Math. Ann. (2016) 364:857–935
DOI 10.1007/s00208-015-1237-5

Mathematische Annalen

Existence and regularity of mean curvature flow with transport term in higher dimensions

Keisuke Takasao[1] · **Yoshihiro Tonegawa**[2]

Received: 18 August 2013 / Revised: 16 April 2015 / Published online: 12 June 2015

Abstract Given an initial C^1 hypersurface and a time-dependent vector field in a Sobolev space, we prove a time-global existence of a family of hypersurfaces which start from the given hypersurface and which move by the velocity equal to the mean curvature plus the given vector field. We show that the hypersurfaces are C^1 for a short time and, even after some singularities occur, almost everywhere C^1 away from the higher multiplicity region.

Mathematics Subject Classification 53C44 · 35K55 · 49Q15

1 Introduction

A family $\{M_t\}_{t\geq 0}$ of hypersurfaces in $\mathbb{R}^n$ is called mean curvature flow (MCF) if the velocity vector v of M_t is equal to its mean curvature vector h at each point and time, that is,

K. Takasao is supported by the JSPS Research Fellowships for Young Scientists and the JSPS Grant-in-Aid for the JSPS fellows 12J06075 and Y. Tonegawa is partially supported by JSPS Grant-in-aid for scientific research 25247008, 26220702.

✉ Yoshihiro Tonegawa
tonegawa@math.titech.ac.jp

Keisuke Takasao
takasao@ms.u-tokyo.ac.jp

[1] Department of Mathematics, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan

[2] Department of Mathematics, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8551, Japan

$$v = h \quad \text{on} \;\; M_t. \tag{1.1}$$

As one of the fundamental geometric evolution problems, the MCF has been studied by numerous researchers in the past few decades. One of many facets of investigations is the time-global existence question of such a family when given an initial hypersurface M_0. In general dimensions, there exists a unique smooth family of MCF for finite time until singularities such as vanishing and pinch-off occur. Though the classical MCF ceases to exist at this point, it is well-known that a unique time-global solution $\{M_t\}_{t\geq 0}$ exists in a weak viscosity sense [11,16] despite the occurrence of singularities.

In this paper, we are interested in an aspect of time-global existence theory for a related problem, and the question we ask is the following. Given an initial hypersurface M_0 and a vector field u, is there a family $\{M_t\}_{t\geq 0}$ of hypersurfaces whose velocity vector v is equal to its mean curvature h plus u? What is the minimum regularity assumption on u for the existence and regularity of such a family? To be more precise, since we would be interested in the normal velocity to see the motion, the requirement is

$$v = h + (u \cdot \nu)\nu \quad \text{on} \;\; M_t \tag{1.2}$$

where ν is the unit normal vector field of M_t and $\cdot$ is the inner product in $\mathbb{R}^n$. Motivation to investigate (1.2) is more than just to see what happens when an extra lower order term is added. While the MCF is of premier importance, one wonders what is the limit of applicability of various analytic techniques developed for the MCF if one puts a wild perturbation. In a reverse context, if one understands the limit of generality of the MCF, then some of the analytic techniques developed for more general settings may be useful for the MCF. In fact, our investigation on (1.2) has already led us to the development of a local regularity theory [30,46] which gives new insight to the MCF. Physically, one may regard (1.2) as a surface tension driven phase boundary motion with a given background transport effect such as fluid flow or external force field. One can also find such motion law in a coupled system with the Navier-Stokes equation modeling a flow of dry foam (see, for example, [31] for the numerical simulation and references therein).

Though far from complete, in this paper we obtain satisfactory time-global existence and regularity theorems if we assume that M_0 is C^1 and u satisfies

$$\left(\int_0^T \left(\int_{\mathbb{R}^n} |u(x,t)|^p + |\nabla u(x,t)|^p \, dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}} < \infty \tag{1.3}$$

for all $T < \infty$, with $2 < q < \infty$ and $\frac{nq}{2(q-1)} < p < \infty$ ($\frac{4}{3} \leq p$ in addition if $n = 2$). Here $\nabla u = (\partial_{x_1} u, \ldots, \partial_{x_n} u)$ is the weak partial derivatives and $u, \nabla u$ are measurable with the stated integrability. We prove that the hypersurfaces remain C^1 at least for a short time, and it is a.e. C^1 away from a region where M_t develops higher multiplicities. With more regularity assumption on u such as Hölder continuity, we have C^2 instead of C^1 and (1.2) is satisfied classically. For the precise statement of the regularity, see Theorem 2.3.

Here we briefly discuss our approach. If u is regular enough with respect to x, for example Lipschitz continuous, the level set method approach works well with a good order preserving property (see, for example, [22] and [20, Sect. 4.8]). Also for regular enough u, there are a number of short time existence results which are often stated for the MCF but which can be extended to include regular u: (1) solving an evolution equation for the height function from the reference initial manifold [10], (2) solving equations for signed distance function [17] (and elaborated further in [21]), and (3) constructing an approximate solution by time-discrete minimal movement [3], just to name a few examples. On the other hand, with irregular u, one can not expect the order preserving property in general and even the short time existence of solution can be a serious issue. Hence to characterize (1.2), we take an approach pioneered by Brakke [6] using the notion of varifold from geometric measure theory. To construct a sequence of approximate solutions, we use the Allen-Cahn equation [2] with an extra transport term coming from u, (3.5). Much of the analysis of the present paper concerns various ε-independent estimates of quantities associated with φ_ε. We obtain a desired solution by taking a limit $\varepsilon \to 0$. Thus the interest of the present paper can be also the analysis of (3.5) itself. Once we verify that the limit satisfies (1.2) in a weak sense of varifold as in Brakke's formulation, we apply a local regularity theory developed in [30,46] which is tailor-made for the present problem. To our knowledge, under the assumption (1.3) of u, even the short time existence of C^1 solution seems new.

As for the MCF in general, there are a number of books and papers some of which include up-to-date research results on the subject and we mention [4,5,12,14,20,35, 48]. Concerning a time-global existence for the MCF and the related problems, we mention [3,6,11,16,29,34] and references therein. While there are numerous works with varying generalities establishing the connection between the Allen-Cahn equation and the MCF (for example [7,9,13,15,19,39]), analysis of the Allen-Cahn equation using geometric measure theory was pioneered by Ilmanen [28] in which he proved that the limit surface measures are rectifiable and satisfy (1.1) in the sense of Brakke's formulation. The second author proved that the limit surface measures are integral [45]. There are a number of closely related works even if we restrict the scope within some measure theoretic approach to the Allen-Cahn equation, and we further mention [37,40,42,43] and references therein. The existence result of the present paper has been proved by Liu et al. [33] for $n = 2, 3$ and with more restrictive assumptions on p and q. The limitation of the dimensions was due to the use of results by Röger and Schätzle [38], which gives a characterization of limit measures under an assumption of uniform L^2 bound of mean curvature-like quantity. In the present paper, we avoid using [38], and we follow the line of proofs of [28,45] combined with various estimates from [33]. This frees us from any dimensional restriction. As a special case, the first author investigated the graph-like problem of (1.2) with a better regularity assumption on u and showed a unique short time existence [44].

The paper is organized as follows. In Sect. 2 we set our notations and explain the main results. In Sect. 3 we briefly discuss some heuristic aspects of the Allen-Cahn equation. Section 4 deals with the uniform upper density ratio bound and monotonicity formula, and this is the key to control the transport term subsequently. In Sect. 5, we show that there exists a limit surface measure for all $t \geq 0$. Section 6 proves that the limit measure is rectifiable and this part owes much to Ilmanen's work [28]. In Sect. 7, we prove that the limit measure has integer density modulo surface energy constant. There, the idea of proof goes back to [27] and the parabolic version [45]. In Sect. 8 we prove the main results by combining all the results from previous four sections. We record our final remarks in the last Sect. 9. We intended the paper to be as self-contained as possible, only exception being the proof for regularity. There we cite the main local regularity theorem which has a set of assumptions we need to check.

2 Preliminaries and main results

2.1 Basic notation

Let $\mathbb{N}$ be the set of natural numbers and $\mathbb{R}^+ := \{x \geq 0\}$. For $0 < r < \infty$ and $a \in \mathbb{R}^k$ define

$$B^k_r(a) := \{x \in \mathbb{R}^k \ : \ |x - a| < r\}.$$

We write $B^k_r := B^k_r(0)$. When $k = n$, we omit writing n. We often identify $\mathbb{R}^{n-1}$ with $\mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$. On $\mathbb{R}^n$ we denote the Lebesgue measure by $\mathcal{L}^n$ and for $0 \leq k \leq n$, the k-dimensional Hausdorff measure by $\mathcal{H}^k$. Define $\omega_n := \mathcal{L}^n(B_1)$. Given a set $A \subset \mathbb{R}^n$ and a measure μ, the restriction of μ to A is denoted by $\mu\lfloor_A$. The characteristic function of A is denoted by χ_A. Symbol ∇ always refers to a differentiation with respect to the space variables. For a set of finite perimeter (see [24] for the definition) A, we denote the total variation measure of the distributional derivative $\nabla\chi_A$ by $\|\nabla\chi_A\|$.

Throughout the paper, we set Ω to be either $\mathbb{T}^n$, the n-dimensional unit torus, or $\mathbb{R}^n$. For $\Omega = \mathbb{T}^n$ we often regard Ω as the unit square $[0, 1) \times \cdots \times [0, 1) \subset \mathbb{R}^n$ where all the relevant quantities are extended periodically to the entire $\mathbb{R}^n$. Objects such as functions and sets in Ω are understood implicitly in this manner. For any Radon measure μ on $\mathbb{R}^n$ and $\phi \in C_c(\mathbb{R}^n)$ we often write $\mu(\phi)$ for $\int \phi \, d\mu$. We write $\operatorname{spt}\mu$ for the support of μ. Thus $x \in \operatorname{spt}\mu$ if $\forall r > 0$, $\mu(B_r(x)) > 0$. For $1 \leq p \leq \infty$, we write $f \in L^p(\mu)$ if f is μ measurable and $(\int |f|^p \, d\mu)^{1/p} < \infty$. We use the standard notation for Sobolev spaces such as $W^{1,p}(\Omega)$ and $W^{1,p}_{loc}(\Omega)$ from [23].

For $A, B \in \operatorname{Hom}(\mathbb{R}^n; \mathbb{R}^n)$ which we identify with $n \times n$ matrices, we define

$$A \cdot B := \sum_{i,j} A_{ij} B_{ij} \quad \text{and} \quad |A| := \sqrt{A \cdot A}.$$

$\|A\|$ denotes the operator norm. The identity of $\operatorname{Hom}(\mathbb{R}^n; \mathbb{R}^n)$ is denoted by I. For $k \in \mathbb{N}$ with $k < n$, let $\mathbf{G}(n, k)$ be the space of k-dimensional subspaces of $\mathbb{R}^n$. The orthogonal complement of $S \in \mathbf{G}(n, k)$ is denoted by $S^\perp \in \mathbf{G}(n, n-k)$. For $a \in \mathbb{R}^n$, $a \otimes a \in \operatorname{Hom}(\mathbb{R}^n; \mathbb{R}^n)$ is the matrix with the entries $a_i a_j$ $(1 \leq i, j \leq n)$. For $S \in \mathbf{G}(n, k)$, we identify S with the corresponding orthogonal projection of $\mathbb{R}^n$ onto S. In the case of $k = n-1$, we also identify $S \in \mathbf{G}(n, n-1)$ with the unit vector $\pm\nu \in \mathbb{S}^{n-1}$ which is perpendicular to S. Note that we may express the relation by

$S = I - \nu \otimes \nu$. The correspondence is a homeomorphism with respect to the naturally endowed topologies on $\mathbf{G}(n, n-1)$ and $\mathbb{S}^{n-1}/\{\pm 1\}$. For $x, y \in \mathbb{R}^n$ and $t < s$ define

$$\rho_{(y,s)}(x, t) := \frac{1}{(4\pi(s-t))^{\frac{n-1}{2}}} e^{-\frac{|x-y|^2}{4(s-t)}}, \tag{2.1}$$

which is the backward heat kernel with pole at (y, s).

2.2 Varifolds

We recall some definitions from geometric measure theory and refer to [1,6,41] for more details. For any open set $U \subset \mathbb{R}^n$ let $G_k(U) := U \times \mathbf{G}(n, k)$. A general k-varifold in U is a Radon measure on $G_k(U)$. We denote the set of all general k-varifolds in U by $\mathbf{V}_k(U)$. For $V \in \mathbf{V}_k(U)$, let $\|V\|$ be the weight measure of V, namely,

$$\|V\|(\phi) := \int_{G_k(U)} \phi(x)\, dV(x, S), \quad \forall \phi \in C_c(U).$$

We say $V \in \mathbf{V}_k(U)$ is rectifiable if there exist a $\mathcal{H}^k$ measurable countably k-rectifiable set $M \subset U$ and a locally $\mathcal{H}^k$ integrable function θ defined on M such that

$$V(\phi) = \int_M \phi(x, \mathrm{Tan}_x M)\theta(x)\, d\mathcal{H}^k \tag{2.2}$$

for $\phi \in C_c(G_k(U))$. Here $\mathrm{Tan}_x M$ is the approximate tangent space of M at x which exists $\mathcal{H}^k$ a.e. on M. Rectifiable k-varifold is uniquely determined by its weight measure $\|V\| = \theta\, \mathcal{H}^{n-1} \lfloor_M$ through the formula (2.2). For this reason, we naturally say a Radon measure μ on U is rectifiable when one can associate a rectifiable varifold V such that $\|V\| = \mu$. If $\theta \in \mathbb{N}$, $\mathcal{H}^k$ a.e. on M, we say V is integral. The set of all integral k-varifolds in U is denoted by $\mathbf{IV}_k(U)$. If $\theta = 1$, $\mathcal{H}^k$ a.e. on M, we say V is a unit density k-varifold.

For $V \in \mathbf{V}_k(U)$ let δV be the first variation of V, namely,

$$\delta V(g) := \int_{G_k(U)} \nabla g(x) \cdot S\, dV(x, S) \tag{2.3}$$

for $g \in C_c^1(U\, ;\, \mathbb{R}^n)$. If the total variation $\|\delta V\|$ of δV is locally bounded and absolutely continuous with respect to $\|V\|$, by the Radon-Nikodym theorem, we have a $\|V\|$ measurable vector field $h(V, \cdot)$ with

$$\delta V(g) = -\int_U g(x) \cdot h(V, x)\, d\|V\|(x). \tag{2.4}$$

The vector field $h(V, \cdot)$ is called the generalized mean curvature vector of V. For any $V \in \mathbf{IV}_k(U)$ with an integrable $h(V, \cdot)$, Brakke's perpendicularity theorem [6, Chapter 5] says that we have

$$\int_U (\mathrm{Tan}_x M)^{\perp}(g(x)) \cdot h(V,x)\, d\|V\|(x) = \int_U g(x) \cdot h(V,x)\, d\|V\|(x) \tag{2.5}$$

for all $g \in C_c(U;\mathbb{R}^n)$. Here, M is related to V as in (2.2). In the case of $k = n-1$, note that $(\mathrm{Tan}_x M)^{\perp} = \nu(x) \otimes \nu(x)$ for $\|V\|$ a.e. in U, where $\nu(x)$ is the unit normal vector to $\mathrm{Tan}_x M$. With this notation (2.5) may be written as

$$\int_U (g(x)\cdot\nu(x))(h(V,x)\cdot\nu(x))\, d\|V\|(x) = \int_U g(x)\cdot h(V,x)\, d\|V\|(x) \tag{2.6}$$

for $g \in C_c(U;\mathbb{R}^n)$. If $h(V,\cdot) \in L^2(\|V\|)$, by approximation, (2.6) holds even for $g \in L^2(\|V\|)$.

2.3 Weak formulation of velocity

Let $\{M_t\}_{t\geq 0}$ be a family of smooth hypersurfaces in Ω whose normal velocity is denoted by v. To formulate the velocity in a weak sense, observe the following characterization of v: a smooth normal vector field $\tilde{v}$ on M_t is equal to v if and only if

$$\frac{d}{dt}\int_{M_t} \phi\, d\mathcal{H}^{n-1} \leq \int_{M_t} (\nabla\phi - h\phi)\cdot\tilde{v} + \partial_t\phi\, d\mathcal{H}^{n-1} \tag{2.7}$$

holds for all $\phi \in C_c^1(\Omega\times[0,\infty);\mathbb{R}^+)$ and for all $t \geq 0$. Here h is the classical mean curvature vector of M_t. To check this claim, after some calculation, one first sees that v satisfies (2.7) with equality. Conversely, if $\tilde{v}$ satisfies (2.7), and already knowing that v satisfies (2.7) with equality, we obtain

$$0 \leq \int_{M_t} (\nabla\phi - h\phi)\cdot(\tilde{v} - v)\, d\mathcal{H}^{n-1}$$

for $\phi \in C_c^1(\Omega;\mathbb{R}^+)$. For any $\hat{x} \in M_t$ and $\lambda > 0$, let $\phi_\lambda(y) := \lambda^{2-n}\phi(\lambda^{-1}(y-\hat{x}))$. Substitute ϕ_λ and let $\lambda \downarrow 0$. Since $\lambda^{-1}(M_t - \hat{x}) \to \mathrm{Tan}_{\hat{x}} M_t$, we obtain

$$0 \leq \int_{\mathrm{Tan}_{\hat{x}} M_t} \nabla\phi\, d\mathcal{H}^{n-1}\cdot(\tilde{v}(\hat{x}) - v(\hat{x})).$$

The integration by parts shows $\int_{\mathrm{Tan}_{\hat{x}} M_t} \nabla\phi\, d\mathcal{H}^{n-1} \perp \mathrm{Tan}_{\hat{x}} M_t$. On the other hand, one may choose this vector to be $-(\tilde{v}(\hat{x}) - v(\hat{x}))$, for example. Thus we have $\tilde{v}(\hat{x}) = v(\hat{x})$ and we complete the proof of the claim. The characterization (2.7) motivates the following definition.

