i.e. we write  $\bar{\chi}$  instead of  $\bar{\chi}(\cdot, 0)$ . Let  $\bar{\chi} = (\bar{\chi}_1, \dots, \bar{\chi}_N)$  be an initial partition of  $\mathbb{R}^2$  with interfaces  $\partial\{\bar{\chi}_i = 1\} \cap \partial\{\bar{\chi}_j = 1\} =: I_{i,j}$  for distinct  $i, j \in \{1, \dots, N\}$ . We decompose the network of interfaces according to its topological features, i.e. into smooth



two-phase interfaces and triple junctions. Suppose that the network has  $P$  of such topological features  $\mathcal{T}_n, n \in \{1, \dots, P\}$ . We then split  $\{1, \dots, P\} =: \mathcal{C} \cup \mathcal{P}$ , where  $\mathcal{C}$  enumerates the connected components of the two-phase interfaces and  $\mathcal{P}$  enumerates the triple junctions. In particular, if  $p \in \mathcal{P}$ ,  $\mathcal{T}_p$  is a triple junction, whereas if  $c \in \mathcal{C}$ ,  $\mathcal{T}_c$  is a connected component of a two-phase interface  $I_{i,j}$  for some distinct  $i, j \in \{1, \dots, N\}$ .

In the following we use a suitable notion of neighborhood for a single connected component of the network of interfaces provided by [13, Definition 21]. In particular, we adopt the notion of localization radius, which allows one to define the diffeomorphism corresponding to a single connected component of a network as follows. Let  $r_{i,j}$  be a localization radius for the interface  $I_{i,j}$  and let  $\bar{n}_{i,j}$  be the normal vector field to  $I_{i,j}$  pointing towards  $\{\bar{\chi}_j = 1\}$  for some distinct  $i, j \in \{1, \dots, N\}$ . Then the map  $\Psi_{i,j}: I_{i,j} \times (-r_{i,j}, r_{i,j}) \rightarrow \mathbb{R}^2, (x, s) \mapsto x + s\bar{n}_{i,j}(x)$  defines a diffeomorphism, whose inverse can be split as follows:  $\Psi_{i,j}^{-1}: \text{im}(\Psi_{i,j}) \mapsto I_{i,j} \times (-r_{i,j}, r_{i,j}), x \mapsto (P_{I_{i,j}}x, \text{dist}^\pm(x, I_{i,j}))$ , where  $P_{I_{i,j}}: \text{im}(\Psi_{i,j}) \rightarrow I_{i,j}$  represent the projection onto the nearest point on the interface  $I_{i,j}$ , whereas  $\text{dist}^\pm(\cdot, I_{i,j}): \text{im}(\Psi_{i,j}) \rightarrow (-r_{i,j}, r_{i,j})$  is a signed distance function.

Similarly to [13, Definition 24], we provide a notion of admissible localization radius for a triple junction.

**Definition 16.** Let  $d = 2$ . Let  $\bar{\chi} = (\bar{\chi}_1, \dots, \bar{\chi}_N)$  be an initial partition of  $\mathbb{R}^d$  with interfaces  $\partial\{\bar{\chi}_i = 1\}, i = 1, \dots, N$ . Let  $\mathcal{T}$  be a triple junction present in the network of interfaces of  $\bar{\chi}$ , which we assume for simplicity to be formed by the phases 1, 2, and 3. For each  $i \in \{1, 2, 3\}$ , denote by  $\mathcal{T}_{i,i+1}$  the connected component of  $I_{i,i+1}$  with an endpoint at the triple junction  $\mathcal{T}$  and let  $r_{i,i+1} \in (0, 1]$  be an admissible localization radius for the interface  $I_{i,i+1}$  in the sense of [13, Definition 21]. We call a scale  $r = r_{\mathcal{T}} \in (0, r_{1,2} \wedge r_{2,3} \wedge r_{3,1}]$  an admissible localization radius for the triple junction  $\mathcal{T}$  if there exists a wedge decomposition of the neighborhood  $B_r(\mathcal{T})$  of the triple junction in the following sense:

For each  $i \in \{1, 2, 3\}$  there exist sets  $W_{i,i+1}$  and  $W_i$  with the following properties:

First, the sets  $W_{i,i+1}$  and  $W_i$  are nonempty subsets of  $B_r(\mathcal{T})$  with pairwise disjoint interior such that

$$\bigcup_{i \in \{1,2,3\}} \overline{W_{i,i+1}} \cup \overline{W_i} = \overline{B_r(\mathcal{T})}.$$

Second, each of these sets is represented by a cone with apex at the triple junction  $\mathcal{T}$  intersected with  $B_r(\mathcal{T})$ . More precisely, there exist six pairwise distinct unit-length vectors  $(X_{i,i+1}^i, X_{i,i+1}^{i+1})_{i \in \{1,2,3\}}$  such that for all  $i \in \{1, 2, 3\}$  we have

$$\begin{aligned} W_{i,i+1} &= (\mathcal{T} + \{aX_{i,i+1}^i + bX_{i,i+1}^{i+1} : a, b \in (0, \infty)\}) \cap B_r(\mathcal{T}), \\ W_i &= (\mathcal{T} + \{aX_{i,i+1}^i + bX_{i-1,i}^i : a, b \in (0, \infty)\}) \cap B_r(\mathcal{T}). \end{aligned}$$

The opening angles of these cones are numerically fixed by

$$X_{i,i+1}^i \cdot X_{i,i+1}^{i+1} = \cos(\pi/2) = 0, \quad X_{i,i+1}^i \cdot X_{i-1,i}^i = \cos(\pi/6).$$

Third, we require that for all  $i \in \{1, 2, 3\}$  it holds that

$$\begin{aligned} B_r(\mathcal{T}) \cap \mathcal{T}_{i,i+1} &\subset W_{i,i+1} \cup \mathcal{T} \subset \mathbb{H}_{i,i+1}, \\ W_i &\subset \mathbb{H}_{i,i+1} \cap \mathbb{H}_{i,i-1}, \end{aligned}$$

with the domains  $\mathbb{H}_{i,i+1} := \{x \in \mathbb{R}^2 : x \in \text{im}(\Psi_{i,i+1})\} \cap B_r(\mathcal{T})$ , where  $\Psi_{i,i+1}$  is the diffeomorphism defining the neighborhood of  $I_{i,i+1}$  in the sense of [13, Definition 21].

Let  $r := \min_{p \in \mathcal{P}} r_p$ , where  $r_p$  is the admissible localization radius for the triple junction  $\mathcal{T}_p$ . Let  $\rho \in (0, r)$ . Consider a triple junction  $\mathcal{T}$ , which we assume for simplicity are formed by the phases 1, 2, and 3. Let  $\varepsilon < \rho$ ; then for all  $i \in \{1, 2, 3\}$  we define (see Figure 2)

(i) the two-dimensional regions

$$\begin{aligned} W_{i,i+1}^{\rho,+} &:= W_{i,i+1} \cap \{x \in \mathbb{R}^2 : 0 \leq \text{dist}^\pm(x, \mathcal{T}_{i,i+1}) \leq \rho\}, \\ W_{i,i+1}^{\rho,-} &:= W_{i,i+1} \cap \{x \in \mathbb{R}^2 : -\rho \leq \text{dist}^\pm(x, \mathcal{T}_{i,i+1}) \leq 0\}, \\ W_{i,i+1}^\rho &:= W_{i,i+1}^{\rho,+} \cup W_{i,i+1}^{\rho,-}, \\ W_{i,i+1}^{\varepsilon,\pm} &:= W_{i,i+1}^{\rho,\pm} \cap B_\varepsilon(\mathcal{T}), \\ W_{i,i+1}^\varepsilon &:= W_{i,i+1}^{\varepsilon,+} \cup W_{i,i+1}^{\varepsilon,-}, \end{aligned}$$

satisfying the inclusions

$$\mathcal{T}_{i,i+1} \cap B_r(\mathcal{T}) \subset W_{i,i+1}^\rho, \quad \mathcal{T}_{i,i+1} \cap B_\varepsilon(\mathcal{T}) \subset W_{i,i+1}^\varepsilon \subset W_{i,i+1}^\rho;$$

(ii) the one-dimensional segments resp. arcs

$$\begin{aligned} R_{i,i+1}^i &:= (\mathcal{T} + \{aX_{i,i+1}^i : a \in (0, \infty)\}) \cap \partial W_{i,i+1}^\rho, \\ R_{i,i+1}^+ &:= R_{i,i+1}^{i+1}, \quad R_{i,i+1}^- := R_{i,i+1}^i, \\ H_{i,i+1}^\varepsilon &:= W_{i,i+1}^\rho \cap \partial B_\varepsilon(\mathcal{T}), \\ H_{i,i+1}^{\varepsilon,+} &:= H_{i,i+1}^\varepsilon \cap \{x \in \mathbb{R}^2 : 0 \leq \text{dist}^\pm(x, \mathcal{T}_{i,i+1}) \leq \rho\}, \\ H_{i,i+1}^{\varepsilon,-} &:= H_{i,i+1}^\varepsilon \cap \{x \in \mathbb{R}^2 : -\rho \leq \text{dist}^\pm(x, \mathcal{T}_{i,i+1}) \leq 0\}; \end{aligned}$$

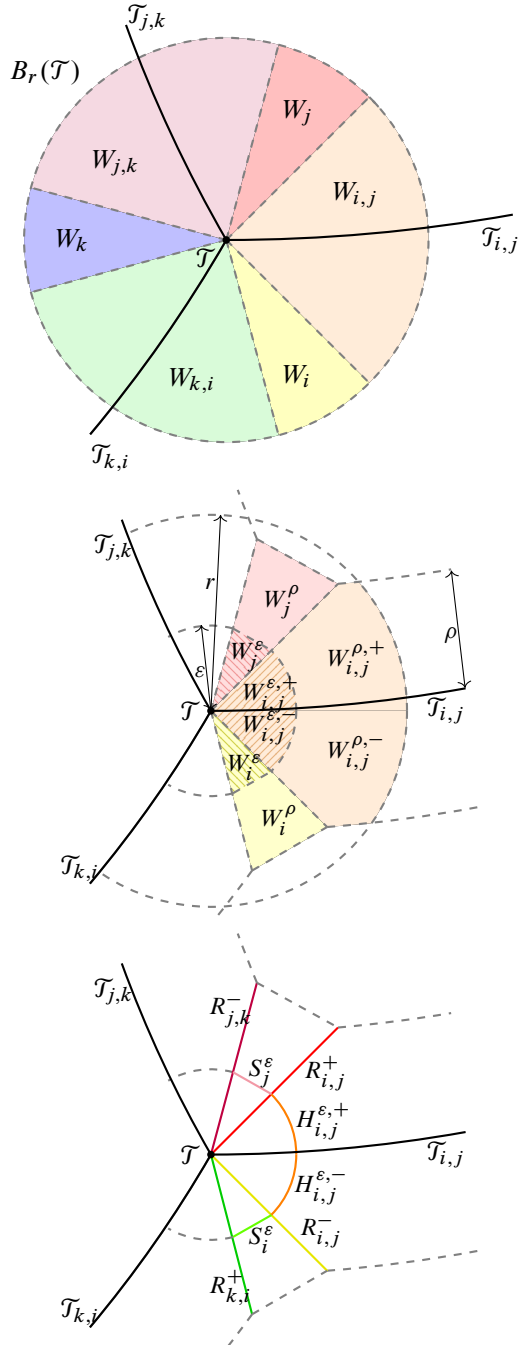
(iii)  $S_i^\varepsilon$  as the segment connecting  $R_{i,i+1}^i \cap \partial B_\varepsilon(\mathcal{T})$  to  $R_{i,i-1}^i \cap \partial B_\varepsilon(\mathcal{T})$ ;

(iv) the two-dimensional triangular regions  $W_i^\rho$  resp.  $W_i^\varepsilon$  as the one with sides  $R_{i,i+1}^i$  and  $R_{i,i-1}^i$  resp. the one delimited by  $S_i^\varepsilon$ ,  $R_{i,i+1}^i$ , and  $R_{i,i-1}^i$ , thus satisfying the inclusions

$$W_i^\varepsilon \subset W_i^\rho \subset W_i.$$

Furthermore, we introduce

$$\begin{aligned} P_{R_{i,i+1}^+} &: W_i^\rho \cup (W_{i,i+1}^\rho \cap \{x \in \Psi_{i,i+1}(I_{i,i+1} \times [0, \rho])\}) \rightarrow R_{i,i+1}^+, \\ P_{R_{i,i+1}^-} &: W_i^\rho \cup (W_{i,i+1}^\rho \cap \{x \in \Psi_{i,i+1}(I_{i,i+1} \times (-\rho, 0])\}) \rightarrow R_{i,i+1}^-, \\ P_{H_{i,i+1}^\varepsilon} &: W_i^\varepsilon \rightarrow H_{i,i+1}^\varepsilon, \\ P_{S_i^\varepsilon} &: W_i^\varepsilon \rightarrow S_i^\varepsilon \end{aligned}$$



**Figure 2.** Notation and geometry for the construction of well-prepared initial data at a triple junction  $\mathcal{T}$  ( $d = 2$ ) formed by the phases  $i$ ,  $j$ , and  $k$  for mutually distinct  $i, j, k \in \{1, \dots, N\}$ .

as the orthogonal projections onto the nearest point on  $R_{i,i+1}^+$ ,  $R_{i,i+1}^-$ ,  $H_{i,i+1}^\varepsilon$ , and  $S_i^\varepsilon$ , respectively.