Definition 2.1 A family of varifolds $\{V_t\}_{t\geq 0} \subset \mathbf{V}_{n-1}(\Omega)$ is a generalized solution of (1.2) if the following four conditions are satisfied.

(a) $V_t \in \mathbf{IV}_{n-1}(\Omega)$ for a.e. $t \geq 0$.

(b) For all $T > 0$,

$$\sup_{t\in[0,T]} \|V_t\|(\Omega) < \infty \quad \text{and} \quad \sup_{t\in[0,T],\, B_r(x)\subset\Omega} \frac{\|V_t\|(B_r(x))}{\omega_{n-1}r^{n-1}} < \infty. \tag{2.8}$$

(c) For all $T > 0$,

$$\int_0^T dt \int_\Omega |h|^2 + |u|^2 \, d\|V_t\| < \infty. \tag{2.9}$$

(d) For all $\phi \in C_c^1(\Omega \times [0,\infty); \mathbb{R}^+)$ and $0 \le t_1 < t_2 < \infty$,

$$\|V_t\|(\phi(\cdot,t))\Big|_{t=t_1}^{t_2} \le \int_{t_1}^{t_2} dt \int_\Omega (\nabla\phi - h\phi)\cdot\{h + (u\cdot\nu)\nu\} + \partial_t\phi \, d\|V_t\| \tag{2.10}$$

holds, where we abbreviated $h(V_t, x)$ by h.

The condition (b) may appear out of place in the definition of velocity. In fact, if u is 0 or a bounded function and if $\|V_0\|$ satisfies (2.8), one can derive (2.8) as a consequence of (2.10) via Huisken's monotonicity formula. However, if u is not bounded, it is not clear how to obtain (2.8) from (2.10). The other important point is that, unless one has (2.8), it is unclear how to make sense of (2.9) and (2.10). The difficulty is, $u(\cdot,t)$ needs to be defined as a $\|V_t\|$ measurable function for a.e. $t \ge 0$. In general, $u(\cdot,t)$ is assumed to be in some Sobolev space on Ω, and we need to define $\|V_t\|$ measurable $u(\cdot,t)$ as a trace function. If we have (2.8), we may define the trace using the following inequality.

Theorem 2.1 *For a Radon measure μ on $\mathbb{R}^n$ with $D := \sup_{B_r(x)\subset\mathbb{R}^n} \frac{\mu(B_r(x))}{\omega_{n-1}r^{n-1}}$ and $1 \le p < n$,*

$$\int_{\mathbb{R}^n} |\phi|^{\frac{p(n-1)}{n-p}} \, d\mu \le c(n,p) D \left(\int_{\mathbb{R}^n} |\nabla\phi|^p \, dx\right)^{\frac{n-1}{n-p}} \tag{2.11}$$

holds for $\phi \in C_c^1(\mathbb{R}^n)$.

See [36] and [49] for the proof in the case of $p = 1$. The above inequality for $1 < p < n$ may be derived by the Hölder and Sobolev inequalities.

Suppose that we have (2.8). We only need to define u as a function in $L^2_{loc}(\|V_t\| \times dt)$ to make sense of (2.9) and (2.10). Since $W^{1,p'}_{loc} \subset W^{1,p}_{loc}$ if $p' > p$, we need to consider only $1 \le p < n$. Using the Hölder inequality and (2.11), we obtain (with $D := \sup_{B_r(x)\subset\Omega} \frac{\|V_t\|(B_r(x))}{\omega_{n-1}r^{n-1}}$)

$$\begin{aligned}\int_\Omega |\phi|^2 \, d\|V_t\| &\le \left(\int_\Omega |\phi|^{\frac{p(n-1)}{n-p}} \, d\|V_t\|\right)^{\frac{2(n-p)}{p(n-1)}} (\|V_t\|(\operatorname{spt}\phi))^{\frac{pn+p-2n}{p(n-1)}} \\ &\le (c(n,p)D)^{\frac{2(n-p)}{p(n-1)}} \left(\int_\Omega |\nabla\phi|^p \, dx\right)^{\frac{2}{p}} (\|V_t\|(\operatorname{spt}\phi))^{\frac{pn+p-2n}{p(n-1)}}.\end{aligned} \tag{2.12}$$

for $\phi \in C_c^1(\Omega)$. Here, we also need to assume that

$$p \geq \frac{2n}{n+1} \tag{2.13}$$

so that $\frac{p(n-1)}{n-p} \geq 2$. Since we will assume (2.14) in the next subsection, which implies $p > \frac{n}{2}$ in particular, (2.13) will be relevant only for $n = 2$ and we will assume $p \geq \frac{4}{3}$ when $n = 2$. With this restriction, we may define u as an $L^2_{loc}(\|V_t\| \times dt)$ function on $\Omega \times [0, T]$ uniquely as long as $u \in L^2_{loc}([0, \infty); (W^{1,p}_{loc}(\Omega))^n)$ by the standard density argument. The function u in (2.9) and (2.10) is defined in this sense.

2.4 Main results

First we present some existence result for (1.2) when given a vector field u and an initial hypersurface M_0.

Theorem 2.2 *Suppose $n \geq 2$,*

$$2 < q < \infty, \quad \frac{nq}{2(q-1)} < p < \infty \quad \left(\frac{4}{3} \leq p \ \ \textit{in addition if} \ \ n = 2\right) \tag{2.14}$$

and $\Omega = \mathbb{R}^n$ or $\mathbb{T}^n$. Given any

$$u \in L^q_{loc}([0, \infty); (W^{1,p}(\Omega))^n) \tag{2.15}$$

and a non-empty bounded domain $\Omega_0 \subset \Omega$ with C^1 boundary $M_0 = \partial\Omega_0$, there exist

(1) *a family of varifolds $\{V_t\}_{t \geq 0} \subset \mathbf{V}_{n-1}(\Omega)$ which is a generalized solution of (1.2) as in Definition 2.1 with $\|V_0\| = \mathcal{H}^{n-1}\lfloor_{M_0}$ and*

(2) *a function $\varphi \in BV_{loc}(\Omega \times [0, \infty)) \cap C^{\frac{1}{2}}_{loc}([0, \infty); L^1(\Omega))$ with the following properties.*

  (2a) *$\varphi(\cdot, t)$ is a characteristic function for all $t \in [0, \infty)$,*

  (2b) *$\|\nabla\varphi(\cdot, t)\|(\phi) \leq \|V_t\|(\phi)$ for all $t \in [0, \infty)$ and $\phi \in C_c(\Omega; \mathbb{R}^+)$,*

  (2c) *$\varphi(\cdot, 0) = \chi_{\Omega_0}$ a.e. on Ω,*

  (2d) *writing $\|V_t\| = \theta\mathcal{H}^{n-1}\lfloor_{M_t}$ and $\|\nabla\varphi(\cdot, t)\| = \mathcal{H}^{n-1}\lfloor_{\tilde{M}_t}$ for a.e. $t > 0$, we have*

$$\mathcal{H}^{n-1}(\tilde{M}_t \backslash M_t) = 0 \tag{2.16}$$

and

$$\theta(x, t) = \begin{cases} \textit{even integer} \ \geq 2 & \textit{if} \ x \in M_t \backslash \tilde{M}_t, \\ \textit{odd integer} \ \geq 1 & \textit{if} \ x \in \tilde{M}_t \end{cases} \tag{2.17}$$

for $\mathcal{H}^{n-1}$ a.e. $x \in M_t$.

(3) *If $p < n$, then for any $T > 0$, setting $s := \frac{p(n-1)}{n-p}$, we have*

$$\left(\int_0^T \left(\int_\Omega |u|^s \, d\|V_t\|\right)^{\frac{q}{s}} dt\right)^{\frac{1}{q}} < \infty. \tag{2.18}$$

If $p = n$, then we have (2.18) locally for $U \subset\subset \Omega$ for any $2 \le s < \infty$ and if $p > n$, then we have (2.18) with L^s norm replaced by $C^{1-\frac{n}{p}}$ norm on Ω.

(4) *There exists $T_1 > 0$ such that V_t has unit density for a.e. $t \in [0, T_1)$. In addition $\|\nabla\varphi(\cdot, t)\| = \|V_t\|$ for a.e. $t \in [0, T_1)$.*

The condition (2.14) on u is a dimensionally sharp condition in the following sense. Consider a natural parabolic change of variables $\tilde{x} := \frac{x}{\lambda}$ and $\tilde{t} := \frac{t}{\lambda^2}$ with $\lambda > 0$. Since u is a velocity field, it should behave just like x/t, thus it is natural to consider $\tilde{u} := \lambda u$. Then we have

$$\left(\int_0^\infty \left(\int_{\mathbb{R}^n} |\nabla u|^p \, dx\right)^{\frac{q}{p}} dt\right)^{\frac{1}{q}} = \lambda^{\frac{n}{p}+\frac{2}{q}-2}\left(\int_0^\infty \left(\int_{\mathbb{R}^n} |\nabla \tilde{u}|^p \, d\tilde{x}\right)^{\frac{q}{p}} d\tilde{t}\right)^{\frac{1}{q}}$$

and $\frac{n}{p} + \frac{2}{q} - 2 < 0$ is equivalent to the second inequality in (2.14). This guarantees that u locally behaves more like a perturbative term. In (3), if $p > n$, then the result follows from the standard Sobolev inequality on $\mathbb{R}^n$.

To understand what V_t and φ are, assume for a moment that no singular behaviors occur and we have a smooth family $\{M_t\}_{t\ge0}$ with the velocity given by (1.2). Then we should have $\text{spt}\,\|V_t\| = \partial\{\varphi(\cdot, t) = 1\} = M_t$. Since (1.2) is stated in terms of V_t, it may first appear that φ is redundant. However, beside the fact that φ is obtained naturally from the approach of the present paper, it has a few important roles. First, φ helps to guarantee that V_t is non-trivial. Since $\varphi(\cdot, t)$ is continuous in $L^1(\Omega)$ by (2), $\|\varphi(\cdot, t)\|_{L^1(\Omega)}$ cannot vanish instantaneously at some arbitrary time. As long as $\varphi(\cdot, t)$ is not identically zero or identically 1, $\|V_t\|$ is non-zero measure. Note that, given arbitrary $t_0 > 0$, by re-defining $V_t := 0$ for all $t > t_0$, we obtain another generalized solution of (1.2) due to the inequality in (2.10). Obviously, this is not a solution we would like to obtain in the end. The second role of φ is that it gives some restriction on the possible singularities of $\text{spt}\,\|V_t\|$. For example, consider in the $n = 2$ case. One can see that a unit density V_t cannot form a triple junction since $\partial\{\varphi(\cdot, t) = 1\}$ cannot be a triple junction. Thus, having φ as an auxiliary object may be a useful tool to obtain some better regularity results. As for the actual occurrence of the higher multiplicities, Bronsard and Stoth [8] showed that one can have solution with $\theta \ge 2$ for a limit of the Allen-Cahn equation, thus we may indeed have such solution in general.

We next state the regularity property of $\text{spt}\,\|V_t\|$, which is obtained as an application of [30,46]. To state the result, we recall some definitions from there.

Definition 2.2 A point $x \in \text{spt}\,\|V_t\|$ is said to be a $C^{1,\zeta}$ regular point if there exists some open neighborhood O in $\mathbb{R}^{n+1}$ containing (x, t) such that $O \cap \cup_{s>0}(\text{spt}\,\|V_s\| \times \{s\})$ is an embedded n-dimensional manifold with $C^{1,\zeta}$ regularity in space and $C^{(1+\zeta)/2}$ regularity in time. Similarly, we define a $C^{2,\alpha}$ regular point by replacing the respective regularities by $C^{2,\alpha}$ in space and $C^{1,\alpha/2}$ in time.

Theorem 2.3 *Let $\{V_t\}_{t\geq 0}$ be as in Theorem 2.2 .*

(1) *Suppose that there exist an open set $U \subset \Omega$ and an interval (t_1, t_2) such that V_t is unit density in U for a.e. $t \in (t_1, t_2)$. Then for a.e. $t \in (t_1, t_2)$, there exists a closed set $G_t \subset U$ with $\mathcal{H}^{n-1}(G_t) = 0$ such that $(U \cap \operatorname{spt}\|V_t\|)\setminus G_t$ is a set of $C^{1,\zeta}$ regular points where $\zeta := 2 - \frac{n}{p} - \frac{2}{q}$ if $p < n$. If $p \geq n$, one may take any ζ with $0 < \zeta < 1 - \frac{2}{q}$.*

(2) *There exists $T_2 > 0$ such that every point of $\operatorname{spt}\|V_t\|$ is a $C^{1,\zeta}$ regular point for all $t \in (0, T_2)$ (that is, $G_t = \emptyset$), where ζ is as in (1).*

(3) *If u is Hölder continuous with exponent α in the parabolic sense, i.e.,*

$$\sup_{\Omega\times[0,T]} |u| + \sup_{x,y\in\Omega, 0\leq t_1<t_2\leq T} \frac{|u(x,t_1) - u(y,t_2)|}{\max\{|x-y|^\alpha, |t_1 - t_2|^{\alpha/2}\}} < \infty$$

for all $0 < T < \infty$, then the same results for (1) and (2) hold true with $C^{1,\zeta}$ there replaced by $C^{2,\alpha}$ and (1.2) is satisfied pointwise.

(4) *We have $\lim_{t\downarrow 0} t^{-\frac{1}{2}} \operatorname{dist}(M_0, \operatorname{spt}\|V_t\|) = 0$ and $\operatorname{spt}\|V_t\|$ converges to M_0 in C^1 topology as $t \downarrow 0$. Namely, given $\varepsilon > 0$ there exists a finite number of sets $\{U_i = x_i + O_i(B_r^{n-1} \times (-r, r))\}_{i=1}^N$, where O_i is an orthogonal rotation and $x_i \in M_0$, such that $M_0 \subset \cup_{i=1}^N U_i$, and C^1 norms of difference of graphs representing M_0 and $\operatorname{spt}\|V_t\|$ over $x_i + O_i(B_r^{n-1})$ in U_i are less than ε for all sufficiently small $t > 0$.*

The claim (1) says that wherever V_t is unit density in some space-time neighborhood, $\operatorname{spt}\|V_t\|$ is locally a hypersurface with regularity of $C^{1,\zeta}$ in space and $C^{(1+\zeta)/2}$ in time, almost everywhere in space and time. We can guarantee by (2) that there is some time interval $[0, T_2)$ such that $\operatorname{spt}\|V_t\|$ is a $C^{1,\zeta}$ hypersurface. We obtain a lower bound on T_2 in terms of M_0 and the norm of u. On the other hand, T_2 may be much larger than the lower bound and it is the time when a non-$C^{1,\zeta}$ regular point occurs for the first time. In general, $T_2 \leq T_1$ and it is plausible that some non-$C^{1,\zeta}$ regular point first appears at T_2 but V_t may remain unit density for some more time. The claim (4) shows that $\operatorname{spt}\|V_t\|$ has C^1 uniform regularity and convergence as $t \downarrow 0$. As for (3), we first note that we can show the same existence results for Hölder continuous u (and not in $L^q_{loc}([0,\infty); (W^{1,p}(\Omega))^n)$) as in Theorem 2.2. In fact the proof is simpler if u is bounded. $C^{2,\alpha}$ regularity allows one to have pointwise mean curvature vector and velocity vector of $\operatorname{spt}\|V_t\|$ and (1.2) is satisfied pointwise. At this point, we reach a well-defined PDE setting, and $\operatorname{spt}\|V_t\|$ is as regular as what the standard parabolic regularity theory shows depending on any additional regularity assumption imposed on u.

3 Allen-Cahn equation with transport term

As stated in the introduction, the method of proof for the existence is to approximate (1.2) by the Allen-Cahn equation with an extra transport term coming from u. Throughout the paper, we assume that a function W satisfies the following:

$$W : \mathbb{R} \to [0, \infty) \text{ is } C^3 \quad \text{and} \quad W(\pm 1) = W'(\pm 1) = 0. \tag{3.1}$$

$$\text{For some } \gamma \in (-1, 1),\ W' < 0 \ \text{ on } \ (\gamma, 1) \quad \text{and} \quad W' > 0 \ \text{ on } \ (-1, \gamma). \tag{3.2}$$

$$\text{For some } \alpha \in (0, 1) \quad \text{and} \quad \kappa > 0,\ W''(x) \geq \kappa \ \text{ for all } \ 1 \geq |x| \geq \alpha. \tag{3.3}$$

We also define a constant

$$\sigma := \int_{-1}^{1} \sqrt{2W(s)}\, ds. \tag{3.4}$$

Basically, above assumptions require W to be W-shaped with non-degenerate two minima at ± 1. Requiring (3.2) may appear non-essential, but it is used essentially in deriving an upper bound for ξ_ε in Lemma 4.2. Any such W satisfying above can be used. The reader can take a concrete example such as $W(s) = (1 - s^2)^2$ in the following.