### 5.3. Construction of the initial datum

We construct the initial datum  $u_{\varepsilon,0} := u_\varepsilon(\cdot, 0)$  separately in each two-dimensional region identified by the geometry of the network as introduced above. Then we show that in each of these regions the initial relative energy estimate (5.1) holds true.

**Neighborhood of a connected component of a two-phase interface.** Let  $\mathcal{T}_{i,j}$  be a connected component of  $I_{i,j}$  with either one or two endpoints at a triple junction for some distinct  $i, j \in \{1, \dots, N\}$ . Let  $\mathcal{P}_{i,j} \subset \mathcal{P}$  enumerate the numbers of triple junctions as endpoints of  $\mathcal{T}_{i,j}$ . Then we can define

$$M_{i,j} := \left( \{x \in \mathbb{R}^2 : -\rho \leq \text{dist}^\pm(x, \mathcal{T}_{i,j}) \leq \rho\} \setminus \bigcup_{p \in \mathcal{P}_{i,j}} B_r(\mathcal{T}_p) \right) \cup \left( \bigcup_{p \in \mathcal{P}_{i,j}} (W_{i,j}^\rho \setminus W_{i,j}^\varepsilon) \right).$$

In the two-dimensional region  $M_{i,j}$  we define the initial datum  $u_{\varepsilon,0}$  by means of the rescaled one-dimensional equilibrium profile (5.2) as

$$u_{\varepsilon,0}(x) := \theta_{i,j}(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j})) \quad \text{for any } x \in M_{i,j},$$

whence we obtain

$$\begin{aligned} & E_{M_{i,j}}[u_\varepsilon|\xi](0) \\ &:= \int_{M_{i,j}} \frac{1}{2\varepsilon} |(\theta_{i,j})'(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j}))|^2 + \frac{1}{\varepsilon} W(\theta_{i,j}(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j}))) \\ &\quad - \frac{1}{2\varepsilon} (\xi_{i,j} \cdot n_{i,j})(\theta_{i,j})'(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j})) \cdot \partial_u \psi_{i,j}(\theta_{i,j}(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j}))) \, dx \\ &\leq \int_{M_{i,j}} \frac{1}{2\varepsilon} |(\theta_{i,j})'(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j}))|^2 + \frac{1}{\varepsilon} W(\theta_{i,j}(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j}))) \\ &\quad - \frac{1}{\varepsilon} |\xi_{i,j}|^2 \frac{1}{\theta_{i,j}^\pm} 2W(\theta_{i,j}(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j}))) \, dx \\ &\leq \int_{M_{i,j}} (1 - |\xi_{i,j}|^2) \left[ \frac{1}{2\varepsilon} |(\theta_{i,j})'(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j}))|^2 + \frac{1}{\varepsilon} W(\theta_{i,j}(\varepsilon^{-1} \text{dist}^\pm(x, \mathcal{T}_{i,j}))) \right] dx. \end{aligned}$$

Here, we used that  $\psi_k = 0$  along  $\gamma_{i,j}$  for any  $k \in \{1, \dots, N\} \setminus \{i, j\}$  and that  $\partial_u \psi_{i,j}(\gamma_{i,j}) \cdot \frac{\gamma'_{i,j}}{|\gamma'_{i,j}|} = 2\sqrt{2W(\gamma_{i,j})}$ . Indeed, our assumption (A4) implies that  $\partial_u \psi_{i,j}(\gamma_{i,j}) \cdot \gamma'_{i,j} \leq 2\sqrt{2W(\gamma_{i,j})}|\gamma'_{i,j}|$ , and then (A3) gives the equality sign by contradiction. Then, in the

last step, we added a zero and we used  $|(\theta_{i,j})'(s)|^2 \leq \frac{1}{(\bar{\theta}_{i,j}^\pm)^2} 2W(\theta_{i,j}(s))$ . Note that  $1 - |\xi_{i,j}|^2 \leq c \operatorname{dist}^2(\cdot, \mathcal{T}_{i,j})$ , and that  $W(\theta_{i,j}(s))$  has an exponential resp. a power-law decay of order  $\frac{2q}{q-2}$  for  $q = 2$  resp.  $q > 2$  as  $s$  approaches the extrema of  $(-\rho, \rho)$ , and then it vanishes for  $s \in (-\infty, -\rho] \cup [\rho, \infty)$ . As a consequence, we obtain  $E_{M_{i,j}}[u_\varepsilon | \xi](0) \leq C\varepsilon^2$  for some constant  $C > 0$ .

**Pure-phase region.** Let  $i \in \{1, \dots, N\}$ . Let  $\mathcal{P}_i \subset \mathcal{P}$  enumerate the numbers of triple junctions as endpoints of a connected component of an interface between phase  $i$  and any other one. We set  $u_{\varepsilon,0} = \alpha_i$  in the pure-phase region

$$\{\bar{\chi}_i = 1\} \setminus \left( \bigcup_{p \in \mathcal{P}_i} W_i \cup \bigcup_{j:j \neq i} \{x - sn_{i,j}(x), x \in I_{i,j}, s \in [0, \rho]\} \right).$$

Then  $|\nabla u_{\varepsilon,0}| = 0$  and, having  $W(\alpha_i) = 0$ , the initial relative entropy is equal to zero in the pure-phase region.

**Triple junction wedge containing a connected component of a two-phase interface.**

Given a triple junction  $\mathcal{T}$ , let  $i, j \in \{1, \dots, N\}$ ,  $i \neq j$ , be two of the three phases forming  $\mathcal{T}$  and let  $k \in \{1, \dots, N\} \setminus \{i, j\}$  be the third one. The initial datum  $u_{\varepsilon,0}$  in the corresponding wedge  $W_{i,j}^{\varepsilon,\pm}$  is given by interpolation via orthogonal projections  $P_{H_{i,j}^\varepsilon}$ ,  $P_{I_{i,j}}$ , and  $P_{R_{i,j}^\pm}$ , which reads

$$\begin{aligned} u_{\varepsilon,0}(x) &= \frac{\operatorname{dist}(x, I_{i,j}) + \operatorname{dist}(x, R_{i,j}^\pm)}{\operatorname{dist}(x, H_{i,j}^\varepsilon) + \operatorname{dist}(x, I_{i,j}) + \operatorname{dist}(x, R_{i,j}^\pm)} u_{H_{i,j}^\varepsilon}(P_{H_{i,j}^\varepsilon} x) \\ &+ \frac{\operatorname{dist}(x, H_{i,j}^\varepsilon) + \operatorname{dist}(x, R_{i,j}^\pm)}{\operatorname{dist}(x, H_{i,j}^\varepsilon) + \operatorname{dist}(x, I_{i,j}) + \operatorname{dist}(x, R_{i,j}^\pm)} u_{I_{i,j}}(P_{I_{i,j}} x) \\ &+ \frac{\operatorname{dist}(x, H_{i,j}^\varepsilon) + \operatorname{dist}(x, I_{i,j})}{\operatorname{dist}(x, H_{i,j}^\varepsilon) + \operatorname{dist}(x, I_{i,j}) + \operatorname{dist}(x, R_{i,j}^\pm)} u_{R_{i,j}^\pm}(P_{R_{i,j}^\pm} x) \end{aligned}$$

for any  $x \in W_{i,j}^{\varepsilon,\pm}$ , where

$$\begin{aligned} u_{H_{i,j}^\varepsilon}(x) &= \frac{h_{i,j}^\pm(x, \mathcal{T})}{h_{i,j}^{\varepsilon,\pm}} \theta_{i,j}(\pm\varepsilon^{-1} \bar{h}_{i,j}^{\varepsilon,\pm}) + \frac{h_{i,j}^{\varepsilon,\pm} - h_{i,j}^\pm(x, \mathcal{T})}{h_{i,j}^{\varepsilon,\pm}} \theta_{i,j}(0) \quad \text{for any } x \text{ along } H_{i,j}^\varepsilon, \\ u_{I_{i,j}}(x) &= \frac{l_{i,j}(x, \mathcal{T})}{l_{i,j}^\varepsilon} \theta_{i,j}(0) + \frac{l_{i,j}^\varepsilon - l_{i,j}(x, \mathcal{T})}{l_{i,j}^\varepsilon} \bar{\alpha} \quad \text{for any } x \text{ along } I_{i,j}, \\ u_{R_{i,j}^\pm}(x) &= \frac{r_{i,j}^\pm(x, \mathcal{T})}{\varepsilon} \theta_{i,j}(\pm\varepsilon^{-1} \bar{h}_{i,j}^{\varepsilon,\pm}) + \frac{\varepsilon - r_{i,j}^\pm(x, \mathcal{T})}{\varepsilon} \bar{\alpha} \quad \text{for any } x \text{ along } R_{i,j}^\pm, \end{aligned}$$

where  $\bar{\alpha} = \frac{\alpha_i + \alpha_j + \alpha_k}{3}$ ,  $\bar{h}_{i,j}^{\varepsilon,\pm} := \operatorname{dist}^\pm(H_{i,j}^\varepsilon \cap R_{i,j}^\pm, I_{i,j})$ ,  $l_{i,j}^\varepsilon$  resp.  $h_{i,j}^{\varepsilon,\pm}$  is the length of  $I_{i,j} \cap B_\varepsilon(\mathcal{T})$  resp.  $\partial B_\varepsilon(\mathcal{T}) \cap W_{i,j}^{\varepsilon,\pm}$ , whereas  $h_{i,j}^\pm(\cdot, \mathcal{T})$ ,  $l_{i,j}(\cdot, \mathcal{T})$  resp.  $r_{i,j}^\pm(\cdot, \mathcal{T})$  is the length of the path along  $H_{i,j}^\varepsilon$ ,  $I_{i,j}$  resp.  $R_{i,j}^\pm$  connecting to  $\mathcal{T}$ . Since  $l_{i,j}^\varepsilon$  and  $h_{i,j}^{\varepsilon,\pm}$  are

of order  $\varepsilon$ , then this construction gives that  $|\nabla u_{\varepsilon,0}(x)|$  is of order  $1/\varepsilon$  for any  $x \in W_{i,j}^\varepsilon$ . Hence, the area of  $W_{i,j}^\varepsilon$  being of order  $\varepsilon^2$ , we can deduce

$$\begin{aligned} E_{W_{i,j}^\varepsilon}[u_\varepsilon|\xi](0) &:= \int_{W_{i,j}^\varepsilon} \frac{\varepsilon}{2} |\nabla u_{\varepsilon,0}|^2 + \frac{1}{\varepsilon} W(u_{\varepsilon,0}) + \sum_{\ell=1}^N \xi_\ell \cdot \nabla(\psi_\ell \circ u_{\varepsilon,0}) \, dx \\ &\leq \int_{W_{i,j}^\varepsilon} \frac{\varepsilon}{2} |\nabla u_{\varepsilon,0}|^2 + \frac{C}{\varepsilon} + C |\nabla u_{\varepsilon,0}| \, dx \leq C\varepsilon, \end{aligned}$$

for some constant  $C > 0$  varying from line to line.