Given u and M_0 as in Theorem 2.2, the whole scheme of the present paper is to approximate the motion law (1.2) by

$$\partial_t \varphi_\varepsilon + u_\varepsilon \cdot \nabla \varphi_\varepsilon = \Delta \varphi_\varepsilon - \frac{W'(\varphi_\varepsilon)}{\varepsilon^2}, \tag{3.5}$$

where $\varepsilon > 0$ is a small parameter tending to 0 and u_ε is a smooth approximation of u. For readers who are not familiar with the Allen-Cahn equation, we give a quick heuristic argument. Assume that u is smooth and that we have a family of domains Ω_t with smooth boundaries $M_t = \partial \Omega_t$. Let $d(\cdot, t)$ be the signed distance function to M_t so that $d(\cdot, t) > 0$ inside of Ω_t. We let $\Psi : \mathbb{R} \to (-1, 1)$ be an ODE solution of $\Psi'' = W'(\Psi)$ with $\lim_{x \to \pm\infty} \Psi(x) = \pm 1$. Such solution exists and we may assume $\Psi(0) = 0$. If we postulate that $\varphi_\varepsilon(x, t) \approx \Psi(d(x, t)/\varepsilon)$ and φ_ε satisfies (3.5), then we expect that

$$\Psi' \partial_t d + u_\varepsilon \cdot \Psi' \nabla d \approx \Psi' \Delta d + \varepsilon^{-1}(\Psi'' |\nabla d|^2 - W'(\Psi)). \tag{3.6}$$

Since d is a distance function, $|\nabla d| = 1$, and the last two terms cancel each other. This leaves

$$\partial_t d + u_\varepsilon \cdot \nabla d \approx \Delta d. \tag{3.7}$$

Due to the nature of the distance function, evaluated on M_t, $\partial_t d$ is the outward velocity of M_t, $u_\varepsilon \cdot \nabla d$ is the inward normal component of u_ε and Δd is the mean curvature of M_t. As $\varepsilon \to 0$, this approximation may be expected to get better, and the relation (3.7) motivates that $\{\varphi_\varepsilon(\cdot, t) = 0\}$ should converge to M_t which moves by (1.2). This heuristic argument may be justified if we know in advance that there exists a smooth M_t moving by (1.2). Here, however, u is not smooth and we aim to obtain a time-global existence result which necessitates a framework inclusive of singularities. This is the reason to use the language of varifold in this paper as was done first by Ilmanen [28]. The basic approach is to prove that φ_ε satisfying (3.5) has the property that

$$\mu^\varepsilon := \left(\frac{\varepsilon|\nabla\varphi_\varepsilon|^2}{2} + \frac{W(\varphi_\varepsilon)}{\varepsilon}\right) dx \approx \sigma N(x,t)\mathcal{H}^{n-1}\lfloor_{M_t} \tag{3.8}$$

when ε is small and where $N(x,t)$ is some integer. At the same time we prove that the limiting measure of μ^ε satisfies (2.10). The first key estimate to be established is the analogue of (2.8) for φ_ε which will be discussed in the next section.

4 Density ratio upper bound and energy monotonicity formula

In this section, we prove the upper density ratio bound for diffused interface energy and energy monotonicity formula which are crucial in the limiting process. Estimates in this section are similar to [33, Sect. 3] with some modifications.

4.1 The upper density ratio bound

We state the main theorem concerning the uniform density ratio upper bound independent of ε of the Allen-Cahn equation with extra transport term. The proof takes the entire Sect. 4. We establish the monotonicity formula which is a perturbed version of Ilmanen's monotonicity formula for the Allen-Cahn equation (and Huisken's monotonicity formula for the MCF [26]) along the way.

Theorem 4.1 *Suppose $n \geq 2$, $\Omega = \mathbb{T}^n$ or $\mathbb{R}^n$, p, q satisfy (2.14),*

$$0 < \beta < \frac{1}{2}, \tag{4.1}$$

$0 < \varepsilon < 1$ and φ satisfies

$$\partial_t\varphi + u\cdot\nabla\varphi = \Delta\varphi - \frac{W'(\varphi)}{\varepsilon^2} \quad \text{on } \Omega\times[0,T], \tag{4.2}$$

$$\varphi(x,0) = \varphi_0(x) \quad \text{on } \Omega. \tag{4.3}$$

Assume $u \in C_c^\infty(\Omega\times[0,T])$, $\nabla^j\varphi$, $\partial_t\nabla^k\varphi \in C(\Omega\times[0,T])$ for $k\in\{0,1\}$ and $j\in\{0,1,2,3\}$. Let μ_t^ε be a Radon measure on Ω defined by

$$\int_\Omega \phi(x)\,d\mu_t^\varepsilon(x) := \int_\Omega \phi(x)\left(\frac{\varepsilon|\nabla\varphi(x,t)|^2}{2} + \frac{W(\varphi(x,t))}{\varepsilon}\right) dx \tag{4.4}$$

for $\phi \in C_c(\Omega)$ and define

$$D(t) := \max\left\{1, \mu_t^\varepsilon(\Omega), \sup_{B_r(x)\subset\Omega} \frac{\mu_t^\varepsilon(B_r(x))}{\omega_{n-1}r^{n-1}}\right\}, \quad t\in[0,T]. \tag{4.5}$$

Assume

$$\sup_{\Omega\times[0,T]} |\varphi| \le 1, \tag{4.6}$$

$$\sup_{\Omega} \varepsilon^i |\nabla^i \varphi_0| \le c_1 \quad \textit{for } i \in \{1,2,3\}, \tag{4.7}$$

$$\lim_{R\to\infty} R^k \|\varphi + 1\|_{C^2((\mathbb{R}^n\setminus B_R)\times[0,T])} = 0 \textit{ for any } k \in \mathbb{N} \textit{ in case } \Omega = \mathbb{R}^n, \tag{4.8}$$

$$\sup_{\Omega}\left(\frac{\varepsilon|\nabla\varphi_0|^2}{2} - \frac{W(\varphi_0)}{\varepsilon}\right) \le \varepsilon^{-\beta}, \tag{4.9}$$

$$\sup_{\Omega\times[0,T]} |u| \le \varepsilon^{-\beta}, \quad \sup_{\Omega\times[0,T]} |\nabla u| \le \varepsilon^{-(\beta+1)}, \tag{4.10}$$

$$\|u\|_{L^q([0,T];(W^{1,p}(\Omega))^n)} \le c_2 \tag{4.11}$$

and

$$D(0) \le D_0. \tag{4.12}$$

Then there exist $D_1 = D_1(c_2, n, p, q, D_0, T) > 0$ *and* $\epsilon_1 = \epsilon_1(c_2, n, p, q, D_0, T, c_1, \beta, W) > 0$ *such that*

$$\sup_{t\in[0,T]} D(t) \le D_1 \tag{4.13}$$

as long as $\varepsilon < \epsilon_1$.

Remark 4.1 If $u = 0$, $\mu_t^\varepsilon(\Omega)$ is monotone decreasing, thus it is straightforward to conclude that $\mu_t^\varepsilon(\Omega)$ is bounded uniformly independent of ε if $\mu_0^\varepsilon(\Omega)$ is. The uniform density ratio bound may be also obtained from Ilmanen's monotonicity formula. When $u \ne 0$, however, it is non-trivial even to conclude that the total energy $\mu_t^\varepsilon(\Omega)$ up to time T has a uniform bound independent of ε. We will see that we need the density ratio bound to estimate $\mu_t^\varepsilon(\Omega)$.

4.2 Monotonicity formula

In this subsection as a first step we obtain a modified monotonicity formula analogous to that of Ilmanen [28]. It is still not a very useful formula due to the possible negative contribution coming from ξ_ε defined below. We will show that the negative contribution is small when ε is small.

To localize the computations, fix a radially symmetric cut-off function

$$\eta(x) \in C_c^\infty\left(B_{\frac{1}{2}}\right) \quad \text{with} \quad \eta = 1 \text{ on } B_{\frac{1}{4}},\ 0 \le \eta \le 1. \tag{4.14}$$

Define

$$\tilde\rho_{(y,s)}(x,t) := \rho_{(y,s)}(x,t)\eta(x-y) = \frac{1}{(4\pi(s-t))^{\frac{n-1}{2}}} e^{-\frac{|x-y|^2}{4(s-t)}} \eta(x-y) \tag{4.15}$$

for $t < s$ and $x, y \in \Omega$ and define

$$e_\varepsilon := \frac{\varepsilon|\nabla\varphi|^2}{2} + \frac{W(\varphi)}{\varepsilon}, \quad \xi_\varepsilon := \frac{\varepsilon|\nabla\varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon}. \tag{4.16}$$

Proposition 4.1 *Suppose that φ satisfies (4.2). With the notation of (4.4), (4.15), (4.16) and writing $\tilde\rho = \tilde\rho_{(y,s)}(x,t)$, we have c_3 depending only on n such that*

$$\frac{d}{dt}\int_\Omega \tilde\rho\, d\mu_t^\varepsilon(x) \le \frac{1}{2}\int_\Omega \tilde\rho|u|^2\, d\mu_t^\varepsilon(x) + \frac{1}{2(s-t)}\int_\Omega \xi_\varepsilon\tilde\rho\, dx + c_3 e^{-\frac{1}{128(s-t)}}\mu_t^\varepsilon(B_{\frac{1}{2}}(y)) \tag{4.17}$$

for $y \in \Omega$, $0 < t < s < \infty$ and $t < T$.

Proof We define L as follows and by (4.2),

$$L := \partial_t\varphi + u\cdot\nabla\varphi = \Delta\varphi - \frac{W'(\varphi)}{\varepsilon^2}.$$

By integration by parts we have

$$\begin{aligned}
\frac{d}{dt}\int_\Omega e_\varepsilon\tilde\rho\, dx &= \int_\Omega \{e_\varepsilon\partial_t\tilde\rho - \varepsilon(L - u\cdot\nabla\varphi)(\nabla\tilde\rho\cdot\nabla\varphi + \tilde\rho L)\}\, dx \\
&= \int_\Omega \left\{ e_\varepsilon\partial_t\tilde\rho - \varepsilon\tilde\rho\left(L + \frac{\nabla\tilde\rho\cdot\nabla\varphi}{\tilde\rho}\right)^2 + \varepsilon\left(L\nabla\tilde\rho\cdot\nabla\varphi + \frac{(\nabla\tilde\rho\cdot\nabla\varphi)^2}{\tilde\rho}\right)\right. \\
&\quad \left. + \varepsilon\tilde\rho u\cdot\nabla\varphi\left(L + \frac{\nabla\tilde\rho\cdot\nabla\varphi}{\tilde\rho}\right)\right\} dx \\
&\le \int_\Omega \left\{ e_\varepsilon\partial_t\tilde\rho + \varepsilon\left(L\nabla\tilde\rho\cdot\nabla\varphi + \frac{(\nabla\tilde\rho\cdot\nabla\varphi)^2}{\tilde\rho}\right) + \frac{1}{4}\varepsilon\tilde\rho(u\cdot\nabla\varphi)^2\right\} dx.
\end{aligned} \tag{4.18}$$

Moreover by integration by parts we obtain

$$\int_\Omega \varepsilon L\nabla\tilde\rho\cdot\nabla\varphi\, dx = \int_\Omega -\varepsilon(\nabla\varphi\otimes\nabla\varphi)\cdot\nabla^2\tilde\rho + e_\varepsilon\Delta\tilde\rho\, dx. \tag{4.19}$$

Substitution of (4.19) into (4.18) gives

$$\begin{aligned}
\frac{d}{dt}\int_\Omega e_\varepsilon\tilde\rho\, dx \le \int_\Omega &(-\xi_\varepsilon)(\partial_t\tilde\rho + \Delta\tilde\rho) + \varepsilon|\nabla\varphi|^2\left(\partial_t\tilde\rho + \Delta\tilde\rho - \frac{\nabla\varphi\otimes\nabla\varphi}{|\nabla\varphi|^2}\cdot\nabla^2\tilde\rho\right. \\
&\left. + \frac{(\nabla\tilde\rho\cdot\nabla\varphi)^2}{\tilde\rho|\nabla\varphi|^2}\right) + \frac{1}{4}\varepsilon\tilde\rho(u\cdot\nabla\varphi)^2\, dx.
\end{aligned} \tag{4.20}$$

We remark that ρ (without multiplication by η) satisfies the following:

$$\partial_t \rho + \Delta\rho = -\frac{\rho}{2(s-t)}, \quad \partial_t \rho + \Delta\rho - \frac{\nabla\varphi \otimes \nabla\varphi}{|\nabla\varphi|^2} \cdot \nabla^2\rho + \frac{(\nabla\rho \cdot \nabla\varphi)^2}{\rho|\nabla\varphi|^2} = 0. \tag{4.21}$$

When one computes (4.21) with $\tilde{\rho}$ instead of ρ, we have additional terms coming from differentiation of η. The integration of these terms can be bounded by $c\mu_t^\varepsilon(B_{1/2}(y))e^{-\frac{1}{128(s-t)}}$ for $c = c(n)$ since $|\nabla^j\rho| \le c(j,n)e^{-\frac{1}{128(s-t)}}$ for any $x, y \in \Omega$ with $|x-y| > \frac{1}{4}$ and $j = 0, 1$. Thus, with an appropriate choice of c_3 depending only on n, we obtain (4.17). □

4.3 Some estimates on $\Omega \times [0, T]$

Lemma 4.1 *Suppose that φ satisfies (4.2), (4.3), (4.6), (4.7) and (4.10). Then there exists $c_4 > 0$ depending only on n, c_1, W such that*

$$\sup_{\Omega\times[0,T]} \varepsilon|\nabla\varphi| + \sup_{x,y\in\Omega,\ t\in[0,T]} \varepsilon^{\frac{3}{2}} \frac{|\nabla\varphi(x,t) - \nabla\varphi(y,t)|}{|x-y|^{\frac{1}{2}}} \le c_4. \tag{4.22}$$

Proof Take any domain $B_{3\varepsilon}(x_0) \times [t_0, t_0 + 2\varepsilon^2] \subset \Omega \times [0, T]$. Define $\tilde{\varphi}(x,t) := \varphi(\varepsilon x + x_0, \varepsilon^2 t + t_0)$ and $\tilde{u}(x,t) := u(\varepsilon x + x_0, \varepsilon^2 t + t_0)$ for $(x,t) \in B_3 \times [0, 2]$. By (4.2) we have

$$\partial_t\tilde{\varphi} + \varepsilon\tilde{u} \cdot \nabla\tilde{\varphi} = \Delta\tilde{\varphi} - W'(\tilde{\varphi}). \tag{4.23}$$

Using the estimate of [32, p. 342, Theorem 9.1], if $\partial_t v - \Delta v = f$ on $B_2 \times [0, 2]$ then we have

$$\|\partial_t v, \nabla^2 v\|_{L^r(B_1\times[j,2])} \le c(n,r)(\|f, \nabla v, v\|_{L^r(B_2\times[0,2])}) + (1-j)\|v(\cdot,0)\|_{W^{2,r}(B_2)}) \tag{4.24}$$

for $j = 0$ (up to $t = 0$) or $j = 1$ (interior estimate) and for $r \in (1, \infty)$. Let $\phi \in C_c^1(B_3)$ be a cut-off function and multiply $\phi^2\tilde{\varphi}$ to (4.23), then by integration by parts, (4.6), (4.7) and (4.10), we have

$$\int_0^2 \int_{B_2} |\nabla\tilde{\varphi}|^2\, dxdt \le c(W). \tag{4.25}$$

Hence by (4.6), (4.7), (4.10), (4.24) ($r = 2$) and (4.25) we obtain

$$\int_0^2 \int_{B_1} |\tilde{\varphi}_t|^2 + |\nabla^2\tilde{\varphi}|^2\, dxdt \le c(n, c_1, W).$$

By applying (4.24) to the equation

$$\partial_t(\tilde{\varphi}_{x_i}) - \Delta\tilde{\varphi}_{x_i} = -\varepsilon\tilde{u}_{x_i}\cdot\nabla\tilde{\varphi} - \varepsilon\tilde{u}\cdot\nabla\tilde{\varphi}_{x_i} - W''(\tilde{\varphi})\tilde{\varphi}_{x_i},$$

and using (4.6), (4.7) and (4.10) again, we obtain

$$\int_0^2\int_{B_1}|\nabla\tilde{\varphi}_t|^2 + |\nabla^3\tilde{\varphi}|^2\,dxdt \le c(n, c_1, W).$$

Therefore we obtain the $W^{1,2}$ estimates of $\nabla\tilde{\varphi}$ on $B_1\times[0,2]$, and by the Sobolev inequality we have

$$\|\nabla\tilde{\varphi}\|_{L^{\frac{2(n+1)}{n-1}}(B_1\times[0,2])} \le c(n, c_1, W).$$

We can use this estimate to (4.23) and (4.24) with $r = \frac{2(n+1)}{n-1}$. We repeat this argument until r is large enough so that $W^{1,r}\subset C^{\frac{1}{2}}$ with appropriate modifications of the domain. Then we obtain the desired estimate

$$\|\nabla\tilde{\varphi}\|_{C^{\frac{1}{2}}(B_1\times[0,2])} \le c(n, c_1, W).$$

Since the domain was arbitrary, after returning to the original coordinate system, we obtain (4.22). □