**Triple junction wedge not containing any connected component of a two-phase interface.** Given a triple junction  $\mathcal{T}$ , let  $i, j, k \in \{1, \dots, N\}$  be three distinct phases forming  $\mathcal{T}$ . The initial datum  $u_{\varepsilon,0}$  in the region  $W_j^\varepsilon$  is given by interpolation via orthogonal projections  $P_{S_j^\varepsilon}$ ,  $P_{R_{j,k}^-}$ , and  $P_{R_{i,j}^+}$ , which reads

$$\begin{aligned} u_{\varepsilon,0}(x) &= \frac{\text{dist}(x, R_{j,k}^-) + \text{dist}(x, R_{i,j}^+)}{\text{dist}(x, S_j^\varepsilon) + \text{dist}(x, R_{j,k}^-) + \text{dist}(x, R_{i,j}^+)} u_{H_{i,j}^\varepsilon}(P_{S_j^\varepsilon} x) \\ &+ \frac{\text{dist}(x, S_j^\varepsilon) + \text{dist}(x, R_{i,j}^+)}{\text{dist}(x, S_j^\varepsilon) + \text{dist}(x, R_{j,k}^-) + \text{dist}(x, R_{i,j}^+)} u_{R_{j,k}^-}(P_{R_{j,k}^-} x) \\ &+ \frac{\text{dist}(x, S_j^\varepsilon) + \text{dist}(x, R_{j,k}^-)}{\text{dist}(x, S_j^\varepsilon) + \text{dist}(x, R_{j,k}^-) + \text{dist}(x, R_{i,j}^+)} u_{R_{i,j}^+}(P_{R_{i,j}^+} x), \end{aligned}$$

for any  $x \in W_j^\varepsilon$ , where

$$\begin{aligned} u_{S_j^\varepsilon}(x) &= \frac{s_j(x)}{s_j^\varepsilon} \theta_{i,j}(\varepsilon^{-1} \bar{h}_{i,j}^\varepsilon) + \frac{s_j^\varepsilon - s_j(x)}{s_j^\varepsilon} \theta_{j,k}(-\varepsilon^{-1} \bar{h}_{j,k}^\varepsilon) \quad \text{for any } x \text{ along } S_j^\varepsilon, \\ u_{R_{j,k}^-}(x) &= \frac{r_{j,k}(x, \mathcal{T})}{\varepsilon} \theta_{j,k}(-\varepsilon^{-1} \bar{h}_{j,k}^\varepsilon) + \frac{\varepsilon - r_{j,k}(x, \mathcal{T})}{\varepsilon} \bar{\alpha} \quad \text{for any } x \text{ along } R_{j,k}^-, \\ u_{R_{i,j}^+}(x) &= \frac{r_{i,j}(x, \mathcal{T})}{\varepsilon} \theta_{i,j}(\varepsilon^{-1} \bar{h}_{i,j}^\varepsilon) + \frac{\varepsilon - r_{i,j}(x, \mathcal{T})}{\varepsilon} \bar{\alpha} \quad \text{for any } x \text{ along } R_{i,j}^+, \end{aligned}$$

where  $\bar{h}_{i,j}^\varepsilon := \text{dist}(H_{i,j}^\varepsilon \cap R_{i,j}^+, I_{i,j})$ ,  $\bar{h}_{j,k}^\varepsilon := \text{dist}(H_{j,k}^\varepsilon \cap R_{j,k}^-, I_{j,k})$ ,  $s_j^\varepsilon$  is the length of the segment  $S_j^\varepsilon$ , whereas  $s_j$  resp.  $r_{i,j}(\cdot, \mathcal{T})$  is the length of the path along  $S_j^\varepsilon$  resp.  $R_{i,j}^+$  connecting to  $R_{j,k}^-$  resp.  $\mathcal{T}$ . Since  $s_j^\varepsilon$  is of order  $\varepsilon$ , we have that  $|\nabla u_{\varepsilon,0}(x)|$  is of order  $1/\varepsilon$  for any  $x \in W_j^\varepsilon$ , whence as above we can deduce that

$$E_{W_j^\varepsilon}[u_\varepsilon|\xi](0) := \int_{W_j^\varepsilon} \frac{\varepsilon}{2} |\nabla u_{\varepsilon,0}|^2 + \frac{1}{\varepsilon} W(u_{\varepsilon,0}) + \sum_{\ell=1}^N \xi_\ell \cdot \nabla(\psi_\ell \circ u_{\varepsilon,0}) \, dx \leq C\varepsilon,$$

for some constant  $C > 0$ .

**Interpolation between two rescaled one-dimensional equilibrium profiles.** First, we introduce for any  $x \in W_j^\rho \setminus W_j^\varepsilon$  the set of coordinates  $(s, h)$ , where  $h$  denotes the distance from  $S_j^\varepsilon$  while  $s$  is such that  $\nabla s \cdot \nabla h = 0$  and  $s = 0$  whenever  $x \in R_{j,k}^-$ . Hence,  $h \in [0, \tilde{r}_\varepsilon]$ , where  $\tilde{r}_\varepsilon := (r - \varepsilon) \cos(\frac{\pi}{12})$  is fixed, and  $s \in [0, 2\tilde{h}_\varepsilon \sin(\frac{\pi}{12})]$ , where  $\tilde{h}_\varepsilon := (\cos(\frac{\pi}{12}))^{-1}h + \varepsilon$ . The initial datum  $u_{\varepsilon,0}$  in  $W_j^\rho \setminus W_j^\varepsilon$  is given as

$$u_{\varepsilon,0}(x) = \frac{s}{2\tilde{h}_\varepsilon \sin(\frac{\pi}{12})} \theta_{j,k}(-\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) \\ + \frac{2\tilde{h}_\varepsilon \sin(\frac{\pi}{12}) - s}{2\tilde{h}_\varepsilon \sin(\frac{\pi}{12})} \theta_{i,j}(\varepsilon^{-1} \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j})),$$

for any  $x \in W_j^\rho \setminus W_j^\varepsilon$ , where  $P_{R_{j,k}^-}^s$  and  $P_{R_{i,j}^+}^s$  are projections along the  $s$ -axis onto  $R_{j,k}^-$  and  $R_{i,j}^+$ , respectively. Hence, we compute

$$\partial_s u_{\varepsilon,0}(x) = \frac{1}{2\tilde{h}_\varepsilon \sin(\frac{\pi}{12})} (\theta_{j,k}(-\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) - \theta_{i,j}(\varepsilon^{-1} \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j}))),$$

and

$$\partial_h u_{\varepsilon,0}(x) \\ = \frac{s}{2\tilde{h}_\varepsilon^2 \sin(\frac{\pi}{12}) \cos(\frac{\pi}{12})} (\theta_{j,k}(-\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) - \theta_{i,j}(\varepsilon^{-1} \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j}))) \\ - \frac{s}{2\varepsilon\tilde{h}_\varepsilon \sin(\frac{\pi}{12})} \partial_h \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k}) (\theta_{j,k})'(-\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) \\ + \frac{2\tilde{h}_\varepsilon \sin(\frac{\pi}{12}) - s}{2\varepsilon\tilde{h}_\varepsilon \sin(\frac{\pi}{12})} \partial_h \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j}) (\theta_{i,j})'(\varepsilon^{-1} \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j})).$$

For the sake of brevity, we introduce the notation  $\tilde{\lambda}(x) := \frac{s}{2\tilde{h}_\varepsilon \sin(\frac{\pi}{12})} \in [0, 1]$  for any  $x \in W_j^\rho \setminus W_j^\varepsilon$ . By adding zeros and using Young's inequality, together with the fact that  $s^2 \leq 4\tilde{h}_\varepsilon^2 \sin^2(\frac{\pi}{12})$ , one can obtain

$$\frac{\varepsilon}{2} |\nabla u_{\varepsilon,0}(x)|^2 = \frac{\varepsilon}{2} |\partial_s u_{\varepsilon,0}(x)|^2 + \frac{\varepsilon}{2} |\partial_h u_{\varepsilon,0}(x)|^2 \\ \leq C \frac{\varepsilon}{\tilde{h}_\varepsilon^2} |\theta_{j,k}(-\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) - \alpha_j|^2 \\ + C \frac{\varepsilon}{\tilde{h}_\varepsilon^2} |\theta_{i,j}(\varepsilon^{-1} \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j})) - \alpha_j|^2 \\ + C \tilde{\lambda}^2(x) \frac{1}{\varepsilon} |(\theta_{j,k})'(-\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k}))|^2 \\ + C (1 - \tilde{\lambda}(x))^2 \frac{1}{\varepsilon} |(\theta_{i,j})'(\varepsilon^{-1} \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j}))|^2,$$

for some constant  $C > 0$ . First, we consider

$$\begin{aligned}
 & \int_{W_j^\rho \setminus W_j^\varepsilon} \frac{\varepsilon}{\tilde{h}_\varepsilon^2} |\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) - \alpha_j|^2 dx \\
 &= 2 \sin\left(\frac{\pi}{12}\right) \int_0^{\tilde{r}_\varepsilon} \frac{\varepsilon}{\tilde{h}_\varepsilon} |\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{k,j})) - \alpha_j|^2 dh \\
 &\leq C \int_0^{\tilde{r}_\varepsilon} |\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) - \alpha_j|^2 dh \\
 &\leq C \int_0^{\tilde{r}_\varepsilon} \frac{h}{\varepsilon} |\theta_{j,k}(\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) - \alpha_j| |(\theta_{j,k})'(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k}))| dh \\
 &\leq C \int_0^{\tilde{r}_\varepsilon} \frac{h}{\varepsilon} |(\theta_{j,k})'(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k}))| dh,
 \end{aligned}$$

where we integrated by parts and  $C > 0$  is a suitable constant varying from line to line. Then we observe that  $\operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k})$  is a homogeneous and increasing function with respect to  $h$ . On the other hand, we recall that  $|(\theta_{j,k})'(s)|$  has an exponential resp. a power-law decay of order  $\frac{q}{q-2}$  for  $q = 2$  resp.  $q > 2$  as  $s$  approaches the extrema of  $(-\rho, \rho)$ , and then it vanishes for  $s \in (-\infty, -\rho] \cup [\rho, \infty)$ . As a consequence, we obtain

$$\int_{W_j^\rho \setminus W_j^\varepsilon} \frac{\varepsilon}{\tilde{h}_\varepsilon^2} |\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) - \alpha_j|^2 dx \leq C \varepsilon,$$

and analogously

$$\int_{W_j^\rho \setminus W_j^\varepsilon} \frac{\varepsilon}{\tilde{h}_\varepsilon^2} |\theta_{i,j}(\varepsilon^{-1} \operatorname{dist}(P_{R_{i,j}^+}^s x, I_{i,j})) - \alpha_j|^2 dx \leq C \varepsilon.$$

Second, we have

$$\begin{aligned}
 & \int_{W_j^\rho \setminus W_j^\varepsilon} \tilde{\lambda}^2(x) \frac{1}{\varepsilon} |(\theta_{j,k})'(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k}))|^2 dx \\
 &= 2 \sin\left(\frac{\pi}{12}\right) \int_0^{\tilde{r}_\varepsilon} \tilde{\lambda}^2(x) \frac{\tilde{h}_\varepsilon}{\varepsilon} |(\theta_{j,k})'(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k}))|^2 dh \\
 &\leq C \int_0^{\tilde{r}_\varepsilon} \left(\frac{h}{\varepsilon} + 1\right) |(\theta_{j,k})'(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^-}^s x, I_{j,k}))|^2 dh \\
 &\leq C \varepsilon,
 \end{aligned}$$

and analogously

$$\int_{W_j^\rho \setminus W_j^\varepsilon} (1 - \tilde{\lambda}(x))^2 \frac{1}{\varepsilon} |(\theta_{i,j})'(\varepsilon^{-1} \operatorname{dist}(P_{R_{i,j}^+}^s x, I_{i,j}))|^2 dx \leq C \varepsilon.$$



Observe that by adding zeros we can write

$$\begin{aligned}
 W(u_{\varepsilon,0}(x)) &= \tilde{\lambda}(x)(W(u_{\varepsilon,0}(x)) - W(\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k})))) \\
 &\quad + (1 - \tilde{\lambda}(x))(W(u_{\varepsilon,0}(x)) - W(\theta_{i,j}(\varepsilon^{-1} \operatorname{dist}(P_{R_{i,j}^s}^+ x, I_{i,j})))) \\
 &\quad + \tilde{\lambda}(x)W(\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k}))) \\
 &\quad + (1 - \tilde{\lambda}(x))W(\theta_{i,j}(\varepsilon^{-1} \operatorname{dist}(P_{R_{i,j}^s}^+ x, I_{i,j}))),
 \end{aligned}$$

where  $\tilde{\lambda}(x), 1 - \tilde{\lambda}(x) \in [0, 1]$  for any  $x \in W_j^\rho \setminus W_j^\varepsilon$ . Since  $W$  is Lipschitz (see (A1) in Section 2), then by adding zeros we obtain

$$\begin{aligned}
 &\left| \frac{1}{\varepsilon} W(u_{\varepsilon,0}(x)) - \frac{1}{\varepsilon} W(\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k}))) \right| \\
 &\leq \frac{C}{\varepsilon} |u_{\varepsilon,0}(x) - \theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k}))| \\
 &\leq \frac{C}{\varepsilon} (1 - \tilde{\lambda}(x)) (|\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k})) - \alpha_j| \\
 &\quad + |\theta_{i,j}(\varepsilon^{-1} \operatorname{dist}(P_{R_{i,j}^s}^+ x, I_{i,j})) - \alpha_j|)
 \end{aligned}$$

and

$$\begin{aligned}
 &\left| \frac{1}{\varepsilon} W(u_{\varepsilon,0}(x)) - \frac{1}{\varepsilon} W(\theta_{i,j}(\varepsilon^{-1} \operatorname{dist}(P_{R_{i,j}^s}^+ x, I_{i,j}))) \right| \\
 &\leq \frac{C}{\varepsilon} |u_{\varepsilon,0}(x) - \theta_{i,j}(\varepsilon^{-1} \operatorname{dist}(P_{R_{i,j}^s}^+ x, I_{i,j}))| \\
 &\leq \frac{C}{\varepsilon} \tilde{\lambda}(x) (|\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k})) - \alpha_j| \\
 &\quad + |\theta_{i,j}(\varepsilon^{-1} \operatorname{dist}(P_{R_{i,j}^s}^+ x, I_{i,j})) - \alpha_j|).
 \end{aligned}$$

In particular, we have

$$\begin{aligned}
 &\int_{W_j^\rho \setminus W_j^\varepsilon} \frac{1}{\varepsilon} |\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k})) - \alpha_j| \, dx \\
 &= 2 \sin\left(\frac{\pi}{12}\right) \int_0^{\tilde{r}_\varepsilon} \frac{\tilde{h}_\varepsilon}{\varepsilon} |\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k})) - \alpha_j| \, dh \\
 &\leq C \int_0^{\tilde{r}_\varepsilon} \frac{(\cos(\frac{\pi}{12}))^{-1} h + \varepsilon}{\varepsilon} |\theta_{j,k}(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k})) - \alpha_j| \, dh \\
 &\leq C \int_0^{\tilde{r}_\varepsilon} \left(\frac{h^2}{\varepsilon^2} + \frac{h}{\varepsilon}\right) |(\theta_{j,k})'(-\varepsilon^{-1} \operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k}))| \, dh,
 \end{aligned}$$

where we integrated by parts and  $C > 0$  is a suitable constant varying from line to line. Once again, since  $\operatorname{dist}(P_{R_{j,k}^s}^- x, I_{j,k})$  is a homogeneous and increasing function with respect

to  $h$ , recalling the decay of  $|(\theta_{j,k})'(s)|$  in  $s$  mentioned above, then we get

$$\int_{W_j^\rho \setminus W_j^\varepsilon} \frac{1}{\varepsilon} |\theta_{j,k}(-\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k})) - \alpha_j| dx \leq C \varepsilon,$$

and also

$$\int_{W_j^\rho \setminus W_j^\varepsilon} \frac{1}{\varepsilon} |\theta_{i,j}(\varepsilon^{-1} \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j})) - \alpha_j| dx \leq C \varepsilon.$$

Similarly, we estimate

$$\begin{aligned} & \int_{W_j^\rho \setminus W_j^\varepsilon} \frac{1}{\varepsilon} W(\theta_{j,k}(\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k}))) dx \\ & \leq C \int_0^{\tilde{r}_\varepsilon} \frac{\tilde{h}_\varepsilon}{\varepsilon} W(\theta_{j,k}(-\varepsilon^{-1} \text{dist}(P_{R_{j,k}^-}^s x, I_{j,k}))) dh \leq C \varepsilon, \end{aligned}$$

due to the fact that  $\text{dist}(P_{R_{j,k}^-}^s x, I_{j,k})$  is a homogeneous and increasing function with respect to  $h$  and the decay of  $W(\theta_{j,k}(s))$  in  $s$ . Analogously, we have

$$\int_{M_j} \frac{1}{\varepsilon} W(\theta_{i,j}(\varepsilon^{-1} \text{dist}(P_{R_{i,j}^+}^s x, I_{i,j}))) dx \leq C \varepsilon.$$

Finally, (3.3) together with (A4) (see Section 2) and an application of Young's inequality give

$$\begin{aligned} \left| \sum_{\ell=1}^N \int_{W_j^\rho \setminus W_j^\varepsilon} \xi_\ell \cdot \nabla(\psi_\ell \circ u_{\varepsilon,0}) dx \right| & \leq C \int_{W_j^\rho \setminus W_j^\varepsilon} \sqrt{2W(u_{\varepsilon,0})} |\nabla u_{\varepsilon,0}| dx \\ & \leq C \left[ \int_{W_j^\rho \setminus W_j^\varepsilon} \frac{\varepsilon}{2} |\nabla u_{\varepsilon,0}|^2 dx + \int_{W_j^\rho \setminus W_j^\varepsilon} \frac{1}{\varepsilon} W(u_{\varepsilon,0}) dx \right]. \end{aligned}$$

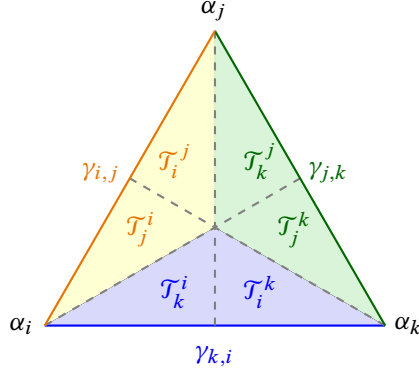
Then, from the estimates above, we can conclude that

$$E_{W_j^\rho \setminus W_j^\varepsilon}[u_\varepsilon | \xi](0) := \int_{W_j^\rho \setminus W_j^\varepsilon} \frac{\varepsilon}{2} |\nabla u_{\varepsilon,0}|^2 + \frac{1}{\varepsilon} W(u_{\varepsilon,0}) + \sum_{\ell=1}^N \xi_\ell \cdot \nabla(\psi_\ell \circ u_{\varepsilon,0}) dx \leq C \varepsilon$$

for some constant  $C > 0$ .

## 6. Suitable multi-well potentials and a construction for the $\psi_i$

We next proceed to show that the class of  $N$ -well potentials satisfying assumptions (A1)–(A4) is in fact sufficiently broad.



**Figure 3.** Illustration of the disjoint partition  $\{\mathcal{T}_j^i\}_{i,j:i \neq j}$  of the simplex  $\Delta^2$ .

### 6.1. A class of multi-well potentials

Let  $\Delta^{N-1}$  be an  $(N-1)$ -simplex with edges of unit length in  $\mathbb{R}^{N-1}$ . We denote by  $\{\gamma_{i,j}\}_{i,j:i \neq j}$  its edges and by  $\{\alpha_i\}_i$  its  $N$  vertices, so that  $|\alpha_i - \alpha_j| = 1$  for any mutually distinct  $i, j \in \{1, \dots, N\}$ . We decompose  $\Delta^{N-1}$  (almost) symmetrically into a disjoint partition  $\{\mathcal{T}_{i,j}\}_{i,j:i < j}$  such that each point  $x \in \Delta^{N-1}$  is assigned to the set  $\mathcal{T}_{i,j}$  if  $\gamma_{i,j}$  is the edge of  $\Delta^{N-1}$  that is closest to  $x$ , with  $x$  being assigned to the edge with the lowest  $i$  and  $j$  in the case of ties. Each  $\mathcal{T}_{i,j}$  can be further split nearly symmetrically into  $\mathcal{T}_j^i$  and  $\mathcal{T}_i^j$  by defining  $\mathcal{T}_j^i$  to consist of the points in  $\mathcal{T}_{i,j}$  that are closer to  $\alpha_i$  than to  $\alpha_j$ . For an illustration of this partition, we refer to Figure 3 for  $N = 3$ .

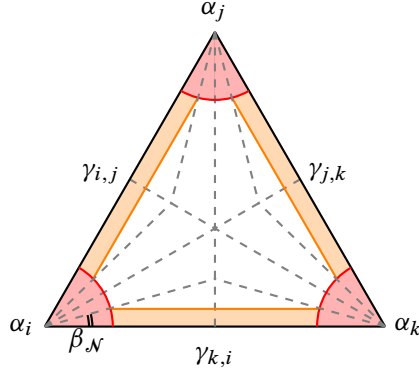
For the purpose of our construction of the  $\psi_i$  from condition (A4), we introduce some further notation:

- For  $i \in \{1, \dots, N\}$ , we denote by  $\mathcal{U}_i := \overline{B_{r_{\mathcal{U}}}(\alpha_i)}$  a ball around the vertex  $\alpha_i$  with radius  $r_{\mathcal{U}} \in (0, \frac{1}{4}]$ .
- For  $i < j, i, j \in \{1, \dots, N\}$ , we denote by  $\mathcal{N}_{i,j} := \{u \in \mathbb{R}^{N-1} : \text{dist}(u, \gamma_{i,j}) \leq r_{\mathcal{U}} \sin(\beta_{\mathcal{N}})\}$  a neighborhood of the edge  $\gamma_{i,j}$ . Here,  $\beta_{\mathcal{N}} \in (0, \frac{\pi}{12(N-2)}]$  is a fixed positive angle.

For a depiction of the resulting partition in the case  $N = 3$ , we refer to Figure 4. We furthermore make use of a couple of additional abbreviations.

- For any  $i < j, i, j \in \{1, \dots, N\}$ , we denote by  $P_{i,j}: \mathcal{T}_{i,j} \rightarrow \gamma_{i,j}$  the standard orthogonal projection onto  $\gamma_{i,j}$ , i.e. the projection onto the nearest point on  $\gamma_{i,j}$ .
- We denote by  $P_{i,j}^{\text{rad},i}: \mathcal{T}_{i,j} \rightarrow \gamma_{i,j}$  the radial projection onto  $\gamma_{i,j}$  with respect to  $\alpha_i$ , i.e.  $P_{i,j}^{\text{rad},i} u$  denotes the point on  $\gamma_{i,j}$  with  $|P_{i,j}^{\text{rad},i} u - \alpha_i| = |u - \alpha_i|$ .
- For any  $u \in \mathcal{T}_{i,j}$ , we denote by  $\beta_j^i(u)$  the angle formed by  $u - \alpha_i$  and  $\gamma_{i,j}$ .

For an illustration of these notions, we refer to Figure 5 (again in the case  $N = 3$ ).



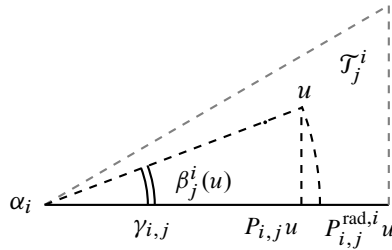
**Figure 4.** Illustration of the partition of the simplex  $\Delta^2$  given by  $\{\mathcal{U}_i \cap \Delta^2\}_i$ ,  $\{(\mathcal{N}_{i,j} \setminus \mathcal{U}_i) \cap \Delta^2\}_{i < j}$ , and  $\Delta^2 \setminus \mathcal{N}_{i,j}$ .