Lemma 4.2 *There exists* $\epsilon_2 = \epsilon_2(n, W, \beta) > 0$ *such that, if* $\varepsilon < \epsilon_2$ *and under the assumptions of* (4.1)–(4.3), (4.6), (4.7), (4.9) *and* (4.10), *we have*

$$\frac{\varepsilon|\nabla\varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \le 10\varepsilon^{-\beta} \quad \text{on } \Omega\times[0,T]. \tag{4.26}$$

Proof Rescale the domain by $x\mapsto\frac{x}{\varepsilon}$ and $t\mapsto\frac{t}{\varepsilon^2}$. Under the change of variables, we continue to use the same notations for φ and u. Define

$$\xi := \frac{|\nabla\varphi|^2}{2} - W(\varphi) - G(\varphi), \tag{4.27}$$

where G will be chosen later. We compute $\partial_t\xi + \varepsilon u\cdot\nabla\xi - \Delta\xi$ and obtain

$$\begin{aligned}
&\partial_t\xi + \varepsilon u\cdot\nabla\xi - \Delta\xi \\
&\quad = \nabla\varphi\cdot\nabla\partial_t\varphi - (W' + G')\partial_t\varphi + \varepsilon(u\otimes\nabla\varphi)\cdot\nabla^2\varphi - \varepsilon(W' + G')u\cdot\nabla\varphi \\
&\qquad -|\nabla^2\varphi|^2 - \nabla\varphi\cdot\nabla(\Delta\varphi) + (W' + G')\Delta\varphi + (W'' + G'')|\nabla\varphi|^2.
\end{aligned} \tag{4.28}$$

Here, we denoted and will denote $W'(\varphi)$ as W', $G(\varphi)$ as G and so forth for simplicity. Differentiate (4.23) with respect to x_j, multiply φ_{x_j} and sum over j to obtain

$$\begin{aligned}&\nabla\varphi\cdot\nabla\partial_t\varphi+\varepsilon\nabla u\cdot(\nabla\varphi\otimes\nabla\varphi)+\varepsilon(u\otimes\nabla\varphi)\cdot\nabla^2\varphi\\&\quad=\nabla\varphi\cdot\nabla(\Delta\varphi)-W''|\nabla\varphi|^2.\end{aligned}\tag{4.29}$$

By (4.23), (4.28) and (4.29) we have

$$\begin{aligned}\partial_t\xi+\varepsilon u\cdot\nabla\xi-\Delta\xi&=W'(W'+G')-|\nabla^2\varphi|^2\\&\quad-\varepsilon\nabla u\cdot(\nabla\varphi\otimes\nabla\varphi)+G''|\nabla\varphi|^2.\end{aligned}\tag{4.30}$$

Differentiating (4.27) with respect to x_j and by using the Cauchy-Schwarz inequality we have

$$\begin{aligned}\sum_{j=1}^n\left(\sum_{i=1}^n\varphi_{x_i}\varphi_{x_ix_j}\right)^2&=\sum_{j=1}^n\left(\xi_{x_j}+(W'+G')\varphi_{x_j}\right)^2\\&=|\nabla\xi|^2+2(W'+G')\nabla\xi\cdot\nabla\varphi+(W'+G')^2|\nabla\varphi|^2\\&\le|\nabla\varphi|^2|\nabla^2\varphi|^2.\end{aligned}\tag{4.31}$$

On $\{|\nabla\varphi|>0\}$, divide (4.31) by $|\nabla\varphi|^2$ and substitute into (4.30) to obtain

$$\begin{aligned}&\partial_t\xi+\varepsilon u\cdot\nabla\xi-\Delta\xi\\&\quad\le W'(W'+G')-\frac{1}{|\nabla\varphi|^2}(|\nabla\xi|^2+2(W'+G')\nabla\xi\cdot\nabla\varphi+(W'+G')^2|\nabla\varphi|^2)\\&\qquad-\varepsilon\nabla u\cdot(\nabla\varphi\otimes\nabla\varphi)+G''|\nabla\varphi|^2\\&\quad\le-(G')^2-W'G'-\frac{2(W'+G')}{|\nabla\varphi|^2}\nabla\xi\cdot\nabla\varphi-\varepsilon\nabla u\cdot(\nabla\varphi\otimes\nabla\varphi)\\&\qquad+G''|\nabla\varphi|^2.\end{aligned}\tag{4.32}$$

By $|\nabla\varphi|^2=2(\xi+W+G)$ and (4.32) we have on $\{|\nabla\varphi|>0\}$

$$\begin{aligned}\partial_t\xi+\varepsilon u\cdot\nabla\xi-\Delta\xi&\le-(G')^2-W'G'+2G''(\xi+W+G)\\&\quad-\frac{2(W'+G')}{|\nabla\varphi|^2}\nabla\xi\cdot\nabla\varphi-\varepsilon\nabla u\cdot(\nabla\varphi\otimes\nabla\varphi).\end{aligned}\tag{4.33}$$

Let $\phi(x,t)=\phi(x)\in C^\infty(B_{3\varepsilon^{-1}})$ be such that

$$\phi=\begin{cases}M:=\sup_{\mathbb{R}^n\times[0,\varepsilon^{-2}T]}\left(\frac{|\nabla\varphi|^2}{2}-W(\varphi)\right)&\text{on }B_{3\varepsilon^{-1}}\setminus B_{2\varepsilon^{-1}},\\0\quad\text{on }B_{\varepsilon^{-1}},\end{cases}$$

and

$$0\le\phi\le M,\ |\nabla\phi|\le2\varepsilon M,\ |\Delta\phi|\le2n\varepsilon^2M.$$

Note that M may be bounded depending only on n, c_1, W by Lemma 4.1. Note also that we may assume $M > 0$ since $M \le 0$ implies our conclusion (4.26) immediately. Let

$$\tilde{\xi} := \xi - \phi \quad \text{and} \quad G(\varphi) := \varepsilon^{\frac{1}{2}}\left(1 - \frac{1}{8}(\varphi - \gamma)^2\right),$$

where γ is as in (3.2). To derive a contradiction, suppose that

$$\sup_{B_{\varepsilon^{-1}}\times[0,\varepsilon^{-2}T]} \xi \ge \varepsilon^{\frac{1}{2}}.$$

Since $\tilde{\xi} \le 0$ on $(B_{3\varepsilon^{-1}}\setminus B_{2\varepsilon^{-1}}) \times [0, \varepsilon^{-2}T]$, $\tilde{\xi} \le \varepsilon^{1-\beta}$ on $B_{3\varepsilon^{-1}} \times \{0\}$ by (4.9) and $\sup_{B_{\varepsilon^{-1}}\times[0,\varepsilon^{-2}T]} \tilde{\xi} \ge \varepsilon^{\frac{1}{2}}$, there exists some interior maximum point (x_0, t_0) of $\tilde{\xi}$ where

$$\partial_t\tilde{\xi} \ge 0,\ \nabla\tilde{\xi} = 0,\ \Delta\tilde{\xi} \le 0 \quad \text{and} \quad \tilde{\xi} \ge \varepsilon^{\frac{1}{2}}$$

hold. By the definition of ϕ we have at the point (x_0, t_0)

$$\partial_t\xi \ge 0,\ |\nabla\xi| \le 2\varepsilon M,\ \Delta\xi \le 2n\varepsilon^2 M \quad \text{and} \quad |\nabla\varphi|^2 \ge 2\varepsilon^{\frac{1}{2}}. \tag{4.34}$$

Substitute (4.34) into (4.33). Using $\varepsilon\nabla u \cdot (\nabla\varphi \otimes \nabla\varphi) \le 2\varepsilon|\nabla u|(\xi + W + G)$ and (4.10), we have

$$\begin{aligned} 0 \le\ & 2n\varepsilon^2 M - (G')^2 - W'G' + 2G''(\xi + W + G) + \frac{4(|W'| + |G'|)\varepsilon M}{\left(2\varepsilon^{\frac{1}{2}}\right)^{\frac{1}{2}}} \\ & + 2\varepsilon^{1-\beta}(\xi + W + G) + 2\varepsilon^{2-\beta}M. \end{aligned} \tag{4.35}$$

Since $\beta < \frac{1}{2}$ and $G'' = -\varepsilon^{\frac{1}{2}}/4$, for sufficient small ε depending only on β and W,

$$2G''(\xi + W + G) + 2\varepsilon^{1-\beta}(\xi + W + G) \le G''(W + G). \tag{4.36}$$

If $|\varphi(x_0, t_0)| \le \alpha$, then

$$G''(\varphi(x_0, t_0))W(\varphi(x_0, t_0)) \le -\frac{\varepsilon^{\frac{1}{2}}}{4} \min_{|z|\le\alpha} W(z),$$

which is a 'big' negative number compared to the rest, and one can check that this and (4.36) (as well as $W'G' \ge 0$ and $G > 0$) lead to a contradiction in (4.35). If $|\varphi(x_0, t_0)| \ge \alpha$, then we would have 'big' negative contributions coming from (all evaluated at (x_0, t_0))

$$(G')^2 \ge \frac{\varepsilon(\alpha - |\gamma|)^2}{64} \quad \text{and} \quad -W'G' \le -\frac{\varepsilon^{\frac{1}{2}}(\alpha - |\gamma|)}{4}|W'|,$$

which again lead to a contradiction in (4.35) for sufficiently small ε. This shows that

$$\sup_{B_{\varepsilon^{-1}}\times[0,\varepsilon^{-2}T]}\left(\frac{|\nabla\varphi|^2}{2}-W(\varphi)\right)\le 2\varepsilon^{\frac{1}{2}},$$

where $G\le\varepsilon^{\frac{1}{2}}$ is used. Now repeat the same argument, this time with M replaced by $2\varepsilon^{\frac{1}{2}}$ and G replaced by $8\varepsilon^{1-\beta}(1-\frac{1}{8}(\varphi-\gamma)^2)$. If we assume

$$\sup_{B_{\varepsilon^{-1}}\times[0,\varepsilon^{-2}T]}\xi\ge 2\varepsilon^{1-\beta},$$

$\tilde{\xi}=\xi-\phi$ would attain some interior maximum in $B_{3\varepsilon^{-1}}\times[0,\varepsilon^{-2}T]$ by (4.9) and by the subtraction of ϕ. This time we would have $\partial_t\xi\ge 0,\ |\nabla\xi|\le 4\varepsilon^{\frac{3}{2}},\ \Delta\xi\le 4n\varepsilon^{\frac{5}{2}}$ and $|\nabla\varphi|^2\ge 4\varepsilon^{1-\beta}$. With this (4.35) is

$$0\le 4n\varepsilon^{\frac{5}{2}}-(G')^2-W'G'+2G''(\xi+W+G)+\frac{8(|W'|+|G'|)\varepsilon^{\frac{3}{2}}}{(4\varepsilon^{1-\beta})^{\frac{1}{2}}}$$
$$+2\varepsilon^{1-\beta}(\xi+W+G)+4\varepsilon^{\frac{5}{2}-\beta}.$$

Exactly the same type of argument as before shows that we have a contradiction, and since $G\le 8\varepsilon^{1-\beta}$ and $\xi-G\le 2\varepsilon^{1-\beta}$, we have (4.26). □

Lemma 4.3 *Let μ_s^ε, $D(t)$ and $\tilde{\rho}_{(y,s)}$ be defined as in* (4.4), (4.5) *and* (4.15). *Let s, R, r be positive with $0\le s-(\frac{R}{r})^2\le T$ and $R\in(0,\frac{1}{2})$. Set $\tilde{s}=s-(\frac{R}{r})^2$. Then there exists $c_5=c_5(n)\ge 1$ such that, for any $y\in\Omega$, we have*

$$\int_\Omega\tilde{\rho}_{(y,s)}(x,\tilde{s})\,d\mu_{\tilde{s}}^\varepsilon(x)\le\left(\frac{r}{\sqrt{4\pi}R}\right)^{n-1}\left\{\mu_{\tilde{s}}^\varepsilon(B_R(y))+\mu_{\tilde{s}}^\varepsilon(B_{\frac{1}{2}}(y))\exp\left(-\frac{r^2}{16R^2}\right)\right\}$$
$$+c_5D(\tilde{s})\exp\left(-\frac{r^2}{8}\right).$$

Proof First, on $B_R(y)$ we compute

$$\int_{B_R(y)}\tilde{\rho}_{(y,s)}(x,\tilde{s})\,d\mu_{\tilde{s}}^\varepsilon\le\left(\frac{r}{\sqrt{4\pi}R}\right)^{n-1}\int_{B_R(y)}e^{-\frac{r^2|x-y|^2}{4R^2}}\,d\mu_{\tilde{s}}^\varepsilon$$
$$\le\left(\frac{r}{\sqrt{4\pi}R}\right)^{n-1}\mu_{\tilde{s}}^\varepsilon(B_R(y)).$$

On $\Omega\backslash B_R(y)$ we have

$$
\begin{aligned}
&\left(\frac{\sqrt{4\pi}R}{r}\right)^{n-1}\int_{\Omega\backslash B_R(y)}\tilde{\rho}_{(y,s)}(x,\tilde{s})\,d\mu^{\varepsilon}_{\tilde{s}} \le \int_{B_{\frac{1}{2}}(y)\backslash B_R(y)} e^{-\frac{r^2|x-y|^2}{4R^2}}\,d\mu^{\varepsilon}_{\tilde{s}}\\
&\le \int_0^1 \mu^{\varepsilon}_{\tilde{s}}\left(\left(B_{\frac{1}{2}}(y)\backslash B_R(y)\right)\cap\left\{x \mid e^{-\frac{r^2|x-y|^2}{4R^2}}\ge\lambda\right\}\right)d\lambda\\
&\le \int_0^{\exp(-\frac{r^2}{16R^2})}\mu^{\varepsilon}_{\tilde{s}}\left(B_{\frac{1}{2}}(y)\backslash B_R(y)\right)d\lambda+\int_{\exp(-\frac{r^2}{16R^2})}^{\exp(-\frac{r^2}{4})}\mu^{\varepsilon}_{\tilde{s}}(B_{\frac{2R}{r}\sqrt{\log\lambda^{-1}}}(y))\,d\lambda\\
&\le \mu^{\varepsilon}_{\tilde{s}}\left(B_{\frac{1}{2}}(y)\right)e^{-\frac{r^2}{16R^2}}+D(\tilde{s})\omega_{n-1}\left(\frac{2R}{r}\right)^{n-1}\int_{\frac{r^2}{4}}^{\frac{r^2}{16R^2}} l^{\frac{n-1}{2}}e^{-l}\,dl\\
&\le \mu^{\varepsilon}_{\tilde{s}}\left(B_{\frac{1}{2}}(y)\right)e^{-\frac{r^2}{16R^2}}+c(n)D(\tilde{s})\left(\frac{2R}{r}\right)^{n-1}e^{-\frac{r^2}{8}}.
\end{aligned}
\tag{4.37}
$$

Here we used the fact that there exists $c = c(n) > 0$ such that $l^{\frac{n-1}{2}}e^{-l} \le ce^{-\frac{l}{2}}$ for any $l > 0$. □

4.4 Proof of Theorem 4.1

In this subsection, we always work under the assumptions of Theorem 4.1. In particular, results from the two preceding subsections are available. Furthermore, from now on until Proposition 4.2, we assume

$$D(t) \le D_1 \tag{4.38}$$

holds for $t \in [0, T_1]$ and $T_1 \le T$. Here, $D_1 \ge 2D_0$ is a constant depending only on c_2, n, p, q, T, D_0, and not on ε, and which will be determined after Proposition 4.2. We need to be careful about the dependence of constants so that we do not end up a circular argument. Any constant depending on D_1 will be again a constant depending on c_2, n, p, q, T, D_0. Note that such $T_1 > 0$ exists because $D_1 > D_0$ and by the continuity of $D(t)$ in time. Such continuity follows from that of φ in the case of $\Omega = \mathbb{T}^n$, and additionally from (4.8) in the case of $\Omega = \mathbb{R}^n$. T_1 may depend on ε in general, but in the end, we prove that $T_1 = T$ as long as ε is sufficiently small. First, under this assumption we have the following a-priori estimate:

Lemma 4.4 *There exists c_6 depending only on n, p, q such that for any $0 \le t_0 < t_1$ we have*

$$
\begin{aligned}
&\sup_{t\in[t_0,t_1]}\mu^{\varepsilon}_t(\Omega)+\frac{1}{2}\int_{t_0}^{t_1}\int_{\Omega}\varepsilon\left(\Delta\varphi-\frac{W'(\varphi)}{\varepsilon^2}\right)^2 dxdt\\
&\quad\le \mu^{\varepsilon}_{t_0}(\Omega)+c_6(t_1-t_0)^{1-\frac{2}{q}}\|u\|^2_{L^q([t_0,t_1];(W^{1,p}(\Omega))^n)}\sup_{t\in[t_0,t_1]}D(t).
\end{aligned}
\tag{4.39}
$$

In particular, there exists E_0 *depending only on* c_2, n, p, q, T, D_0 *such that*

$$\sup_{t\in[0,T_1]} \mu_t^\varepsilon(\Omega) + \frac{1}{2}\int_0^{T_1}\int_\Omega \varepsilon\left(\Delta\varphi - \frac{W'(\varphi)}{\varepsilon^2}\right)^2 dxdt \le E_0. \tag{4.40}$$

Proof By (4.2) we can compute

$$\frac{d}{dt}\mu_t^\varepsilon(\Omega) \le -\frac{1}{2}\int_\Omega \varepsilon\left(\Delta\varphi - \frac{W'(\varphi)}{\varepsilon^2}\right)^2 dx + \varepsilon\int_\Omega (u\cdot\nabla\varphi)^2\,dx. \tag{4.41}$$