**Definition 17** (Strongly coercive  $N$ -well potential on the simplex). We call a function  $W: \Delta^{N-1} \rightarrow [0, \infty)$  a *strongly coercive symmetric  $N$ -well potential on the simplex* if it satisfies the following list of properties:

- (1) The nonnegative function  $W \in C^{1,1}(\Delta^{N-1}; [0, \infty))$  vanishes exactly in the  $N$  vertices  $\{\alpha_1, \dots, \alpha_N\}$  of the simplex  $\Delta^{N-1}$ . It furthermore has the same symmetry properties as the simplex  $\Delta^{N-1}$ .
- (2) Given the geodesic distance

$$\text{dist}_W(v, w) := \inf\left\{\int_0^1 \sqrt{2W(\gamma(s))} |\gamma'(s)| \, ds : \gamma \in C^1([0, 1]; \mathbb{R}^2) \text{ with } \gamma(0) = v, \gamma(1) = w\right\},$$

the infimum for  $\text{dist}_W(\alpha_i, \alpha_j)$  is achieved by  $\gamma_{i,j}$  and  $\text{dist}_W(\alpha_i, \alpha_j) = 1$  for any  $i, j \in \{1, \dots, N\}, i \neq j$ .



**Figure 5.** Projections of  $u \in \mathcal{T}_j^i$  onto  $\gamma_{i,j} \subset \Delta^2$ .

- (3) (Growth near the minima  $\alpha_i$  depending on the angle). For any distinct  $i, j \in \{1, \dots, N\}$  and any  $u \in \mathcal{U}_i \cap \mathcal{T}_{i,j}$ , we have the estimate

$$(1 + \omega(\beta_j^i(u)))W(P_{i,j}^{\text{rad},i}u) \leq W(u),$$

where  $\omega: [0, \frac{\pi}{6(N-2)}] \rightarrow [0, \infty)$  is a  $C^1$  increasing function such that

$$\omega(\beta) = 0 \quad \text{for } \beta = 0, \quad (6.1a)$$

$$\omega(\beta) \geq 0 \quad \text{for } \beta \in (0, \beta_{\mathcal{N}}), \quad (6.1b)$$

$$\omega(\beta) > C_\omega \quad \text{for } \beta \in \left[ \beta_{\mathcal{N}}, \frac{\pi}{6(N-2)} \right], \quad (6.1c)$$

where  $C_\omega > 0$  is a suitable large constant depending on  $N$  and where  $\beta_{\mathcal{N}} \leq \frac{\pi}{12(N-2)}$ .

- (4) (Growth properties of  $W$  and Lipschitz estimate for  $\sqrt{2W(u)}$  on the edges  $\gamma_{i,j}$ ). There exist constants  $c_\gamma, C_\gamma > 0$  such that

$$c_\gamma(u - \alpha_i)^2(u - \alpha_j)^2 \leq W(u) \leq C_\gamma(u - \alpha_i)^2(u - \alpha_j)^2 \quad (6.2)$$

holds for all  $u \in \gamma_{i,j}$  and any distinct  $i, j \in \{1, \dots, N\}$ . Furthermore, there exists a constant  $L_\gamma > 0$  such that for any  $u_1, u_2 \in \gamma_{i,j}$ ,

$$|\sqrt{2W(u_1)} - \sqrt{2W(u_2)}| \leq L_\gamma|u_1 - u_2|. \quad (6.3)$$

- (5) (Growth behavior as one leaves the shortest paths  $\gamma_{i,j}$ ). For any distinct  $i, j \in \{1, \dots, N\}$  and any  $u \in \mathcal{T}_{i,j} \cap (\mathcal{N}_{i,j} \setminus (\mathcal{U}_i \cup \mathcal{U}_j))$ , the lower bound

$$(1 + C_{\mathcal{N}} \text{dist}^2(u, \gamma_{i,j}))W(P_{i,j}u) \leq W(u) \quad (6.4)$$

holds, where  $C_{\mathcal{N}} > 0$  is a suitable large constant depending on  $r_{\mathcal{U}}, \beta_{\mathcal{N}}, L_\gamma, c_\gamma$ .

- (6) (Lower bound away from the paths  $\gamma_{i,j}$ ). For any  $u \in \Delta \setminus (\bigcup_i \mathcal{U}_i \bigcup_{i < j} \mathcal{N}_{i,j})$ ,

$$\max_{v \in \bigcup_{i < j} \gamma_{i,j}} W(v) \leq \frac{1}{C_{\text{int}}} W(u), \quad (6.5)$$

where  $C_{\text{int}} > 0$  is a suitable large constant depending on  $N, r_{\mathcal{U}}, \beta_{\mathcal{N}}, c_\gamma, C_\gamma$ .

Note that, by following the procedure sketched above Proposition 8, one may indeed construct a large variety of potentials  $W$  satisfying the conditions of Definition 17.

## 6.2. Construction of the approximate phase indicator functions $\psi_i$

In this subsection we provide an ansatz for the set of functions  $\psi_i: \Delta^{N-1} \rightarrow [0, 1]$ ,  $1 \leq i \leq N$ , in the case of a strongly coercive  $N$ -well potential on the simplex  $W: \Delta^{N-1} \rightarrow [0, \infty)$ . Recall that the goal is to construct the  $\psi_i$  to satisfy condition (A4) (as introduced Section 2).

Since  $(\psi_1 \circ u_\varepsilon, \dots, \psi_N \circ u_\varepsilon)$  represents an approximation of the limit partition  $(\bar{\chi}_1, \dots, \bar{\chi}_N)$  and since by assumption we have  $\text{dist}_W(\alpha_i, \alpha_j) = \delta_{ij}$ , our ansatz for  $\psi_i$  on the edges  $\gamma_{i,j}$  is

$$\psi_i(u_\varepsilon) := \begin{cases} 1 - \text{dist}_W(\alpha_i, u_\varepsilon) & \text{along } \gamma_{i,j} \text{ for } j \in \{1, \dots, N\} \setminus \{i\}, \\ 0 & \text{along } \gamma_{j,k} \text{ for } j, k \in \{1, \dots, N\} \setminus \{i\} : j < k. \end{cases}$$

In the following we extend this definition of the set of functions  $\psi_i$  on the domain  $\Delta^{N-1}$ . In order to do this, we introduce three interpolation and/or cutoff functions.

**Lemma 18** (Interpolation functions). *Let  $\beta_{\mathcal{N}} \in (0, \frac{\pi}{12(N-2)}]$  and  $r_{\mathcal{U}} \in (0, \frac{1}{4}]$ . The following statements hold:*

- (1) *There exists a function  $\lambda: [0, \frac{\pi}{6}] \rightarrow [0, 1]$  of at least  $C^1$  regularity satisfying the properties  $\lambda(\beta) = 0$  and  $\partial_\beta \lambda = 0$  for all  $\beta \in [0, \beta_{\mathcal{N}}]$ ,  $\lambda(\frac{\pi}{6(N-2)}) = 1$ , and*

$$\max_{\beta \in [0, \frac{\pi}{6(N-2)}]} |\partial_\beta \lambda| \leq 4(N-2). \quad (6.6)$$

- (2) *There exists a function  $\eta: [0, 1] \rightarrow [0, 1]$  of at least  $C^1$  regularity satisfying the properties  $\eta(s) = 1$  for  $s \in [0, r_{\mathcal{U}}]$ ,  $\eta(s) = 0$  for  $s \in [1 - r_{\mathcal{U}}, 1]$ ,  $\eta(s) + \eta(1-s) = 1$  for all  $s \in [0, 1]$ , and*

$$\max_{s \in [0, 1]} |\partial_s \eta| \leq \frac{5}{2}. \quad (6.7)$$

We omit the proof of the lemma, as it is straightforward. We finally proceed to the construction of the functions  $\psi_i$  from (A4) in the case of a strongly coercive  $N$ -well potential on the simplex.

**Construction 19.** *Let  $W: \Delta^{N-1} \rightarrow [0, \infty)$  be a strongly coercive symmetric  $N$ -well potential on the simplex in the sense of Definition 17. We define the associated set of functions  $\psi_i: \Delta \rightarrow [0, 1]$ ,  $1 \leq i \leq N$ , as follows. For  $i \in \{1, \dots, N\}$ , we construct  $\psi_i$  on the edge between  $\alpha_i$  and  $\alpha_j$  ( $j \in \{1, \dots, N\} \setminus \{i\}$ ) by*

$$\psi_i(u) := 1 - \text{dist}_W(\alpha_i, u) \quad \text{for } u \in \gamma_{i,j}. \quad (6.8)$$

Let  $j \in \{1, \dots, N\} \setminus \{i\}$ . For any  $u \in \mathcal{T}_j^i$ , we set

$$\psi_k(u) := 0 \quad \text{for any } k \in \{1, \dots, N\} \setminus \{i, j\}.$$

Furthermore, we define  $\psi_i$  and  $\psi_j$  on  $\mathcal{T}_j^i \cap (\mathcal{N}_{i,j} \cup \mathcal{U}_i)$  as follows:

- If  $u \in \mathcal{U}_i \cap \mathcal{T}_j^i$ , we set

$$\psi_i(u) := \psi_i(P_{i,j}^{\text{rad},i} u), \quad (6.9a)$$

$$\psi_j(u) := (1 - \lambda(\beta_j^i(u))) \psi_j(P_{i,j}^{\text{rad},i} u). \quad (6.9b)$$

- If  $u \in (\mathcal{N}_{i,j} \setminus \mathcal{U}_i) \cap \mathcal{T}_j^i$ , we set

$$\psi_i(u) := \eta(|P_{i,j}u - \alpha_i|)\psi_i(P_{i,j}^{\text{rad},i}u) + \eta(|P_{i,j}u - \alpha_j|)\psi_i(P_{i,j}^{\text{rad},j}u), \quad (6.10a)$$

$$\psi_j(u) := \eta(|P_{i,j}u - \alpha_i|)\psi_j(P_{i,j}^{\text{rad},i}u) + \eta(|P_{i,j}u - \alpha_j|)\psi_j(P_{i,j}^{\text{rad},j}u). \quad (6.10b)$$

Finally, outside the domain  $\mathcal{M}_i := \bigcup_j \mathcal{U}_j \cup \bigcup_{j:j \neq i} \mathcal{N}_{i,j} \cup \bigcup_{j,k:j,k \neq i, j < k} \mathcal{T}_{j,k}$  on which we have defined  $\psi_i$  so far, we define  $\psi_i$  as a suitable  $C^{1,1}$  extension:

- If  $u \in \Delta^{N-1} \setminus \mathcal{M}_i$ , we define

$$\psi_i(u_\varepsilon) := \psi_i^{\text{int}}(u) \quad (6.11)$$

where  $\psi_i^{\text{int}}: \Delta^{N-1} \rightarrow [0, 1]$  is a suitable  $C^{1,1}$  extension of  $\psi_i: \mathcal{M}_i \cap \Delta^{N-1} \rightarrow [0, 1]$  that almost preserves the Lipschitz constant of  $\psi_i: \mathcal{M}_i \cap \Delta^{N-1} \rightarrow [0, 1]$ .

### 6.3. Existence of a set of suitable approximate phase indicator functions

*Proof of Proposition 8.* It directly follows from Construction 19 that the set of functions  $\psi_i: \Delta \rightarrow [0, 1]$ ,  $1 \leq i \leq N$ , satisfy  $\psi_i = 1$  at  $\alpha_i$  and  $\psi_i(u) < 1$  for  $u \neq \alpha_i$ .

We next show the validity of (A4) in a given set  $\mathcal{T}_j^i$ , which we further decompose into  $\mathcal{U}_i \cap \mathcal{T}_j^i$ ,  $(\mathcal{N}_{i,j} \setminus \mathcal{U}_i) \cap \mathcal{T}_j^i$ , and  $\mathcal{T}_j^i \setminus (\mathcal{U}_i \cup \mathcal{N}_{i,j})$ .

*Step 1: Proof of (A4) in  $\mathcal{U}_i \cap \mathcal{T}_j^i$ .* Let  $u \in \mathcal{U}_i \cap \mathcal{T}_j^i$ . Recall  $\psi_0 := 1 - \sum_{\ell=1}^N \psi_\ell$ . Due to  $\text{dist}_W(\alpha_i, \alpha_j) = 1$  and (6.8), it follows that  $\psi_i(P_{i,j}^{\text{rad},i}u) = 1 - \psi_j(P_{i,j}^{\text{rad},i}u)$ . Thus, we have  $\psi_0(u) = \lambda(\beta_j^i(u))\psi_j(P_{i,j}^{\text{rad},i}u)$ . We also have  $(\frac{\alpha_j - \alpha_i}{|\alpha_j - \alpha_i|} \cdot \nabla)\psi_i(P_{i,j}^{\text{rad},i}u) = \sqrt{2W(P_{i,j}^{\text{rad},i}u)}$ . Using (6.9), we can compute

$$\begin{aligned} \partial_u \psi_{i,j}(u) &= (2 - \lambda(\beta_j^i(u)))\sqrt{2W(P_{i,j}^{\text{rad},i}u)}\vec{e}_{\text{rad},i}(u) \\ &\quad - \partial_\beta \lambda(\beta_j^i(u))\frac{1}{|u - \alpha_i|}\psi_j(P_{i,j}^{\text{rad},i}u)\vec{e}_{\beta_j^i}(u), \\ \partial_u \psi_0(u) &= \lambda(\beta_j^i(u))\sqrt{2W(P_{i,j}^{\text{rad},i}u)}\vec{e}_{\text{rad},i}(u) \\ &\quad + \partial_\beta \lambda(\beta_j^i(u))\frac{1}{|u - \alpha_i|}\psi_j(P_{i,j}^{\text{rad},i}u)\vec{e}_{\beta_j^i}(u), \end{aligned}$$

where  $\vec{e}_{\text{rad},i}(u)$ ,  $\vec{e}_{\beta_j^i}(u)$  are orthogonal vectors associated to the  $(N - 1)$ -dimensional spherical coordinates pointing in the direction of steepest ascent of  $|u - \alpha_i|$  resp.  $\beta_j^i(u)$ ; i.e. in particular we have  $\vec{e}_{\text{rad},i}(u) := \frac{u - \alpha_i}{|u - \alpha_i|}$ . For the sake of brevity, we omit the dependencies on  $\beta_j^i(u)$  in the following. Then it follows that

$$\begin{aligned} |\partial_u \psi_{i,j}(u)|^2 &\leq ((2 - \lambda)^2 + |\partial_\beta \lambda|^2)2W(P_{i,j}^{\text{rad},i}u), \\ |\partial_u \psi_0(u)|^2 &\leq (\lambda^2 + |\partial_\beta \lambda|^2)2W(P_{i,j}^{\text{rad},i}u), \\ |\partial_u \psi_{i,j}(u) \cdot \partial_u \psi_0(u)| &\leq (\lambda(2 - \lambda) + |\partial_\beta \lambda|^2)2W(P_{i,j}^{\text{rad},i}u). \end{aligned}$$

For  $\delta > 0$  small enough, we have

$$\begin{aligned}
 & \left| \frac{1}{2} \partial_u \psi_{i,j}(u) \right|^2 + \left( \frac{5}{4} + \delta \right) \left| \frac{1}{2} \partial_u \psi_0(u) \right|^2 + \delta |\partial_u \psi_{i,j}(u) \cdot \partial_u \psi_0(u)| \\
 & \leq \left[ \frac{1}{4} (2 - \lambda)^2 + \frac{1}{4} |\partial_\beta \lambda|^2 + \left( \frac{5}{4} + \delta \right) \frac{1}{4} (\lambda^2 + |\partial_\beta \lambda|^2) \right. \\
 & \quad \left. + \delta (\lambda(2 - \lambda) + |\partial_\beta \lambda|^2) \right] 2W(P_{i,j}^{\text{rad},i} u) \\
 & \leq \left[ 1 - \lambda + \frac{9}{16} \lambda^2 + \frac{9}{16} |\partial_\beta \lambda|^2 + C\delta \right] 2W(P_{i,j}^{\text{rad},i} u) \\
 & \leq (1 + \omega(\beta_j^i(u))) 2W(P_{i,j}^{\text{rad},i} u) \\
 & \leq 2W(u),
 \end{aligned}$$

where we used (6.6) and (6.1), together with the fact that  $\delta$  can be chosen arbitrarily small.

*Step 2: Proof of (A4) in  $u \in (\mathcal{N}_{i,j} \setminus \mathcal{U}_i) \cap \mathcal{T}_j^i$ .* Let  $u \in (\mathcal{N}_{i,j} \setminus \mathcal{U}_i) \cap \mathcal{T}_j^i$ . First, note that  $\psi_0 = 1 - \psi_i - \psi_j \equiv 0$  on  $(\mathcal{N}_{i,j} \setminus \mathcal{U}_i) \cap \mathcal{T}_j^i$ . Using (6.10), we compute

$$\begin{aligned}
 \partial_u \psi_{i,j}(u) &= \eta(|P_{i,j}u - \alpha_i|) 2\sqrt{2W(P_{i,j}^{\text{rad},i} u)} \vec{e}_{\text{rad},i}(u) \\
 & \quad - \eta(|P_{i,j}u - \alpha_j|) 2\sqrt{2W(P_{i,j}^{\text{rad},j} u)} \vec{e}_{\text{rad},j}(u) \\
 & \quad + \partial_u \eta(|P_{i,j}u - \alpha_i|) \\
 & \quad \times [\psi_j(P_{i,j}^{\text{rad},i} u) - \psi_j(P_{i,j}^{\text{rad},j} u) + \psi_i(P_{i,j}^{\text{rad},j} u) - \psi_i(P_{i,j}^{\text{rad},i} u)],
 \end{aligned}$$

where  $\vec{e}_{\text{rad},i}(u) := \frac{u - \alpha_i}{|u - \alpha_i|}$  and  $\vec{e}_{\text{rad},j}(u) := \frac{u - \alpha_j}{|u - \alpha_j|}$ . Note that we have

$$\begin{aligned}
 |P_{i,j}^{\text{rad},i} u - \alpha_i| &\leq |P_{i,j}u - \alpha_i| + \frac{\text{dist}^2(u, \gamma_{i,j})}{2|P_{i,j}u - \alpha_i|}, \\
 |P_{i,j}^{\text{rad},i} u - P_{i,j}^{\text{rad},j} u| &\leq \frac{\text{dist}^2(u, \gamma_{i,j})}{2|P_{i,j}u - \alpha_i| |P_{i,j}u - \alpha_j|}.
 \end{aligned} \tag{6.12}$$

As a consequence, we obtain

$$\begin{aligned}
 & \max \{ \psi_j(P_{i,j}^{\text{rad},i} u) - \psi_j(P_{i,j}^{\text{rad},j} u), \psi_i(P_{i,j}^{\text{rad},j} u) - \psi_i(P_{i,j}^{\text{rad},i} u) \} \\
 & \leq \sqrt{2W(v_u)} \frac{\text{dist}^2(u, \gamma_{i,j})}{2|P_{i,j}u - \alpha_i| |P_{i,j}u - \alpha_j|},
 \end{aligned}$$

where  $v_u \in \gamma_{i,j}$  is the maximum of  $\sqrt{2W}$  on the segment connecting  $P_{i,j}^{\text{rad},j} u$  to  $P_{i,j}^{\text{rad},i} u$ . From (6.7) and  $\eta(|P_{i,j}u - \alpha_i|) + \eta(|P_{i,j}u - \alpha_j|) = 1$  it follows that

$$\left| \frac{1}{2} \partial_u \psi_{i,j}(u) \right|^2 \leq \left( 1 + \frac{5}{4|P_{i,j}u - \alpha_i| |P_{i,j}u - \alpha_j|} \text{dist}^2(u, \gamma_{i,j}) \right)^2 2W(v_u). \tag{6.13}$$



Using  $1 - |P_{i,j}u - \alpha_i| = |P_{i,j}u - \alpha_j|$  and  $|P_{i,j}u - \alpha_i|^{-1}(1 - |P_{i,j}u - \alpha_i|)^{-1} \leq r_{\mathcal{U}}^{-1}(1 - r_{\mathcal{U}})^{-1}$ , one can obtain

$$\left(1 + \frac{5}{4|P_{i,j}u - \alpha_i||P_{i,j}u_{\varepsilon} - \alpha_j|} \text{dist}^2(u, \gamma_{i,j})\right)^2 \leq 1 + C_1 \text{dist}^2(u, \gamma_{i,j})$$

for  $C_1 = \frac{5}{2r_{\mathcal{U}}(1-r_{\mathcal{U}})} + \frac{25 \sin^2(\beta_{\mathcal{N}})}{16(1-r_{\mathcal{U}})^2}$  since  $\text{dist}(u, \gamma_{i,j}) \leq r_{\mathcal{U}} \sin(\beta_{\mathcal{N}})$ . On the other hand, first by adding a zero and using (6.3), then noting that  $|v_u - \alpha_i| \leq |P_{i,j}^{\text{rad},j}u - \alpha_i|$  and using (6.12), from (6.2) together with the fact that  $|P_{i,j}u - \alpha_i| \geq \frac{1}{2}$  we can deduce

$$\begin{aligned} 2W(v_u) &= 2W(P_{i,j}u) \left(1 + \frac{\sqrt{2W(v_u)} - \sqrt{2W(P_{i,j}u)}}{\sqrt{2W(P_{i,j}u)}}\right)^2 \\ &\leq 2W(P_{i,j}u) \left(1 + \frac{L_{\gamma}}{2|P_{i,j}u - \alpha_i|\sqrt{2W(P_{i,j}u)}} \text{dist}^2(u, \gamma_{i,j})\right)^2 \\ &\leq 2W(P_{i,j}u) \left(1 + \frac{L_{\gamma}}{\sqrt{2c_{\gamma}}} \tan^2(\beta_j^i(u))\right)^2 \\ &\leq 2W(P_{i,j}u)(1 + C_2 \tan^2(\beta_j^i(u))), \end{aligned}$$

where  $C_2 = 2 \frac{L_{\gamma}}{\sqrt{2c_{\gamma}}} + \tan^2(\beta_{\mathcal{N}}) \frac{L_{\gamma}^2}{2c_{\gamma}}$ . Moreover, one can compute

$$\begin{aligned} &(1 + C_1 \text{dist}^2(u, \gamma_{i,j}))(1 + C_2 \tan^2(\beta_j^i(u))) \\ &\leq 1 + C_1 \text{dist}^2(u, \gamma_{i,j}) + \frac{C_2}{r_{\mathcal{U}}^2 \cos^2 \beta_{\mathcal{N}}} \text{dist}^2(u, \gamma_{i,j}) + C_1 C_2 \tan^2(\beta_{\mathcal{N}}) \text{dist}^2(u, \gamma_{i,j}) \\ &\leq 1 + C_{\mathcal{N}} \text{dist}^2(u, \gamma_{i,j}) \end{aligned}$$

for  $C_{\mathcal{N}} = C_1 + \frac{C_2}{r_{\mathcal{U}}^2 \cos^2 \beta_{\mathcal{N}}} + C_1 C_2 \tan^2(\beta_{\mathcal{N}})$ . Using our assumption (6.4), we can conclude from (6.13) and the preceding three estimates that

$$\begin{aligned} \left|\frac{1}{2} \partial_u \psi_{i,j}(u)\right|^2 &\leq (1 + C_{\mathcal{N}} \text{dist}^2(u, \gamma_{i,j})) 2W(P_{i,j}u) \\ &\leq 2W(u). \end{aligned}$$