To estimate the last term of (4.41), we consider two cases $p < 2$ and $p \ge 2$ separately. In addition we consider $\Omega = \mathbb{T}^n, \mathbb{R}^n$ separately, and let us consider $\mathbb{T}^n$ first. Let $\{\psi_\alpha\}_\alpha$ be a partition of unity on Ω such that $\psi_\alpha \in C_c^\infty(\Omega)$, diam (spt ψ_α) $\le 1/2$ and $\|\psi_\alpha\|_{C^2} \le c(n)$. Consider $p < 2$ case first. Just as in (2.12), by setting $s := \frac{p(n-1)}{n-p} \ge 2$, we have

$$\begin{aligned}
\varepsilon\int_\Omega (u\cdot\nabla\varphi)^2\,dx &\le \left(\int_\Omega |u|^s\varepsilon|\nabla\varphi|^2\,dx\right)^{\frac{2}{s}} (2\mu_t^\varepsilon(\Omega))^{1-\frac{2}{s}}\\
&\le \left(\sum_\alpha c(n,p)\int_\Omega |\psi_\alpha u|^s\varepsilon|\nabla\varphi|^2\,dx\right)^{\frac{2}{s}} (2D(t))^{1-\frac{2}{s}}\\
&\le \left(\sum_\alpha c(n,p)D(t)\left(\int_{\mathrm{spt}\,\psi_\alpha} |u|^p + |\nabla u|^p\,dx\right)^{\frac{s}{p}}\right)^{\frac{2}{s}} (2D(t))^{1-\frac{2}{s}}\\
&\le c(n,p)D(t)\|u(\cdot,t)\|^2_{W^{1,p}(\Omega)}
\end{aligned} \tag{4.42}$$

where each constant is different. We used the local finiteness of $\{\psi_\alpha\}_\alpha$ and $\sum_\alpha A_\alpha^{\frac{s}{p}} \le (\sum_\alpha A_\alpha)^{\frac{s}{p}}$ since $\frac{s}{p} \ge 1$. For $p \ge 2$, we have

$$\begin{aligned}
\varepsilon\int_\Omega (u\cdot\nabla\varphi)^2\,dx &\le \left(\int_\Omega |u|^p\varepsilon|\nabla\varphi|^2\,dx\right)^{\frac{2}{p}} (2\mu_t^\varepsilon(\Omega))^{1-\frac{2}{p}}\\
&\le \left(\sum_\alpha c(n,p)\int_\Omega |\psi_\alpha u|^p\,\varepsilon|\nabla\varphi|^2\,dx\right)^{\frac{2}{p}} (2D(t))^{1-\frac{2}{p}}\\
&\le \left(\sum_\alpha c(n,p)D(t)\int_{\mathrm{spt}\,\psi_\alpha} |u|^p + |u|^{p-1}|\nabla u|\,dx\right)^{\frac{2}{p}} (2D(t))^{1-\frac{2}{p}}\\
&\le c(n,p)D(t)\|u(\cdot,t)\|^2_{W^{1,p}(\Omega)}.
\end{aligned} \tag{4.43}$$

Here we used (2.11) with $p = 1$ there and $\phi = |\psi_\alpha u|^p$. Integration of (4.39) over $[t_0, t_1]$ using (4.42) or (4.43) gives (4.39). We define E_0 to be $D_0 + c_6 T^{1-\frac{2}{q}} c_2^2 D_1$. In case of $\Omega = \mathbb{R}^n$, we do not need to take the partition of unity and the proof proceeds similarly. $\square$

In the following we define β' by

$$\beta' := \frac{1+\beta}{2}.$$

In fact, any number $\beta' \in (\beta, 1)$ can be used. To fix the idea, we specify such β', and suppose that β' depends on β for simplicity.

Lemma 4.5 *There exist $c_7 > 1$, $1 > c_8 > 0$ and $\epsilon_3 > 0$ with $0 < \epsilon_3 \le \epsilon_2$ depending only on $n, c_1, c_2, p, q, T, W, \beta$ and D_0 with the following property. Assume $\varepsilon \in (0, \epsilon_3)$ and $|\varphi(y,s)| \le \alpha < 1$ with $s \in (0, T_1]$. Here α is from (3.3). Then for any $t \in [0, T_1]$ with* $\max\{0, s - 2\varepsilon^{2\beta'}\} \le t \le s$ *we have*

$$c_8 \le \frac{1}{R^{n-1}} \mu_t^\varepsilon (B_R(y)), \tag{4.44}$$

where $R = c_7(s + \varepsilon^2 - t)^{\frac{1}{2}}$.

Proof We will choose $\epsilon_3 < \epsilon_2$ and assume for the moment that $\varepsilon < \epsilon_2$. Set $\tilde\rho = \tilde\rho_{(y,s+\varepsilon^2)}(x,t)$ in this proof. Assume $|\varphi(y,s)| \le \alpha < 1$. We have

$$\int_\Omega \tilde\rho \, d\mu_s^\varepsilon(x) = \int_{\varepsilon^{-1}\Omega} \frac{e^{-\frac{|\tilde x|^2}{4}}}{(\sqrt{4\pi})^{n-1}} \eta(\varepsilon \tilde x) \left(\frac{|\nabla \tilde\varphi|^2}{2} + W(\tilde\varphi) \right) d\tilde x,$$

where $\tilde\varphi(\tilde x, s) = \varphi(\varepsilon \tilde x + y, s)$. By $|\tilde\varphi(0,s)| \le \alpha < 1$ and Lemma 4.1 there exists $0 < c_9 = c_9(n, c_1, W) < 1$ such that

$$5c_9 \le \int_\Omega \tilde\rho \, d\mu_s^\varepsilon(x). \tag{4.45}$$

From (4.10), (4.17), (4.26), (4.40) and $\varepsilon < \epsilon_2$ we have for $\lambda \in [t, s)$

$$\frac{d}{d\lambda} \int_\Omega \tilde\rho \, d\mu_\lambda^\varepsilon \le \varepsilon^{-2\beta} \int_\Omega \tilde\rho \, d\mu_\lambda^\varepsilon + \frac{10\sqrt{\pi}\varepsilon^{-\beta}}{\sqrt{s-\lambda}} + c_3 e^{\frac{-1}{128(s+\varepsilon^2-\lambda)}} D_1. \tag{4.46}$$

Here $\int_\Omega \tilde\rho \, dx \le \sqrt{4\pi(s-t)}$ is used. Multiply (4.46) by $e^{\varepsilon^{-2\beta}(s-\lambda)}$ and integrate over $[t, s]$. By $t \ge \max\{0, s - 2\varepsilon^{2\beta'}\}$ we have

$$\begin{aligned} e^{\varepsilon^{-2\beta}(s-\lambda)} \int_\Omega \tilde\rho d\mu_\lambda^\varepsilon(x) \Big|_{\lambda=t}^{s} &\le \varepsilon^{\beta'-\beta} e^{2\varepsilon^{2(\beta'-\beta)}} 20\sqrt{2\pi} \\ &\quad + 2c_3 D_1 e^{2\varepsilon^{2(\beta'-\beta)}} e^{\frac{-1}{128(\varepsilon^2+2\varepsilon^{2\beta'})}} \varepsilon^{2\beta'}. \end{aligned} \tag{4.47}$$

By (4.45) and (4.47) for sufficiently small ε depending only on D_1, β, n and c_3 we have

$$2c_9 \leq \int_\Omega \tilde{\rho}\, d\mu_t^\varepsilon(x). \tag{4.48}$$

Next we use Lemma 4.37 with $r := \sqrt{8\log(2c_5 D_1 c_9^{-1})}$, where we may assume that $c_5, D_1 > 1$ and $c_9 < 1$. We chose this r so that

$$c_5 D_1 e^{-\frac{r^2}{8}} = \frac{c_9}{2}. \tag{4.49}$$

In Lemma 4.37, we replace s and $s - (\frac{R}{r})^2$ by $s + \varepsilon^2$ and t respectively. Remark that $R := r(s+\varepsilon^2 - t)^{\frac{1}{2}} \leq r(\varepsilon^2 + 2\varepsilon^{2\beta'})^{\frac{1}{2}}$ since $s - t \leq 2\varepsilon^{2\beta'}$. Hence we have $R < \frac{1}{2}$ by restricting ε depending only on c_5, D_1 and c_9. From (4.38), (4.40) and Lemma 4.37 we have

$$\int_\Omega \tilde{\rho}\, d\mu_t^\varepsilon(x) \leq \left(r/(\sqrt{4\pi}R)\right)^{n-1} \{\mu_t^\varepsilon(B_R(y)) + E_0 e^{-r^2/(16R^2)}\} + c_5 D_1 e^{-r^2/8}. \tag{4.50}$$

Note that $r/R \geq \varepsilon^{-\beta'}/\sqrt{3}$. By (4.50), (4.48) and (4.49) for sufficiently small ε we obtain

$$c_9 \leq \left(r/(\sqrt{4\pi}R)\right)^{n-1} \mu_t^\varepsilon(B_R(y)).$$

Set $c_7 := r = \sqrt{8\log(2c_5 D_1 c_9^{-1})}$ and $c_8 = r^{1-n}(\sqrt{4\pi})^{n-1} c_9$ and we have the desired estimate (4.44). Note that the restriction on ε depends on c_3, c_5, D_1, c_9. Examining the dependence, we may conclude the proof. □

Lemma 4.6 *There exists* $0 < \epsilon_4 \leq \epsilon_3$ *and* c_{10} *depending only on* n, c_1, c_2, p, q, T, W, β *and* D_0 *with the following property. For any* $r \in (\varepsilon^{\beta'}, \frac{1}{2})$ *and* $t \in [2\varepsilon^{2\beta'}, T] \cap [0, T_1]$, *we have*

$$\int_{B_r(y)} \left(\frac{\varepsilon|\nabla\varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon}\right)_+ (x,t)\, dx \leq c_{10}\varepsilon^{\beta'-\beta} r^{n-1} \tag{4.51}$$

provided $\varepsilon \leq \epsilon_4$.

Proof We only need to prove the claim when $T_1 \geq 2\varepsilon^{2\beta'}$ since the claim is vacuously true otherwise. Let $y \in \Omega$, $r \in (\varepsilon^{\beta'}, \frac{1}{2})$ and $t_* \in [2\varepsilon^{2\beta'}, T] \cap [0, T_1]$ be arbitrary and fixed. We define

$$\tilde{A} := \left\{x \in B_{2r}(y) \ : \ \text{for some } \tilde{t} \text{ with } t_* - \varepsilon^{2\beta'} \leq \tilde{t} \leq t_*,\ |\varphi(x,\tilde{t})| \leq \alpha\right\},$$

$$A := \left\{ x \in B_{2r+2c_7\varepsilon^{\beta'}}(y) \, : \, \mathrm{dist}(\tilde{A}, x) < 2c_7\varepsilon^{\beta'} \right\}.$$

By Vitali's covering theorem applied to $\mathcal{F} = \{\bar{B}_{2c_7\varepsilon^{\beta'}}(x) \, : \, x \in \tilde{A}\}$ (note $A \subset \cup_{B\in\mathcal{F}} B$), there exists a set of pairwise disjoint balls $\{B_{2c_7\varepsilon^{\beta'}}(x_i)\}_{i=1}^N$ such that

$$x_i \in \tilde{A} \text{ for each } i = 1, \dots, N \quad \text{and} \quad A \subset \cup_{i=1}^N \bar{B}_{10c_7\varepsilon^{\beta'}}(x_i). \tag{4.52}$$

By the definition of $\tilde{A}$, for each x_i there exists $\tilde{t}_i$ such that

$$t_* - \varepsilon^{2\beta'} \le \tilde{t}_i \le t_*, \quad |\varphi(x_i, \tilde{t}_i)| \le \alpha. \tag{4.53}$$

Define $\hat{t} := t_* - 2\varepsilon^{2\beta'}$. Since $t_* \ge 2\varepsilon^{2\beta'}$, we have $\hat{t} \ge 0$. By (4.53),

$$\varepsilon^{2\beta'} \le \tilde{t}_i - \hat{t} \le 2\varepsilon^{2\beta'} \tag{4.54}$$

and the assumption of Lemma 4.5 is satisfied for $s = \tilde{t}_i, \ y = x_i, \ t = \hat{t}$ and $R_i := c_7(\tilde{t}_i + \varepsilon^2 - \hat{t})^{\frac{1}{2}}$ if $\varepsilon < \epsilon_3$. Hence we may conclude that

$$c_8 R_i^{n-1} \le \mu_{\hat{t}}^{\varepsilon}(B_{R_i}(x_i)) \quad \text{for } i = 1, \dots, N. \tag{4.55}$$

By (4.54), we have $c_7(\varepsilon^{2\beta'} + \varepsilon^2)^{\frac{1}{2}} \le R_i \le c_7(2\varepsilon^{2\beta'} + \varepsilon^2)^{\frac{1}{2}} \le 2c_7\varepsilon^{\beta'}$, which shows

$$c_{11}\varepsilon^{\beta'(n-1)} \le \mu_{\hat{t}}^{\varepsilon}(B_{2c_7\varepsilon^{\beta'}}(x_i)) \tag{4.56}$$

from (4.55) with $c_{11} := c_8 c_7^{n-1}$. Since $\{B_{2c_7\varepsilon^{\beta'}}(x_i)\}_{i=1}^N$ are pairwise disjoint and $B_{2c_7\varepsilon^{\beta'}}(x_i) \subset B_{2r+2c_7\varepsilon^{\beta'}}(y)$, (4.56) gives

$$N c_{11}\varepsilon^{\beta'(n-1)} \le \mu_{\hat{t}}^{\varepsilon}(B_{2r+2c_7\varepsilon^{\beta'}}(y)). \tag{4.57}$$

Hence the n-dimensional volume of A is estimated by (4.52) and (4.57)

$$\mathcal{L}^n(A) \le N\omega_n(10c_7\varepsilon^{\beta'})^n \le \frac{\omega_n(10c_7)^n\varepsilon^{\beta'}}{c_{11}}\mu_{\hat{t}}^{\varepsilon}(B_{2r+2c_7\varepsilon^{\beta'}}(y)).$$

By (4.38) and $r \ge \varepsilon^{\beta'}$,

$$\mathcal{L}^n(A) \le \frac{\omega_n(10c_7)^n\varepsilon^{\beta'}}{c_{11}} D_1\omega_{n-1}(2r + 2c_7\varepsilon^{\beta'})^{n-1} \le c_{12}\varepsilon^{\beta'} r^{n-1}, \tag{4.58}$$

where $c_{12} := \omega_n\omega_{n-1}(10c_7)^n(2 + 2c_7)^{n-1}D_1c_{11}^{-1}$. Hence by (4.26) and (4.58)

$$\int_{A\cap B_r(y)} \left(\frac{\varepsilon|\nabla\varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \right)_+ (x, t_*)\, dx \le \mathcal{L}^n(A) 10\varepsilon^{-\beta} \le 10c_{12}\varepsilon^{\beta'-\beta} r^{n-1}. \tag{4.59}$$

Next we estimate the surface energy on the complement of A which decays very quickly. Define $\phi \in \mathrm{Lip}(B_{2r}(y))$ such that

$$\phi(x) := \begin{cases} 1 & \text{if } x \in B_r(y)\backslash A, \\ 0 & \text{if } \mathrm{dist(x, B_r(y)\backslash A)} \geq \varepsilon^{\beta'}, \end{cases}$$

$$|\nabla\phi| \leq 2\varepsilon^{-\beta'} \quad \text{and} \quad 0 \leq \phi \leq 1.$$

By $r \geq \varepsilon^{\beta'}$, $2c_7\varepsilon^{\beta'} > \varepsilon^{\beta'}$ and the definitions of $\tilde{A}$ and ϕ, we have $\mathrm{spt}\phi \cap \tilde{A} = \emptyset$, hence

$$|\varphi(x,s)| \geq \alpha, \quad \text{for } x \in \mathrm{spt}\phi,\ \mathrm{s} \in [\mathrm{t}_* - \varepsilon^{2\beta'}, \mathrm{t}_*]. \tag{4.60}$$

For each j differentiate the Eq. (4.2) with respect to x_j, multiply $\phi^2 \frac{\partial\varphi}{\partial x_j}$, sum over j and integrate to obtain

$$\begin{aligned} &\frac{d}{dt}\int_\Omega \frac{1}{2}|\nabla\varphi|^2\phi^2\,dx + \int_\Omega (u \otimes \nabla\varphi \cdot \nabla^2\varphi + \nabla\varphi \otimes \nabla\varphi \cdot \nabla u)\phi^2\,dx \\ &= \int_\Omega \left(\nabla\varphi \cdot \Delta\nabla\varphi - \frac{W''(\varphi)}{\varepsilon^2}|\nabla\varphi|^2\right)\phi^2\,dx. \end{aligned} \tag{4.61}$$

By integration by parts and the Cauchy-Schwarz inequality (4.61) gives

$$\begin{aligned} \frac{d}{dt}\int_\Omega \frac{1}{2}|\nabla\varphi|^2\phi^2\,dx \ &\leq \frac{1}{2}\int_\Omega |u|^2|\nabla\varphi|^2\phi^2\,dx + \int_\Omega |\nabla\varphi|^2|\nabla u|\phi^2\,dx \\ &\quad + 4\int_\Omega |\nabla\phi|^2|\nabla\varphi|^2\,dx - \int_\Omega \frac{W''(\varphi)}{\varepsilon^2}|\nabla\varphi|^2\phi^2\,dx. \end{aligned} \tag{4.62}$$