*Step 3: Proof of (A4) in  $u \in \mathcal{T}_j^i \setminus \mathcal{M}_i$ .* Let  $u \in \mathcal{T}_j^i \setminus \mathcal{M}_i$ . By (6.11) we have

$$\begin{aligned} \psi_{i,j}(u_{\varepsilon}) &= \psi_j^{\text{int}}(u) - \psi_i^{\text{int}}(u), \\ \psi_0(u_{\varepsilon}) &= 1 - \psi_i^{\text{int}}(u) - \psi_j^{\text{int}}(u), \end{aligned}$$

where  $\psi_{\ell}^{\text{int}}$ ,  $\ell \in \{i, j\}$ , is a  $C^{1,1}$  extension of  $\psi_{\ell}$  from  $\mathcal{M}_{\ell} \cap \Delta^{N-1}$  to  $\Delta^{N-1}$  that approximately preserves the Lipschitz constant  $L_{\text{int},\ell} > 0$ . Thus, we have

$$|\partial_u \psi_{i,j}| \leq (1 + \delta)(L_{\text{int},i} + L_{\text{int},j}), \quad \text{and similarly} \quad |\partial_u \psi_0| \leq (1 + \delta)(L_{\text{int},i} + L_{\text{int},j}),$$

where  $\delta > 0$  is an arbitrary small constant. It is not too difficult to derive an estimate on the Lipschitz constants  $L_{\text{int},i}$  and  $L_{\text{int},j}$  in terms of  $\max_{w \in \bigcup_{\ell,m:\ell < m} \mathcal{V}_{\ell,m}} 2W(w)$ . To this aim, we first estimate

$$\begin{aligned} |\psi_i(u)| &\leq \max_{w \in \mathcal{Y}_{i,m}} \sqrt{2W(w)} r_{\mathcal{U}} \quad \text{for } u \in \mathcal{Y}_{i,m} \cap \mathcal{U}_m \text{ and } m \neq i, \\ |\psi_i(P_{i,m}^{\text{rad},i} u) - \psi_\ell(P_{i,m}^{\text{rad},m} u)| &\leq \max_{w \in \mathcal{Y}_{i,m}} \sqrt{2W(w)} |P_{i,m}^{\text{rad},i} u - P_{i,m}^{\text{rad},m} u| \\ &\stackrel{(6.12)}{\leq} \max_{w \in \mathcal{Y}_{i,m}} \sqrt{2W(w)} \frac{r_{\mathcal{U}}}{2(1-r_{\mathcal{U}})} \sin^2(\beta_{\mathcal{N}}) \quad \text{for } u \in \mathcal{N}_{i,m} \text{ and } m \neq i. \end{aligned}$$

Using these estimates, the definitions (6.9)–(6.10), and the bounds (6.6) and (6.7), we obtain

$$\begin{aligned} |\partial_u \psi_i(u)| &\leq \max_{w \in \mathcal{Y}_{i,j}} \sqrt{2W(w)} && \text{for } u \in \mathcal{U}_i, \\ |\partial_u \psi_i(u)| &\leq (1 + 4(N-2)r_{\mathcal{U}}) \max_{w \in \mathcal{Y}_{i,m}} \sqrt{2W(w)} && \text{for } u \in \mathcal{U}_m, m \neq i, \\ |\partial_u \psi_i(u)| &\leq \left(1 + \frac{5r_{\mathcal{U}}}{4(1-r_{\mathcal{U}})} \sin^2(\beta_{\mathcal{N}})\right) \max_{w \in \mathcal{Y}_{i,m}} \sqrt{2W(w)} && \text{for } u \in \mathcal{N}_{i,m}, m \neq i. \end{aligned}$$

Furthermore, we have  $|\partial_u \psi_i(u)| = 0$  in  $\Delta^{N-1} \cap (\bigcup_{m < n: m \neq i, n \neq i} \mathcal{T}_{m,n})$ . Defining

$$M := \max\left\{1, 1 + 4(N-2)r_{\mathcal{U}}, 1 + \frac{5r_{\mathcal{U}}}{4(1-r_{\mathcal{U}})} \sin^2(\beta_{\mathcal{N}})\right\} = 1 + 4(N-2)r_{\mathcal{U}},$$

this yields, by (6.5) for  $u \in \mathcal{M}_i \cap \Delta^{N-1}$ ,

$$|\partial_u \psi_i(u)| \leq M \max_{w \in \bigcup_{\ell,m:\ell < m} \mathcal{V}_{\ell,m}} \sqrt{2W(w)} =: M_W$$

for any  $u \in \Delta^{N-1} \cap \mathcal{M}_i$ . In order to estimate the Lipschitz constant  $L_{\text{int},i}$  of  $\psi_i|_{\mathcal{M}_i \cap \Delta^{N-1}}$ , one has to address the issue of nonconvexity of  $\mathcal{M}_i$ . It is not too difficult to see (but rather technical to prove) that for any pair of points  $u, v \in \mathcal{M}_i$  there exists a connecting path  $\tilde{\gamma}$  in  $\mathcal{M}_i$  with  $\text{len}(\tilde{\gamma}) \leq C_{\mathcal{M}}|u - v|$ . This shows  $L_{\text{int},i} \leq C_{\mathcal{M}}M_W$ . Having an upper bound for  $L_{\text{int},i}$ , using the fact that our extension of  $\psi_\ell$  to  $\Delta^{N-1} \setminus \mathcal{M}_i$  approximately preserves the Lipschitz constant, and choosing  $C_{\text{int}} > \frac{9}{4}C_{\mathcal{M}}^2M^2$  in (6.5), we can compute for  $u \in \mathcal{T}_j^i \setminus \mathcal{M}_i$ ,

$$\begin{aligned} &\left|\frac{1}{2}\partial_u \psi_{i,j}(u)\right|^2 + \left(\frac{5}{4} + \delta\right) \left|\frac{1}{2}\partial_u \psi_0(u)\right|^2 + \delta |\partial_u \psi_{i,j}(u) \cdot \partial_u \psi_0(u)| \\ &\leq (1 + \delta)^2 \left(1 + \frac{5}{4} + 5\delta\right) \max_m L_{\text{int},m}^2 \leq 2W(u). \end{aligned}$$

Here, we have used the fact that  $\delta > 0$  can be chosen arbitrarily small. ■

## A. Additional length condition for the gradient-flow calibration

In this appendix we justify the validity of condition (2.3i) for the gradient-flow calibration constructed in [13] (cf. Definition 6). To this aim, first we show that (2.3i) holds on the network of interfaces, and then we motivate the extension of the property (2.3i) to  $\mathbb{R}^d$ . Similarly, one can check that (2.3i) also holds for the gradient-flow calibration for three-dimensional double bubbles (cf. [15]).

The global gradient-flow calibration for a network is obtained by gluing together suitable local constructions at each topological feature, i.e. a two phase interface or a triple junction (for more details see [13]). More precisely, a partition of unity is defined in order to localize around each topological feature, and then the global vector fields are defined by gluing together locally constructed vector fields. In the following, we denote by  $\xi_{i,j}^{I_{ij}}$  resp.  $\xi_\ell^{\mathcal{T}_{ijk}}$  the local construction of the gradient-flow calibration in a neighborhood of a single connected component of  $I_{ij}$  resp. of the triple junction  $\mathcal{T}_{i,j,k}$  (cf. [13, Sections 5–6]).

As a starting point, we recall some useful properties of the gradient-flow calibration. For any distinct  $i, j \in \{1, \dots, N\}$ , we have

$$\xi_{i,j} = \xi_{i,j}^{I_{ij}} = \xi_i^{I_{ij}} - \xi_j^{I_{ij}} = \bar{n}_{i,j}, \quad \xi_i^{I_{ij}} = \frac{1}{2}\xi_{i,j}^{I_{ij}},$$

and  $\xi_k = \xi_k^{I_{ij}}$  coincides with the null vector for any  $k \in \{1, \dots, N\} \setminus \{i, j\}$  on  $I_{i,j}$  outside a neighborhood of the triple junctions where  $I_{i,j}$  ends. Let  $\mathcal{T}_{i,j,k}$  be the triple junction where the phases  $i, j$ , and  $k$  meet, for mutually distinct  $i, j, k \in \{1, \dots, N\}$ . Then, at the triple junction  $\mathcal{T}_{i,j,k}$ ,

$$\xi_\ell = \xi_\ell^{\mathcal{T}_{ijk}}, \quad |\xi_\ell| = \frac{1}{\sqrt{3}}, \quad \xi_\ell \cdot \xi_m = -\frac{1}{6} \quad \text{for any distinct } \ell, m \in \{i, j, k\},$$

and  $\xi_\ell = \xi_\ell^{\mathcal{T}_{ijk}}$  coincides with the null vector for any  $\ell \in \{1, \dots, N\} \setminus \{i, j, k\}$ .

Let  $i, j \in \{1, \dots, N\}$  such that  $i \neq j$ . If we restrict to the network of interfaces, the vector field  $\xi_{i,j}$  is nonzero only on each connected component of  $I_{i,j}, I_{i,k}, I_{j,k}$  for any  $k \in \{1, \dots, N\} \setminus \{i, j\}$  and at the triple junctions where either the phase  $i$  or the phase  $j$  meets other phases. Using the properties of the gradient-flow calibration listed above, one can easily see that (2.3i) is satisfied at each point belonging to a connected component of  $I_{i,j}, I_{i,k}$ , or  $I_{j,k}$  and outside a neighborhood of the triple junctions where it ends, as well as at the triple junctions where either the phase  $i$  or the phase  $j$  meets other phases. As a consequence, in order to conclude about (2.3i) on the network of interface, we only need to check its validity in a neighborhood of triple junctions where either the phase  $i$  or  $j$  meets other phases. In particular, it suffices to consider a neighborhood of  $\mathcal{T}_{i,j,k}$  and a neighborhood of  $\mathcal{T}_{i,k,k'}$  for some distinct  $k, k' \in \{1, \dots, N\} \setminus \{i, j\}$ .

In a neighborhood of  $\mathcal{T}_{i,j,k}$ , one can write

$$\begin{aligned} \xi_{i,j} &= \eta \xi_{i,j}^{\mathcal{T}_{ijk}} + (1 - \eta) \xi_{i,j}^{I_{ij}}, & \xi_k &= \eta \xi_k^{\mathcal{T}_{ijk}} & \text{along } I_{i,j}, \\ \xi_{i,j} &= \eta \xi_{i,j}^{\mathcal{T}_{ijk}} + (1 - \eta) \xi_{i,j}^{I_{ik}}, & \xi_k &= \eta \xi_k^{\mathcal{T}_{ijk}} + (1 - \eta) \xi_k^{I_{ik}} & \text{along } I_{i,k}, \end{aligned}$$

where  $\eta: \mathbb{R}^d \rightarrow [0, 1]$  is a cutoff function decreasing quadratically with distance from  $\mathcal{T}_{i,j,k}$  and vanishing outside a neighborhood of  $\mathcal{T}_{i,j,k}$ . Note that  $\xi_{i,j}^{\mathcal{T}_{ijk}} = \xi_{i,j}^{I_{ij}}$  along  $I_{i,j}$  in a neighborhood of  $\mathcal{T}_{i,j,k}$ . It follows that  $|\xi_{i,j}|^2 = 1$  and  $\xi_{i,j} \cdot \xi_k = 0$  for  $k \in \{1, \dots, N\} \setminus \{i, j\}$  (cf. [13, Section 7.2]), hence (2.3i) holds along  $I_{i,j}$  in a neighborhood of  $\mathcal{T}_{i,j,k}$ . As a next step, we compute  $|\xi_{i,j}|^2$  and  $|\sqrt{3}\xi_{i,j} \cdot \xi_k|^2$  for  $k \in \{1, \dots, N\} \setminus \{i, j\}$ , and then show that (2.3i) with  $\delta_{\text{cal}} = 0$  holds along  $I_{i,k}$  in a neighborhood of  $\mathcal{T}_{i,j,k}$ . Since  $\xi_{i,j}^{I_{ik}} = \xi_{i,j}^{I_{ik}} = \frac{1}{2}\xi_{i,k}^{I_{ik}} = \frac{1}{2}\xi_{i,k}^{\mathcal{T}_{ijk}}$  and  $\xi_k^{I_{ik}} = -\frac{1}{2}\xi_{i,k}^{I_{ik}} = -\frac{1}{2}\xi_{i,k}^{\mathcal{T}_{ijk}}$  along  $I_{i,k}$  in a neighborhood of  $\mathcal{T}_{i,j,k}$ , one can deduce

$$\begin{aligned} |\xi_{i,j}|^2 &= \eta^2 |\xi_{i,j}^{\mathcal{T}_{ijk}}|^2 + (1-\eta)^2 \left| \frac{1}{2} \xi_{i,k}^{\mathcal{T}_{ijk}} \right|^2 + \eta(1-\eta) \xi_{i,j}^{\mathcal{T}_{ijk}} \cdot \xi_{i,k}^{\mathcal{T}_{ijk}} \\ &= \eta^2 + (1-\eta)^2 \frac{1}{4} + \eta(1-\eta) \frac{1}{2}, \\ \xi_{i,j} \cdot \xi_k &= -\frac{1}{2} \eta(1-\eta) \xi_{i,j}^{\mathcal{T}_{ijk}} \cdot \xi_{i,k}^{\mathcal{T}_{ijk}} + \frac{1}{2} \eta(1-\eta) \xi_{i,k}^{\mathcal{T}_{ijk}} \cdot \xi_k^{\mathcal{T}_{ijk}} - \frac{1}{4} (1-\eta)^2 |\xi_{i,k}^{\mathcal{T}_{ijk}}|^2 \\ &= -\frac{1}{2} \eta(1-\eta) - \frac{1}{4} (1-\eta)^2, \end{aligned}$$

where we used  $\xi_{i,j}^{\mathcal{T}_{ijk}} \cdot \xi_k^{\mathcal{T}_{ijk}} = 0$  in a neighborhood of  $\mathcal{T}_{i,j,k}$  (for more details see [13, Section 7.2]). As a consequence, one can see that

$$|\xi_{i,j}|^2 + 4|\sqrt{3}\xi_{i,j} \cdot \xi_k|^2 \leq 1,$$

thus (2.3i) with  $\delta_{\text{cal}} = 0$  is satisfied along  $I_{i,k}$  in a neighborhood of  $\mathcal{T}_{i,j,k}$ .