By (4.60), $W''(\varphi) \geq \kappa$ on $\mathrm{spt}\phi$ for $t \in [t_* - \varepsilon^{2\beta'}, t_*]$. By (4.10) and the definition of ϕ, (4.62) gives

$$\begin{aligned} \frac{d}{dt}\int_\Omega \frac{1}{2}|\nabla\varphi|^2\phi^2\,dx &\leq \int_\Omega \left(\frac{\varepsilon^{-2\beta}}{2} + \varepsilon^{-1-\beta}\right)|\nabla\varphi|^2\phi^2\,dx + 16\varepsilon^{-2\beta'}\int_{\mathrm{spt}\phi} |\nabla\varphi|^2\,dx \\ &\quad - \frac{\kappa}{\varepsilon^2}\int_\Omega |\nabla\varphi|^2\phi^2\,dx \\ &\leq -\frac{\kappa}{2\varepsilon^2}\int_\Omega |\nabla\varphi|^2\phi^2\,dx + 16\varepsilon^{-2\beta'}\int_{\mathrm{spt}\phi} |\nabla\varphi|^2\,dx \end{aligned} \tag{4.63}$$

for small ε. By integrating (4.63) over $[t_* - \varepsilon^{2\beta'}, t_*]$, we obtain

$$\begin{aligned} \int_\Omega \frac{1}{2}|\nabla\varphi|^2\phi^2(x,t_*)\,dx &\leq e^{-\kappa\varepsilon^{2(\beta'-1)}}\int_\Omega \frac{1}{2}|\nabla\varphi|^2\phi^2(x, t_* - \varepsilon^{2\beta'})\,dx \\ &\quad + \int_{t_*-\varepsilon^{2\beta'}}^{t_*} e^{-\frac{\kappa}{\varepsilon^2}(t_*-\lambda)}16\varepsilon^{-2\beta'}\left(\int_{\mathrm{spt}\phi} |\nabla\varphi|^2(x,\lambda)\,dx\right)d\lambda. \end{aligned} \tag{4.64}$$

Define

$$M := \sup_{\lambda \in [t_* - \varepsilon^{2\beta'}, t_*]} \int_{\operatorname{spt}\phi} \frac{1}{2} |\nabla \varphi|^2 (x, \lambda) \, dx.$$

By (4.64) we have

$$\int_\Omega \frac{1}{2} |\nabla \varphi|^2 \phi^2 (x, t_*) \, dx \le \left(e^{-\kappa \varepsilon^{2(\beta'-1)}} + 32\kappa^{-1} \varepsilon^{2-2\beta'} \right) M. \tag{4.65}$$

By $\operatorname{spt}\phi \subset \mathrm{B}_{2\mathrm{r}}(\mathrm{y})$ and (4.38)

$$\varepsilon M \le \omega_{n-1} D_1 (2r)^{n-1}. \tag{4.66}$$

Since $B_r(y) \setminus A \subset \{\phi = 1\}$, we have

$$\int_{B_r(y)\setminus A} \frac{\varepsilon}{2} |\nabla \varphi|^2 (x, t_*) \, dx \le \int_\Omega \frac{\varepsilon}{2} |\nabla \varphi|^2 (x, t_*) \phi^2 \, dx. \tag{4.67}$$

Recall that $\beta' < 1$. By (4.65)–(4.67), we obtain for sufficiently small ε (depending only on κ)

$$\int_{B_r(y)\setminus A} \frac{\varepsilon}{2} |\nabla \varphi|^2 (x, t_*) \, dx \le 33\kappa^{-1} \varepsilon^{2-2\beta'} D_1 \omega_{n-1} (2r)^{n-1}. \tag{4.68}$$

By (4.59) and (4.68), and since $\beta' - \beta = \frac{1-\beta}{2} < 2 - 2\beta' = 1 - \beta$, we obtain (4.51) with an appropriate choice of c_{10}. □

Later in Sect. 7, we use the following estimate which follows from Lemma 4.6.

Corollary 4.1 *For any* $0 < r < \frac{1}{2}$, $\varepsilon \le \epsilon_4$ *and* $t \in [2\varepsilon^{2\beta'}, T] \cap [0, T_1]$, *we have*

$$\int_0^r \frac{d\tau}{\tau^n} \int_{B_\tau(y)} \left(\frac{\varepsilon |\nabla \varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \right)_+ (x, t) \, dx \le c_{10} \varepsilon^{\beta'-\beta} |\log \varepsilon| + 10\omega_n \varepsilon^{\beta'-\beta}. \tag{4.69}$$

Proof For the integration over the range $\tau \in (0, \varepsilon^{\beta'})$, we simply use the estimate (4.26). For the range $\tau \in (\varepsilon^{\beta'}, r)$, we use (4.51). □

Lemma 4.7 *There exists a constant* c_{13} *depending only on* n, c_1, c_2, p, q, T, D_0, W, β *such that for* $\varepsilon < \epsilon_4$, $t \in [0, T_1]$ *and* $t < s$, *we have*

$$\int_0^t \left\{ \frac{1}{2(s-\lambda)} \int_\Omega \left(\frac{\varepsilon |\nabla \varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \right)_+ \tilde{\rho}_{(y,s)}(x, \lambda) \, dx \right\} d\lambda \le c_{13} \varepsilon^{\beta'-\beta} |\log \varepsilon|. \tag{4.70}$$

Proof If $t \le 2\varepsilon^{2\beta'}$ then by using (4.26) and $\int \rho\, dx = \sqrt{4\pi(s-\lambda)}$ we have

$$\int_0^t \left\{ \frac{1}{2(s-\lambda)} \int_\Omega \left(\frac{\varepsilon |\nabla \varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \right)_+ \tilde{\rho}_{(y,s)}(x,\lambda)\, dx \right\} d\lambda$$
$$\le \int_0^t \frac{10\varepsilon^{-\beta}\sqrt{\pi}}{\sqrt{s-\lambda}}\, d\lambda \le 20\sqrt{2\pi}\,\varepsilon^{\beta'-\beta}. \tag{4.71}$$

By the similar argument, if $s > t \ge s - 2\varepsilon^{2\beta'}$ then we have

$$\int_{s-2\varepsilon^{2\beta'}}^t \left\{ \frac{1}{2(s-\lambda)} \int_\Omega \left(\frac{\varepsilon |\nabla \varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \right)_+ \tilde{\rho}_{(y,s)}(x,\lambda)\, dx \right\} d\lambda$$
$$\le 20\sqrt{2\pi}\,\varepsilon^{\beta'-\beta}. \tag{4.72}$$

Hence we only need to estimate integral over $[2\varepsilon^{2\beta'}, t]$ with $t \le s - 2\varepsilon^{2\beta'}$. First we estimate on $B_{\varepsilon^{\beta'}}(y)$. We compute using (4.26) and $s - t \ge 2\varepsilon^{2\beta'}$ that

$$\int_{2\varepsilon^{2\beta'}}^t \frac{1}{2(s-\lambda)} \int_{B_{\varepsilon^{\beta'}}} \left(\frac{\varepsilon |\nabla \varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \right)_+ \tilde{\rho}\, dx d\lambda$$
$$\le \int_{2\varepsilon^{2\beta'}}^t \frac{10\varepsilon^{-\beta}\varepsilon^{n\beta'}\omega_n}{2(s-\lambda)^{\frac{n+1}{2}}(\sqrt{4\pi})^{n-1}}\, d\lambda \le \frac{10\varepsilon^{\beta'-\beta}\omega_n}{(\sqrt{8\pi})^{n-1}(n-1)}. \tag{4.73}$$

On $\Omega \setminus B_{\varepsilon^{\beta'}}(y)$, by (4.51), $s - t \ge 2\varepsilon^{2\beta'}$ and computations similar to (4.37), we have

$$\int_{2\varepsilon^{2\beta'}}^t \frac{1}{2(s-\lambda)} \int_{\Omega \setminus B_{\varepsilon^{\beta'}}} \left(\frac{\varepsilon |\nabla \varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \right)_+ \tilde{\rho}\, dx d\lambda$$
$$\le \int_{2\varepsilon^{2\beta'}}^t \frac{d\lambda}{2(s-\lambda)^{\frac{n+1}{2}}(\sqrt{4\pi})^{n-1}}$$
$$\int_0^1 \left\{ \int_{B_{\frac{1}{2}} \cap \{x : e^{-\frac{|x-y|^2}{4(s-\lambda)}} \ge l\} \setminus B_{\varepsilon^{\beta'}}(y)} \left(\frac{\varepsilon |\nabla \varphi|^2}{2} - \frac{W(\varphi)}{\varepsilon} \right)_+ \right\} dl$$
$$\le c_{10} c(n) \varepsilon^{\beta'-\beta} \int_{2\varepsilon^{2\beta'}}^t \frac{e^{-\frac{1}{16(s-\lambda)}} + (s-\lambda)^{\frac{n-1}{2}}}{(s-\lambda)^{\frac{n+1}{2}}}\, d\lambda$$
$$\le c_{10} c(n) \varepsilon^{\beta'-\beta}(1 + \beta' \log(1/\varepsilon)). \tag{4.74}$$

By (4.71)–(4.74) we obtain the desired estimate. □

To utilize the formula (4.17), we next obtain the estimate for u.

Lemma 4.8 *There exists c_{14} depending only on n, p and q such that for any t_0, t_1 with $s > t_1 > t_0 \geq 0$ we have*

$$\int_{t_0}^{t_1}\int_\Omega \tilde{\rho}_{(y,s)}|u|^2\,d\mu_t^\varepsilon dt \leq c_{14}(t_1-t_0)^{\hat{p}}\|u\|^2_{L^q([t_0,t_1];(W^{1,p}(B_{\frac12}(y)))^n)} \sup_{t\in[t_0,t_1]} D(t), \tag{4.75}$$

where (1) $0 < \hat{p} = \frac{2pq-2p-nq}{pq}$ *when* $p < n$, (2) $\hat{p} < \frac{q-2}{q}$ *may be taken arbitrarily close to* $\frac{q-2}{q}$ *when* $p = n$ *(and c_{14} depends on $\hat{p}$), and* (3) $\hat{p} = \frac{q-2}{q}$ *when* $p > n$.

Proof First, consider the case $p < n$. By the Hölder inequality, for $l := \frac{p(n-1)}{2(n-p)}$ (which is ≥ 1 due to (2.13)) we have

$$\begin{aligned}\int_\Omega \tilde{\rho}|u|^2\,d\mu_t^\varepsilon &\leq \left(\int_\Omega |\eta^{\frac12}u|^{2l}\rho\,d\mu_t^\varepsilon\right)^{\frac1l}\left(\int_{B_{\frac12}(y)}\rho\,d\mu_t^\varepsilon\right)^{\frac{l-1}{l}}\\ &\leq (D(t))^{\frac{l-1}{l}}\left(\int_\Omega |u\eta^{\frac12}|^{2l}\rho\,d\mu_t^\varepsilon\right)^{\frac1l}\\ &\leq (D(t))^{\frac{l-1}{l}}\left(\frac{1}{(4\pi(s-t))^{\frac{n-1}{2}}}\int_\Omega |u\eta^{\frac12}|^{2l}\,d\mu_t^\varepsilon\right)^{\frac1l}.\end{aligned} \tag{4.76}$$

By (4.76) and (2.11) we have

$$\begin{aligned}\int_\Omega \tilde{\rho}|u|^2\,d\mu_t^\varepsilon &\leq \frac{D(t)}{(4\pi(s-t))^{\frac{n-1}{2l}}}\left(c(n,p)\left(\int_{B_{\frac12}(y)}|u|^p+|\nabla u|^p\,dx\right)^{\frac{n-1}{n-p}}\right)^{\frac1l}\\ &\leq \frac{c_{15}D(t)}{(4\pi(s-t))^{\frac{n-p}{p}}}\|u\|^2_{W^{1,p}(B_{\frac12}(y))},\end{aligned} \tag{4.77}$$

where $c_{15} = c_{15}(n,p)$. Hence by the Hölder inequality and (4.77) we obtain (with $\|u\| := \|u\|_{L^q([t_0,t_1];(W^{1,p}(B_{\frac12}(y))^n)}$)

$$\begin{aligned}\int_{t_0}^{t_1}\int_\Omega \tilde{\rho}|u|^2\,d\mu_t^\varepsilon dt &\leq c_{15}\|u\|^2\sup_{t\in[t_0,t_1]}D(t)\left(\int_{t_0}^{t_1}\frac{1}{(s-t)^{\frac{(n-p)q}{p(q-2)}}}\,dt\right)^{\frac{q-2}{q}}\\ &\leq c_{15}\|u\|^2\sup_{t\in[t_0,t_1]}D(t)c(n,p,q)((t_1-t_0)^{\frac{-(n-p)q}{p(q-2)}+1})^{\frac{q-2}{q}}\\ &\leq c(n,p,q)c_{15}(t_1-t_0)^{\frac{2pq-2p-nq}{pq}}\|u\|^2\sup_{t\in[t_0,t_1]}D(t).\end{aligned}$$

We remark that $(s-t_0)^\iota-(s-t_1)^\iota \leq (t_1-t_0)^\iota$ for $\iota\in(0,1)$ and $\frac{-(n-p)q}{p(q-2)}+1\in(0,1)$. By setting $c_{14} := c(n,p,q)c_{15}$, we obtain the desired estimate when $p < n$. For

$p = n$, since $W^{1,n}_{loc} \subset W^{1,p'}_{loc}$ for $p' < n$, we repeat the same argument as above for p close to n. Note that $\frac{2pq-2p-nq}{pq} \uparrow \frac{q-2}{q}$ as $p \uparrow n$. This gives the estimate for $p = n$ case. For $p > n$, $\sup_{B_{\frac{1}{2}}(y)} |\eta^{\frac{1}{2}} u| \le c(n, p)\|u\|_{W^{1,p}(B_{\frac{1}{2}}(y))}$. Thus $\int \tilde{\rho}|u|^2 \, d\mu^{\varepsilon}_t \le c(n, p)D(t)\|u\|^2_{W^{1,p}(B_{\frac{1}{2}}(y))}$. This gives the desired estimate for $p > n$. □

Proposition 4.2 *There exist $c_{16} > 1$ depending only on n, $c_{17} > 0$ depending only on n, p, q and $\epsilon_5 > 0$ depending only on n, p, q, c_1, c_2, D_0, T, W, β with the following property. For t_0, t_1 with $T_1 \ge t_1 > t_0 \ge 0$ and $t_1 - t_0 \le 1$, suppose $D(t_1) = c_{16}D(t_0)$ and $\sup_{t\in[t_0,t_1]} D(t) \le c_{16}D(t_0)$. Then, if $\varepsilon < \epsilon_5$, we have*

$$(t_1 - t_0)^{\hat{p}} \|u\|^2_{L^q([t_0,t_1];(W^{1,p}(\Omega))^n)} \ge c_{17}, \tag{4.78}$$

where $\hat{p}$ is as in Lemma 4.8.