As a next step, we consider a neighborhood of  $\mathcal{T}_{i,k,k'}$ , where we can write

$$\begin{aligned} \xi_{i,j} &= \tilde{\eta} \xi_{i,j}^{\mathcal{T}_{ikk'}} + (1-\tilde{\eta}) \xi_{i,j}^{I_{ik}}, \quad \xi_k = \tilde{\eta} \xi_k^{\mathcal{T}_{ikk'}} + (1-\tilde{\eta}) \xi_k^{I_{ik}}, \quad \xi_{k'} = \tilde{\eta} \xi_{k'}^{\mathcal{T}_{ikk'}} \quad \text{along } I_{i,k}, \\ \xi_{i,j} &= \tilde{\eta} \xi_{i,j}^{\mathcal{T}_{ikk'}} + (1-\tilde{\eta}) \xi_{i,j}^{\mathcal{T}_{ikk'}}, \quad \xi_k = \tilde{\eta} \xi_k^{\mathcal{T}_{ikk'}} + (1-\tilde{\eta}) \xi_k^{I_{kk'}}, \quad \xi_{k'} = \tilde{\eta} \xi_{k'}^{\mathcal{T}_{ikk'}} + (1-\tilde{\eta}) \xi_{k'}^{I_{kk'}} \quad \text{along } I_{k,k'}, \end{aligned}$$

where  $\tilde{\eta}: \mathbb{R}^d \rightarrow [0, 1]$  denotes a cutoff function decreasing quadratically with the distance from  $\mathcal{T}_{i,k,k'}$  and vanishing outside the neighborhood of  $\mathcal{T}_{i,k,k'}$ . Moreover, we have  $\xi_{i,j}^{I_{ik}} = \xi_{i,j}^{I_{ik}} = \frac{1}{2}\xi_{i,k}^{I_{ik}} = \frac{1}{2}\xi_{i,k}^{\mathcal{T}_{ikk'}}$  and  $\xi_k^{I_{ik}} = -\frac{1}{2}\xi_{i,k}^{I_{ik}} = -\frac{1}{2}\xi_{i,k}^{\mathcal{T}_{ikk'}}$  along  $I_{i,k}$ , as well as  $\xi_{k'}^{I_{kk'}} = \frac{1}{2}\xi_{k,k'}^{I_{kk'}} = \frac{1}{2}\xi_{k,k'}^{\mathcal{T}_{ikk'}}$  and  $\xi_{k'}^{I_{kk'}} = -\frac{1}{2}\xi_{k,k'}^{I_{kk'}} = -\frac{1}{2}\xi_{k,k'}^{\mathcal{T}_{ikk'}}$  along  $I_{k,k'}$ . Proceeding as above (cf. [13, Section 7.2]), one can compute  $|\xi_{i,j}|^2$  and  $|\sqrt{3}\xi_{i,j} \cdot \xi_k|^2$  for  $k \in \{1, \dots, N\} \setminus \{i, j\}$ , and then show that (2.3i) holds along both  $I_{i,k}$  and  $I_{k,k'}$  in a neighborhood of  $\mathcal{T}_{i,k,k'}$ . In particular, observe that on the network (2.3i) is satisfied with  $\delta_{\text{cal}} = 0$ .

Finally, observe that the gradient-flow calibration satisfies first-order compatibility conditions at any triple junction (cf. [13]). As a consequence, property (2.3i) can be extended to  $\mathbb{R}^d$  up to errors of second order in the distance with respect to the network. Hence, if one additionally truncates the gradient-flow calibration constructed in [13] by means of an additional cutoff function decreasing quadratically with the distance from the network, then condition (2.3i) follows for an arbitrarily small  $\delta_{\text{cal}} > 0$ .

**Acknowledgments.** The authors thank Sebastian Hensel for useful and helpful comments on the first draft of this work.

**Funding.** 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)  .

## References

- [1] H. Abels and Y. Liu, [Sharp interface limit for a Stokes/Allen–Cahn system](#). *Arch. Ration. Mech. Anal.* **229** (2018), no. 1, 417–502 Zbl 1394.35343 MR 3799097
- [2] H. Abels and M. Moser, [Convergence of the Allen–Cahn equation to the mean curvature flow with  \$90^\circ\$ -contact angle in 2D](#). *Interfaces Free Bound.* **21** (2019), no. 3, 313–365 Zbl 1430.35127 MR 4014391
- [3] N. D. Alikakos, P. W. Bates, and X. Chen, [Convergence of the Cahn–Hilliard equation to the Hele–Shaw model](#). *Arch. Rational Mech. Anal.* **128** (1994), no. 2, 165–205 Zbl 0828.35105 MR 1308851
- [4] F. Bethuel, G. Orlandi, and D. Smets, [Convergence of the parabolic Ginzburg–Landau equation to motion by mean curvature](#). *Ann. of Math. (2)* **163** (2006), no. 1, 37–163 Zbl 1103.35038 MR 2195132
- [5] L. Bronsard and R. V. Kohn, [Motion by mean curvature as the singular limit of Ginzburg–Landau dynamics](#). *J. Differential Equations* **90** (1991), no. 2, 211–237 Zbl 0735.35072 MR 1101239
- [6] L. Bronsard and F. Reitich, [On three-phase boundary motion and the singular limit of a vector-valued Ginzburg–Landau equation](#). *Arch. Rational Mech. Anal.* **124** (1993), no. 4, 355–379 Zbl 0785.76085 MR 1240580
- [7] X. Chen, [Generation and propagation of interfaces for reaction–diffusion equations](#). *J. Differential Equations* **96** (1992), no. 1, 116–141 Zbl 0765.35024 MR 1153311
- [8] X. Chen, [Global asymptotic limit of solutions of the Cahn–Hilliard equation](#). *J. Differential Geom.* **44** (1996), no. 2, 262–311 Zbl 0874.35045 MR 1425577
- [9] P. De Mottoni and M. Schatzman, [Geometrical evolution of developed interfaces](#). *Trans. Amer. Math. Soc.* **347** (1995), no. 5, 1533–1589 Zbl 0840.35010 MR 1672406
- [10] L. C. Evans, H. M. Soner, and P. E. Souganidis, [Phase transitions and generalized motion by mean curvature](#). *Comm. Pure Appl. Math.* **45** (1992), no. 9, 1097–1123 Zbl 0801.35045 MR 1177477
- [11] M. Fei and Y. Liu, [Phase-field approximation of the Willmore flow](#). *Arch. Ration. Mech. Anal.* **241** (2021), no. 3, 1655–1706 Zbl 1476.35021 MR 4284532
- [12] J. Fischer and S. Hensel, [Weak–strong uniqueness for the Navier–Stokes equation for two fluids with surface tension](#). *Arch. Ration. Mech. Anal.* **236** (2020), no. 2, 967–1087 Zbl 1465.35332 MR 4072686
- [13] J. Fischer, S. Hensel, T. Laux, and T. Simon, [The local structure of the energy landscape in multiphase mean curvature flow: Weak–strong uniqueness and stability of evolutions](#). arXiv:2003.05478, to appear in *J. Eur. Math. Soc. (JEMS)*

- [14] J. Fischer, T. Laux, and T. M. Simon, [Convergence rates of the Allen–Cahn equation to mean curvature flow: A short proof based on relative entropies](#). *SIAM J. Math. Anal.* **52** (2020), no. 6, 6222–6233 Zbl [1454.53079](#) MR [4187143](#)
- [15] S. Hensel and T. Laux, [Weak–strong uniqueness for the mean curvature flow of double bubbles](#). *Interfaces Free Bound.* **25** (2023), no. 1, 37–107 Zbl [1520.53081](#) MR [4578339](#)
- [16] S. Hensel and A. Marveggio, [Weak–strong uniqueness for the Navier–Stokes equation for two fluids with ninety degree contact angle and same viscosities](#). *J. Math. Fluid Mech.* **24** (2022), no. 3, article no. 93 Zbl [1496.35279](#) MR [4462807](#)
- [17] S. Hensel and M. Moser, [Convergence rates for the Allen–Cahn equation with boundary contact energy: The non-perturbative regime](#). *Calc. Var. Partial Differential Equations* **61** (2022), no. 6, article no. 201 Zbl [1503.53166](#) MR [4473117](#)
- [18] T. Ilmanen, [Convergence of the Allen–Cahn equation to Brakke’s motion by mean curvature](#). *J. Differential Geom.* **38** (1993), no. 2, 417–461 Zbl [0784.53035](#) MR [1237490](#)
- [19] R. L. Jerrard and D. Smets, [On the motion of a curve by its binormal curvature](#). *J. Eur. Math. Soc. (JEMS)* **17** (2015), no. 6, 1487–1515 Zbl [1327.53086](#) MR [3353807](#)
- [20] R. L. Jerrard and H. M. Soner, [Scaling limits and regularity results for a class of Ginzburg–Landau systems](#). *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **16** (1999), no. 4, 423–466 Zbl [0944.35006](#) MR [1697561](#)
- [21] T. Laux and Y. Liu, [Nematic-isotropic phase transition in liquid crystals: A variational derivation of effective geometric motions](#). *Arch. Ration. Mech. Anal.* **241** (2021), no. 3, 1785–1814 Zbl [1468.76008](#) MR [4284534](#)
- [22] T. Laux and T. M. Simon, [Convergence of the Allen–Cahn equation to multiphase mean curvature flow](#). *Comm. Pure Appl. Math.* **71** (2018), no. 8, 1597–1647 Zbl [1393.35122](#) MR [3847750](#)
- [23] F. H. Lin, [Complex Ginzburg–Landau equations and dynamics of vortices, filaments, and codimension-2 submanifolds](#). *Comm. Pure Appl. Math.* **51** (1998), no. 4, 385–441 Zbl [0932.35121](#) MR [1491752](#)
- [24] C. Mantegazza, M. Novaga, A. Pluda, and F. Schulze, [Evolution of networks with multiple junctions](#). arXiv:[1611.08254](#), to appear in *Astérisque*
- [25] P. Sternberg, [The effect of a singular perturbation on nonconvex variational problems](#). *Arch. Rational Mech. Anal.* **101** (1988), no. 3, 209–260 Zbl [0647.49021](#) MR [930124](#)

Received 23 September 2022; revised 11 October 2023; accepted 2 November 2023.

### **Julian Fischer**

Institute of Science and Technology Austria (IST Austria), Am Campus 1, 3400 Klosterneuburg, Austria; [julian.fischer@ist.ac.at](mailto:julian.fischer@ist.ac.at)

### **Alice Marveggio**

Institute of Science and Technology Austria (IST Austria), Am Campus 1, 3400 Klosterneuburg, Austria; [alice.marveggio@ist.ac.at](mailto:alice.marveggio@ist.ac.at), [alice.marveggio@hcm.uni-bonn.de](mailto:alice.marveggio@hcm.uni-bonn.de)