Proof First, for any $s > t_0$, by direct computation and by the definition of $D(t_0)$, we have

$$\int_{\Omega} \tilde{\rho}_{(y,s)} \, d\mu^{\varepsilon}_{t_0} \le D(t_0). \tag{4.79}$$

Let $c_{16} > 1$ be a constant defined by

$$c_{16} := \max\left\{ \frac{2 \cdot 4^{n-1}}{\omega_{n-1}}, \frac{(2 + c_3)(4\pi)^{\frac{n-1}{2}}}{\omega_{n-1}e^{-\frac{1}{4}}} \right\}. \tag{4.80}$$

By definition, c_{16} depends only on n. Suppose that t_1 satisfies the assumptions. Recalling the definition of $D(t_1)$, we have the following three possibilities, (a) $D(t_1) = \mu^{\varepsilon}_{t_1}(\Omega)$, (b) there exists $B_r(y) \subset \Omega$ such that $D(t_1) = \frac{1}{\omega_{n-1}r^{n-1}}\mu^{\varepsilon}_{t_1}(B_r(y))$ and $r \ge \frac{1}{4}$, and (c) the same as (b) except that $r < \frac{1}{4}$. For (b), we have the following

$$\frac{\omega_{n-1}}{4^{n-1}} D(t_1) \le \omega_{n-1}r^{n-1}D(t_1) = \mu^{\varepsilon}_{t_1}(B_r(y)) \le \mu^{\varepsilon}_{t_1}(\Omega).$$

Since $\omega_{n-1}/4^{n-1} \le 1$, either (a) or (b), we have

$$\frac{\omega_{n-1}}{4^{n-1}} D(t_1) \le \mu^{\varepsilon}_{t_1}(\Omega). \tag{4.81}$$

Then, by (4.39), we obtain with (4.81) that

$$c_{16}D(t_0) = D(t_1) \le \frac{4^{n-1}}{\omega_{n-1}} \mu^{\varepsilon}_{t_1}(\Omega) \le \frac{4^{n-1}}{\omega_{n-1}} \left(D(t_0) + c_6(t_1 - t_0)^{\frac{q-2}{q}} \|u\|^2 c_{16}D(t_0) \right), \tag{4.82}$$

where $\|u\| := \|u\|_{L^q([t_0,t_1];(W^{1,p}(\Omega))^n)}$. By (4.80), $\frac{4^{n-1}}{\omega_{n-1}} \leq \frac{c_{16}}{2}$, thus (4.82) shows

$$\frac{1}{2c_6} \leq (t_1 - t_0)^{\frac{q-2}{q}} \|u\|^2. \tag{4.83}$$

This is the conclusion deduced from (a) and (b). Next consider the case (c). Let $s = t_1 + r^2$. By (4.17), (4.70), (4.75) and (4.39), we have

$$\begin{aligned}\int_\Omega \tilde{\rho}_{(y,s)}\, d\mu^\varepsilon_{t_1} &\leq \int_\Omega \tilde{\rho}_{(y,s)}\, d\mu^\varepsilon_{t_0} + c_{13}\varepsilon^{\beta\prime-\beta}|\log\varepsilon| + c_{14}c_{16}D(t_0)(t_1-t_0)^{\hat{p}}\|u\|^2 \\ &\quad + c_3(t_1-t_0)(D(t_0) + c_6c_{16}D(t_0)(t_1-t_0)^{\frac{q-2}{q}}\|u\|^2).\end{aligned} \tag{4.84}$$

We compute using $\eta = 1$ on $B_{\frac{1}{4}}(y)$ and $r \leq \frac{1}{4}$ that

$$\begin{aligned}\int_\Omega \tilde{\rho}_{(y,s)}\, d\mu^\varepsilon_{t_1} &\geq \int_{B_r(y)} \rho_{(y,s)}\, d\mu^\varepsilon_{t_1} \geq \frac{e^{-\frac{1}{4}}}{(4\pi)^{\frac{n-1}{2}} r^{n-1}} \mu^\varepsilon_{t_1}(B_r(y)) \\ &= \frac{c_{16}D(t_0)\omega_{n-1}e^{-\frac{1}{4}}}{(4\pi)^{\frac{n-1}{2}}} \geq (2 + c_3)D(t_0),\end{aligned} \tag{4.85}$$

where $s = t_1 + r^2$, the properties of t_1 and c_{16} are used. By (4.79), (4.84) and (4.85) give (using also $t_1 - t_0 \leq 1$)

$$\begin{aligned}D(t_0) &\leq c_{13}\varepsilon^{\beta\prime-\beta}|\log\varepsilon| + c_{14}c_{16}D(t_0)(t_1-t_0)^{\hat{p}}\|u\|^2 \\ &\quad + c_3c_6c_{16}D(t_0)(t_1-t_0)^{2-\frac{2}{q}}\|u\|^2.\end{aligned} \tag{4.86}$$

Since $D(t_0) \geq 1$ by definition, we may restrict ε depending on c_{13} (see Lemma 4.7) so that $c_{13}\varepsilon^{\beta\prime-\beta}|\log\varepsilon| < 1/2$, for example. Now, examining the dependence of constants, we obtain (4.78) from (4.83) and (4.86) by choosing an appropriate $c_{17} > 0$. Here we also use $\hat{p} < 2 - \frac{2}{q}$ and $t_1 - t_0 \leq 1$. □

Proof of Theorem 4.1. We first choose $0 < T_b \leq 1$ so that

$$T_b^{\hat{p}} c_2^2 \leq c_{17} \tag{4.87}$$

holds. Due to the dependence of c_{17}, T_b depends only on n, p, q, c_2. Then set

$$D_1 := D_0 c_{16}^{[T/T_b]+1} (\geq 2D_0 \text{ by } (4.80)), \tag{4.88}$$

so that D_1 depends only on n, p, q, c_2, T, D_0. Finally restrict $\varepsilon < \epsilon_5$ as in Proposition 4.2. Now we claim that

$$D(t) \leq D_0 c_{16}^{[t/T_b]+1} \tag{4.89}$$

holds for all $t \in [0, T]$, thus proving $D(t) \le D_1$ for all $t \in [0, T]$ and $T_1 = T$. Suppose there exists $0 < t \le T$ such that (4.89) fails. Then there must exist some $0 < T_1 < T$ such that $D(t) \le D_0 c_{16}^{[t/T_b]+1}$ for all $t \in [0, T_1]$ and $D(T_1) = D_0 c_{16}^{[T_1/T_b]+1}$. Note that $D(t) \le D_1$ for $t \in [0, T_1]$, satisfying (4.38). If $T_1 < T_b$, we apply Proposition 4.2 with $t_0 = 0$ and $t_1 = T_1$. We have $D(T_1) = c_{16} D_0$ and $\sup_{t\in[0,T_1]} D(t) \le c_{16} D_0$. Thus (4.78) shows

$$T_1^{\hat{p}} c_2^2 \ge c_{17},$$

but this contradicts $T_1 < T_b$ and (4.87). Thus, we have $T_1 \ge T_b$. If $T_1 \in [T_b, 2T_b)$, then $D(T_1) = D_0 c_{16}^2$. Thus there must exist $t_0 \in [T_b, T_1)$ such that $D(t_0) = c_{16} D_0$ and $T_1 - t_0 < T_b$ (note that $D(t) \le D_0 c_{16}$ for all $t \in [0, T_b)$). By Proposition 4.2 with $t_1 = T_1$, we have $(T_1 - t_0)^{\hat{p}} c_2^2 \ge c_{17}$, again contradicting $T_1 - t_0 < T_b$ and (4.87). Continuing this manner, we conclude that $T_1 = T$, which is a contradiction. Thus we proved that (4.89) holds for all $t \in [0, T]$. Also this concludes the proof of Theorem 4.1. □

Since we proved $T = T_1$, i.e., the assumption (4.38) is true for all $[0, T]$, all the estimates in this section hold with T_1 replaced by T. In particular, we have the following monotonicity formula which follows from (4.17), (4.75) and (4.70).

Theorem 4.2 *Under the same assumptions of Theorem 4.1, if $\varepsilon < \epsilon_1$ and for $s > t_1 > t_0$, $t_0, t_1 \in [0, T]$, and $y \in \Omega$ we have*

$$\int_\Omega \tilde{\rho}\, d\mu_t^\varepsilon \Big|_{t=t_0}^{t_1} + \int_{t_0}^{t_1} \frac{dt}{2(s-t)} \int_\Omega |\xi_\varepsilon| \tilde{\rho}\, dx \le c_{14} c_2^2 (t_1 - t_0)^{\hat{p}} D_1 + c_{13} \varepsilon^{\beta\prime - \beta} |\log \varepsilon| + c_3 e^{-\frac{1}{128(s-t_0)}} (t_1 - t_0) D_1, \tag{4.90}$$

where $\tilde{\rho} = \tilde{\rho}_{(y,s)}(x, t)$ and ξ_ε are defined as in (4.15) and (4.16), and $\hat{p}$ is as in Lemma 4.8.

The point of the right-hand side is that it is bounded independent of ε, and it can be made arbitrarily small when $\varepsilon \to 0$ and $t_0 \to t_1$.

5 Existence of limit measures

In this section we construct a sequence of approximate diffused interface solution for (1.2), given any bounded hypersurface $M_0 = \partial\Omega_0$ which is C^1, and any vector field u satisfying (2.15). We then prove that we may extract a subsequence which converges to a family of Radon measures $\{\mu_t\}_{t\ge 0}$.

We first construct a convergent sequence of domains Ω_0^i with C^∞ boundary M_0^i which converges in C^1 topology. This can be carried out by locally representing M_0 by a C^1 graph and by some suitable mollification. Let d_i be the signed distance function to M_0^i which is positive inside of Ω_0^i, and which is smooth in some r_i-neighborhood of M_0^i. Let $h_i \in C^\infty(\mathbb{R})$ be a monotone increasing function such that $h_i(s) = s$ for

$0 \le s \le r_i/3$, $h_i(s) = r_i/2$ for $s > 2r_i/3$, $h_i'(s) \le 1$ for $s > 0$ and $h_i(s) = -h_i(-s)$ for $s < 0$. Then define $\tilde{d}_i(x) := h_i(d_i(x))$ for $x \in \Omega$. We next choose a sequence of $\varepsilon_i > 0$ so that

$$\lim_{i\to\infty} \sqrt{\varepsilon_i}/r_i = 0. \tag{5.1}$$

We define the initial data $(\varphi_{\varepsilon_i})$ differently depending on $\Omega = \mathbb{T}^n$ or $\mathbb{R}^n$ as follows.

For $\Omega = \mathbb{T}^n$, we define

$$(\varphi_{\varepsilon_i})_0 := \Psi\left(\frac{\tilde{d}_i(x)}{\varepsilon_i}\right). \tag{5.2}$$

Here and in the following, Ψ is the solution for $\Psi'' = W'(\Psi)$ (and $\Psi' = \sqrt{2W(\Psi)}$) with $\Psi(0) = 0$. For $\Omega = \mathbb{R}^n$, we will truncate the function to be -1 outside of a compact set as follows. Due to the definition, note that for $x \in \mathbb{R}^n$ with $\mathrm{dist}(x, \Omega_0^i) \ge 2r_i/3$, we have $\tilde{d}_i(x) = -r_i/2$. Choose a sufficiently large $R > 0$ such that

$$\{x : \mathrm{dist}(x, \Omega_0^i) \le 2r_i/3\} \subset B_R \tag{5.3}$$

for all i. Then we have $\tilde{d}_i(x) = -r_i/2$ on $\mathbb{R}^n \setminus B_R$. Let $g : \mathbb{R}^+ \to [0, 1]$ be a smooth decreasing function such that $g(r) = 1$ for $0 \le r \le R$, $g(r) = 0$ for $R + 1 \le r < \infty$ and $|g'| \le 2$. Define

$$(\varphi_{\varepsilon_i})_0(x) := g(|x|)\Psi\left(\frac{\tilde{d}_i(x)}{\varepsilon_i}\right) + g(|x|) - 1. \tag{5.4}$$

Then $(\varphi_{\varepsilon_i})_0(x) = \Psi\left(\frac{\tilde{d}_i(x)}{\varepsilon_i}\right)$ on B_R, and it smoothly changes from $\Psi(-r_i/2\varepsilon_i)$ to -1 as $|x|$ increases from R to $R + 1$. We may show from $\Psi' = \sqrt{2W(\Psi)}$ that $0 < \Psi(-r_i/2\varepsilon_i) + 1 \le c\exp(-c' r_i/\varepsilon_i)$ for some positive constants c, c' depending only on W. Thus the difference between $(\varphi_{\varepsilon_i})_0$ and -1 is exponentially small on $B_{R+1} \setminus B_R$ by (5.1), and $(\varphi_{\varepsilon_i})_0(x) = -1$ on $\mathbb{R}^n \setminus B_{R+1}$.

For both cases, one can check that (4.7) is satisfied for $(\varphi_{\varepsilon_i})_0$ with some i-independent c_1, where we may need to take a smaller ε_i depending on the growth of C^3 norm of the graph functions representing M_0^i. We fix β

$$\beta = \frac{1}{4}, \tag{5.5}$$

though any $0 < \beta < 1/2$ can be chosen. Using the fact that Ψ solves $\Psi' = \sqrt{2W(\Psi)}$ and $|\nabla \tilde{d}_i| \le 1$, one can check that (4.9) is satisfied for all i. We may also assume that

$$\begin{aligned} &\lim_{i\to\infty} \int_\Omega \left| \frac{(\varphi_{\varepsilon_i})_0 + 1}{2} - \chi_{\Omega_0} \right| dx = 0, \\ &\lim_{i\to\infty} \Big(\frac{\varepsilon_i |\nabla (\varphi_{\varepsilon_i})_0|^2}{2} + \frac{W((\varphi_{\varepsilon_i})_0)}{\varepsilon_i}\Big)\, dx = \sigma \|\nabla \chi_{\Omega_0}\| = \sigma \mathcal{H}^{n-1}\lfloor_{M_0} \end{aligned} \tag{5.6}$$

where the second identity is in the sense of measure convergence. We may also assume, due to the assumption that M_0 is C^1, that we have some D_0 depending on M_0 such that $D(0)$ as in (4.5) corresponding to $(\varphi_{\varepsilon_i})_0$ is uniformly bounded by D_0 independent of i.

We next let $T_i = i$ so that $\lim_{i\to\infty} T_i = \infty$, and let $\{u_i\}_{i=1}^{\infty}$ be a sequence of C^∞ vector fields with compact support such that $\|u_i - u\|_{L^q([0,T_i];(W^{1,p}(\Omega))^n)} \to 0$ as $i \to \infty$, which can be constructed by the standard density argument. Then for each i we associate $j(i)$ so that (4.10) is satisfied, i.e.,

$$\sup_{\Omega\times[0,T_i]} \{|u_i|,\ \varepsilon_{j(i)}|\nabla u_i|\} \le \varepsilon_{j(i)}^{-\beta} \tag{5.7}$$

for all i, and at the same time, $\varepsilon_{j(i)} < \epsilon_1$ where ϵ_1 is determined by Theorem 4.1 corresponding to D_0, $T = T_i$ and $c_2 = \|u_i\|_{L^q([0,T_i];(W^{1,p}(\Omega))^n)}$. We relabel $\varepsilon_{j(i)}$ as ε_i and u_i as u_{ε_i}.

With these choices, for each $i \in \mathbb{N}$, we solve (4.2) and (4.3) on $\Omega \times [0, T_i]$ with initial data $(\varphi_{\varepsilon_i})_0$ and u replaced by u_{ε_i}. For $\Omega = \mathbb{T}^n$, the standard parabolic PDE theory shows the existence of classical solution which we denote φ_{ε_i}. The maximum principle shows (4.6). Due to the choice of ε_i, for each fixed $T > 0$, we have all the assumptions of Theorem 4.1 satisfied on $[0, T]$ for all sufficiently large i, thus we have (4.13). The same can be said about Theorem 4.2. For $\Omega = \mathbb{R}^n$ and for each fixed i, we construct the solution by domain approximation. Namely, for each $k \in \mathbb{N}$ with $k > 3R$ (where R is defined in (5.3)), solve

$$\begin{cases} \partial_t\varphi + u_{\varepsilon_i}\cdot\nabla\varphi = \Delta\varphi - \frac{W'(\varphi)}{\varepsilon_i^2} & \text{on } B_k \times [0, T_i], \\ \varphi = (\varphi_{\varepsilon_i})_0 & \text{on } B_k \times \{0\}, \\ \varphi = -1 & \text{on } \partial B_k \times [0, T_i]. \end{cases} \tag{5.8}$$

By the standard parabolic existence theory, there exists a classical solution which we denote by $\varphi_{\varepsilon_i,k}$. By the maximum principle, we have $-1 \le \varphi_{\varepsilon_i,k} < 1$. We claim that

$$\varphi_{\varepsilon_i,k}(x,t) < \Psi\left(\frac{3R + t\|u_{\varepsilon_i}\|_{L^\infty} - |x|}{\varepsilon_i}\right) =: \psi_{\varepsilon_i}(x,t) \tag{5.9}$$

for all k by the maximum principle. To see this, on $\partial B_k \times [0, T_i]$, we have $\varphi_{\varepsilon_i,k}(x,t) = -1 < \psi_{\varepsilon_i}(x,t)$ by (5.8) and (5.9). On $B_k \times \{0\}$ where $\varphi_{\varepsilon_i,k} = (\varphi_{\varepsilon_i})_0$, we may check $\psi_{\varepsilon_i} > (\varphi_{\varepsilon_i})_0$ as follows. When $|x| \ge R+1$, $\psi_{\varepsilon_i}(x,0) > -1 = (\varphi_{\varepsilon_i})_0(x)$, and when $R \le |x| \le R+1$, $(\varphi_{\varepsilon_i})_0(x) \approx -1 < \Psi(0) < \psi_{\varepsilon_i}(x,0)$. When $|x| < R$,

$$(\varphi_{\varepsilon_i})_0(x) \le \Psi\left(\frac{\tilde{d}_i(x)}{\varepsilon_i}\right) < \Psi\left(\frac{2R}{\varepsilon_i}\right) \le \Psi\left(\frac{3R - |x|}{\varepsilon_i}\right) = \psi_{\varepsilon_i}(x,0)$$

since $|\tilde{d}_i(x)| \le |d_i(x)| < 2R$ from $M_0^i \subset B_R$. ψ_{ε_i} is a super-solution since, for $|x| \neq 0$,

$$\partial_t \psi_{\varepsilon_i} + u_{\varepsilon_i} \cdot \nabla \psi_{\varepsilon_i} - \Delta \psi_{\varepsilon_i} + \frac{W'(\psi_{\varepsilon_i})}{\varepsilon_i^2} = \frac{\Psi'(\psi_{\varepsilon_i})}{\varepsilon_i} \left(\|u_{\varepsilon_i}\|_{L^\infty} + \frac{n-1}{|x|} - \frac{x}{|x|} \cdot u_{\varepsilon_i} \right) > 0.$$

We note that $\varphi_{\varepsilon_i,k}$ cannot touch ψ_{ε_i} from below at $|x| = 0$. Thus we may prove (5.9) by the standard argument of the maximum principle. Now let $k \to \infty$ and we may prove that $\varphi_{\varepsilon_i,k}$ converge to a solution φ_{ε_i} of (4.2) on $\mathbb{R}^n \times [0, T_i]$ satisfying $-1 \le \varphi_{\varepsilon_i} \le \psi_{\varepsilon_i}$. Hence, we have (4.6). Due to (5.9), for each fixed i, we have the exponential approach of φ_{ε_i} to -1 as $|x| \to \infty$, which is (4.8). Thus, in the case of $\Omega = \mathbb{R}^n$, we have all the assumptions of Theorem 4.1 satisfied and we may obtain the desired conclusion.

We next prove that there exists a family of Radon measures $\{\mu_t\}_{t\ge 0}$ such that, after choosing a subsequence, $\mu_t^{\varepsilon_{i_j}} \to \mu_t$ as $j \to \infty$ for all $t \ge 0$.

Proposition 5.1 *Corresponding to $T > 0$ and $\phi \in C_c^2(\Omega; \mathbb{R}^+)$, there exists $c_{18} > 0$ depending only on n, p, q, T, D_0, c_2 and $\|\phi\|_{C^2(\Omega)}$ such that, for all i with $i > T$ and $\mu_t^{\varepsilon_i}$ constructed as above, the function*

$$\mu_t^{\varepsilon_i}(\phi) - c_{18} \left(\int_0^t \|u_{\varepsilon_i}(\cdot, s)\|^2_{W^{1,p}(\Omega)}\, ds + t \right) \tag{5.10}$$

of t is monotone decreasing on $[0, T]$.

Proof By (4.2) and integration by parts we have

$$\begin{aligned} \frac{d}{dt}\mu_t^{\varepsilon_i}(\phi) = \int_\Omega -\varepsilon_i \phi \left(\Delta\varphi_{\varepsilon_i} - \frac{W'(\varphi_{\varepsilon_i})}{\varepsilon_i^2} \right)^2 - \varepsilon_i \nabla\phi \cdot \nabla\varphi_{\varepsilon_i} \left(\Delta\varphi_{\varepsilon_i} - \frac{W'(\varphi_{\varepsilon_i})}{\varepsilon_i^2} \right) \\ + \varepsilon_i \phi \left(\Delta\varphi_{\varepsilon_i} - \frac{W'(\varphi_{\varepsilon_i})}{\varepsilon_i^2} \right) u_{\varepsilon_i} \cdot \nabla\varphi_{\varepsilon_i} + \varepsilon_i (\nabla\varphi_{\varepsilon_i} \cdot \nabla\phi)(u_{\varepsilon_i} \cdot \nabla\varphi_{\varepsilon_i})\, dx. \end{aligned} \tag{5.11}$$

By the Cauchy-Schwarz inequality and estimating as in the proof of Lemma 4.4, we have

$$\begin{aligned} \frac{d}{dt}\mu_t^{\varepsilon_i}(\phi) &\le \int_\Omega \varepsilon_i |\nabla\varphi_{\varepsilon_i}|^2 \frac{|\nabla\phi|^2}{\phi} + \varepsilon_i \phi |u_{\varepsilon_i}|^2 |\nabla\varphi_{\varepsilon_i}|^2\, dx \\ &\le 4(\sup \|\nabla^2\phi\|) D(t) + \sup|\phi| c(n, p) D(t) \|u_{\varepsilon_i}(\cdot, t)\|^2_{W^{1,p}(\Omega)}. \end{aligned} \tag{5.12}$$

Thus with a suitable constant independent of i and Theorem 4.1, we have (5.10). □

Proposition 5.2 (See [28,33]) *There exist a family of Radon measures $\{\mu_t\}_{t\ge 0}$ and a subsequence* (*denoted by the same index*) *such that for all $t \ge 0$,*

$$\lim_{i\to\infty} \mu_t^{\varepsilon_i} = \mu_t \quad \textit{as Radon measures.}$$

Proof Fix $T > 0$ and $\phi \in C_c^2(\Omega; \mathbb{R}^+)$. By the Cauchy-Schwarz inequality and $q > 2$,

$$\int_{t_1}^{t_2} \|u_{\varepsilon_i}(\cdot, s)\|^2_{W^{1,p}(\Omega)}\, ds \le (t_2 - t_1)^{\frac{q-2}{q}} \|u_{\varepsilon_i}\|^2_{L^q([0,T];(W^{1,p}(\Omega))^n)}$$

for $0 \le t_1 < t_2 \le T$. Hence the last term of (5.10) is uniformly bounded in Hölder continuous norm with exponent $\frac{q-2}{q}$. Thus by the Ascoli-Arzelà compactness theorem, there exists a subsequence which converges uniformly on $[0, T]$. By the monotone decreasing property due to Proposition 5.1, we can choose a subsequence such that $\mu_t^{\varepsilon_i}(\phi)$ converges on a co-countable set $B(\phi) \subset [0, T]$. Choose a countable set $\{\phi_k\}_{k=1}^\infty \subset C_c^2(\Omega; \mathbb{R}^+)$ which is dense in $C_c(\Omega; \mathbb{R}^+)$. By the similar argument we can choose a subsequence such that $\mu_t^{\varepsilon_i}(\phi_k)$ converges on a co-countable set $B = \cap_{k=1}^\infty B(\phi_k)$. For any $k \ge 1$ we define $\mu_t(\phi_k) = \lim_{i\to\infty} \mu_t^{\varepsilon_i}(\phi_k)$ for $t \in B$. Then we may define $\mu_t(\phi) = \lim_{i\to\infty} \mu_t^{\varepsilon_i}(\phi)$ for any $\phi \in C_c(\Omega; \mathbb{R}^+)$ and for any $t \in B$ since $\{\phi_k\}_{k=1}^\infty$ is dense in $C_c(\Omega; \mathbb{R}^+)$ and the measures are uniformly bounded. Since $[0, T]\backslash B$ is countable, we can choose a subsequence so that $\mu_t^{\varepsilon_i}(\phi_k)$ converges on $[0, T]\backslash B$ for any k. Thus we have the limit $\mu_t(\phi)$ for all $\phi \in C_c(\Omega; \mathbb{R}^+)$ and for all $t \in [0, T]$. Now by letting $T \to \infty$ and by diagonal argument, we may choose a subsequence so that $\mu_t^{\varepsilon_i}(\phi)$ converges for all $t \ge 0$ and $\phi \in C_c(\Omega; \mathbb{R}^+)$. □

We also denote, after choosing a further subsequence,

Definition 5.1 Let μ be a measure on $\Omega \times [0, \infty)$ such that $d\mu = \lim_{j\to\infty} d\mu_t^{\varepsilon_j} dt$ locally as measures.

Since $\sup_{t\in[0,T]} \mu_t^{\varepsilon_j}(\Omega)$ is bounded uniformly in j for all T, the dominated convergence theorem shows $d\mu = d\mu_t\, dt$. On the other hand, note that $\mathrm{spt}\mu$ may not be the same as $\cup_{t\ge 0}\mathrm{spt}\mu_t \times \{t\}$. In the following section we also use the following notation.

Definition 5.2 Define $(\mathrm{spt}\mu)_t \subset \Omega$ as $(\mathrm{spt}\mu)_t := \{x \in \Omega \,:\, (x, t) \in \mathrm{spt}\mu\}$.

We have the following inclusion.

Lemma 5.1 *For all $t > 0$,*

$$\mathrm{spt}\, \mu_t \subset (\mathrm{spt}\, \mu)_t. \tag{5.13}$$

Proof Suppose $x \in \mathrm{spt}\, \mu_{t_0}$ and assume for a contradiction that $(x, t_0) \notin \mathrm{spt}\, \mu$. Then there exists $r > 0$ such that $\mu(B_r(x) \times (t_0 - r^2, t_0 + r^2)) = 0$. Take $\phi \in C_c^2(B_r(x); \mathbb{R}^+)$ with $\phi = 1$ on $B_{r/2}(x)$. Since $x \in \mathrm{spt}\, \mu_{t_0}$, we have $\mu_{t_0}(\phi) > 0$. By Proposition 5.1 and 5.13, $\mu_t(\phi) - c_{18}(\int_0^t \|u(\cdot, s)\|^2_{W^{1,p}}\, ds + t)$ is monotone decreasing. Thus one sees that for all sufficiently small $h > 0$, we have $\mu_{t_0 - h}(\phi) \ge \mu_{t_0}(\phi) - o(1) \ge \mu_{t_0}(\phi)/2$ where $o(1) \to 0$ as $h \to 0$. Since $d\mu = d\mu_t dt$, this contradicts $(x, t_0) \notin \mathrm{spt}\, \mu$. □

6 Rectifiability of limit measures

Throughout this section, let φ_{ε_i}, $\mu_t^{\varepsilon_i}$, u_{ε_i}, μ_t and μ be as in Sect. 5 and let $\tilde{\rho}_{(y,s)}$, e_{ε_i} and ξ_{ε_i} be as in (4.15) and (4.16). We fix arbitrary $T > 0$ and let c_2 be as in (4.11) with this T. Note that all the estimates in the previous two sections hold in $[0, T]$ for all sufficiently large i (such that $T_i > T$). For simplicity we often drop i from these quantities. In this section we prove that for a.e. $t \geq 0$, there exists a countably $(n-1)$-rectifiable set M_t such that $\mu_t = \theta(x,t)\mathcal{H}^{n-1}\lfloor_{M_t}$, where θ is a non-negative $\mathcal{H}^{n-1}$ measurable function. The important ingredient for the proof is the vanishing of the discrepancy measure defined below. As stated in the introduction, the content of this section is based on [28] with some modifications coming from the transport term. First we note

Lemma 6.1 *Let φ_{ε_i} and $\mu_t^{\varepsilon_i}$ be the sequences constructed in Sect. 5. Then there exist a subsequence (denoted by the same index) and a Radom measure $|\xi|$ such that*

$$\lim_{i\to\infty}\int_{t_0}^{t_1}\int_\Omega |\xi_{\varepsilon_i}|\phi\,dxdt = \int_{t_0}^{t_1}\int_\Omega \phi\,d|\xi| \tag{6.1}$$

for all $0 \leq t_0 < t_1 < \infty$ and $\phi \in C_c(\Omega\times[0,\infty))$.

Due to the uniform estimate $\sup_{i\in\mathbb{N}}\sup_{t\in[0,T]}\mu_t^{\varepsilon_i}(\Omega)$ for any fixed T, the existence of such subsequence follows from the weak compactness of measures. Since $|\xi|$ measures the difference between the two terms in $\mu_t^{\varepsilon_i}$ in the limit, we may call $|\xi|$ as a discrepancy measure. Unlike $\mu_t^{\varepsilon_i}$, which converges to μ_t for all $t \geq 0$, note that we do not claim any convergence of $|\xi_{\varepsilon_i}(\cdot,t)|\,dx$ in general. Instead, we will prove

Theorem 6.1 $|\xi| = 0$ *on* $\Omega\times[0,\infty)$.

6.1 Forward density lower bound

Lemma 6.2 *There exist $1 > \gamma_1, \eta_1 > 0$ depending only on n, c_1, c_2, p, q, T, W, D_0 and $1 > \eta_2 > 0$ depending only on n, c_1, W with the following property. Given $0 \leq t < s < T/2$ with $s - t \leq \eta_1$, set $r := \sqrt{2(s-t)}$ and $t' := s + r^2/2$. If $x \in \Omega$ satisfies*

$$\int_\Omega \tilde{\rho}_{(y,s)}(x,t)\,d\mu_s(y) < \eta_2, \tag{6.2}$$

then $(B_{\gamma_1 r}(x)\times\{t'\})\cap \mathrm{spt}\mu = \emptyset$.

Remark 6.1 Note that $t < s < t' < T$ with $s = \frac{t'+t}{2}$. The Lemma says that, unless there is at least a certain amount of measure, there would be no measure later in the neighborhood. The monotonicity formula (4.90) plays a crucial role for such conclusion.

Proof Assume for a contradiction that $(x', t') \in \mathrm{spt}\mu$ for some $x' \in B_{\gamma_1 r}(x)$ under the assumption of (6.2), where γ_1 will be chosen later. Then there is a sequence $\{(x_j,t_j)\}_{j=1}^\infty$ and $\{\varepsilon_{i(j)}\}_{j=1}^\infty$ such that $\lim_{j\to\infty}(x_j,t_j) = (x',t')$ and $|\varphi_{\varepsilon_{i(j)}}(x_j,t_j)| < \alpha$ for all j. We relegate its proof to Lemma 6.3. We re-index $i(j)$ as j. Then just as in the proof of (4.45), there exists $\eta_2 = \eta_2(n, c_1, W) > 0$ such that

$$
\begin{aligned}
3\eta_2 &\le \int_{B_{\varepsilon_j}(x_j)} \frac{W(\varphi_{\varepsilon_j}(y,t_j))}{\varepsilon_j} \tilde{\rho}_{(x_j,t_j+\varepsilon_j^2)}(y,t_j)\,dy \\
&\le \int_\Omega \tilde{\rho}_{(x_j,t_j+\varepsilon_j^2)}(y,t_j)\,d\mu_{t_j}^{\varepsilon_j}(y).
\end{aligned} \tag{6.3}
$$

We use Theorem 4.2. By restricting $t' - s \le \eta_1$ small so that

$$
c_{14}c_2^2(t_j - s)^{\hat{p}} D_1 + c_3 e^{-\frac{1}{128(t_j+\varepsilon_j^2-s)}}(t_j - s)D_1 < \eta_2
$$

in (4.90) for all sufficiently large j, we obtain

$$
\begin{aligned}
&\int_\Omega \tilde{\rho}_{(x_j,t_j+\varepsilon_j^2)}(y,t_j)\,d\mu_{t_j}^{\varepsilon_j}(y) \\
&\quad \le \int_\Omega \tilde{\rho}_{(x_j,t_j+\varepsilon_j^2)}(y,s)\,d\mu_s^{\varepsilon_j}(y) + c_{13}\varepsilon_j^{\beta\prime-\beta}|\log\varepsilon_j| + \eta_2.
\end{aligned} \tag{6.4}
$$

Letting $j \to \infty$, we obtain by (6.3) and (6.4)

$$
2\eta_2 \le \int_\Omega \tilde{\rho}_{(x',t')}(y,s)\,d\mu_s(y). \tag{6.5}
$$

We next want to change the center of the kernel from x' to x. Fix $0 < \delta < 1/2$ so that $2\delta D_1 < \eta_2$. Corresponding to δ, a direct computation shows that we may choose $\gamma_1 > 0$ so that

$$
\int_\Omega \tilde{\rho}_{(x',t')}(y,s)\,d\mu_s(y) \le \delta D_1 + (1+\delta)\int_\Omega \tilde{\rho}_{(x,t')}(y,s)\,d\mu_s(y) \tag{6.6}
$$

if $|x - x'| \le \gamma_1 r$. By the choice of δ, (6.5) and (6.6) show

$$
\eta_2 \le \int_\Omega \tilde{\rho}_{(x,t')}(y,s)\,d\mu_s(y). \tag{6.7}
$$

Finally, since $t'-s = s-t$, we have $\tilde{\rho}_{(x,t')}(y,s) = \tilde{\rho}_{(y,s)}(x,t)$. This is a contradiction to (6.2). Thus we proved $(x',t') \notin \mathrm{spt}\mu$. □

Lemma 6.3 *Assume* $(x',t') \in \mathrm{spt}\mu$. *Then there are sequences* $\{(x_j,t_j)\}_{j=1}^\infty$ *and* $\{\varepsilon_{i(j)}\}_{j=1}^\infty$ *such that* $\lim_{j\to\infty}(x_j,t_j) = (x',t')$ *and* $|\varphi_{\varepsilon_{i(j)}}(x_j,t_j)| < \alpha$ *for all* j.

Proof If the claim were not true, there would be $0 < r_0 < 1/2$ such that

$$
\inf_{B_{r_0}(x')\times[t'-r_0^2,t'+r_0^2]} |\varphi_{\varepsilon_i}| \ge \alpha \tag{6.8}
$$

for all sufficiently large i. Let $\phi \in C_c^2(B_{r_0}(x'))$ be a function such that $|\nabla\phi| \le 2/r_0$, $0 \le \phi \le 1$ on $B_{r_0}(x')$ and $\phi = 1$ on $B_{r_0/3}(x')$. Then the same computations following (4.60) using (6.8) show

$$\frac{d}{dt}\int_\Omega \frac{1}{2}|\nabla\varphi_{\varepsilon_i}|^2\phi^2\,dx \le -\frac{\kappa}{2\varepsilon_i^2}\int_\Omega |\nabla\varphi_{\varepsilon_i}|^2\phi^2\,dx + 16r_0^{-2}\int_{\mathrm{spt}\phi}|\nabla\varphi_{\varepsilon_i}|^2\,dx$$

for $t \in [t'-r_0^2, t'+r_0^2]$. Writing $M_i := \sup_{\lambda\in[t'-r_0^2,t'+r_0^2]}\int_{\mathrm{spt}\phi}\frac{1}{2}|\nabla\varphi_{\varepsilon_i}(x,\lambda)|^2\,dx$, and proceeding similarly as in (4.65), we obtain

$$\int_\Omega \frac{1}{2}|\nabla\varphi_{\varepsilon_i}(\cdot,\lambda)|^2\phi^2\,dx \le \left(e^{-\frac{\kappa}{\varepsilon_i^2}(\lambda-t'+r_0^2)} + \frac{32\varepsilon_i^2}{r_0^2\kappa}\right)M_i \tag{6.9}$$

for $\lambda \in [t'-r_0^2, t'+r_0^2]$. Since $\varepsilon_i M_i$ is uniformly bounded, we see from (6.9) that

$$\lim_{i\to\infty}\sup_{\lambda\in[t'-\frac{r_0^2}{2},t'+r_0^2]}\int_\Omega \frac{\varepsilon_i}{2}|\nabla\varphi_{\varepsilon_i}(\cdot,\lambda)|^2\phi^2\,dx = 0. \tag{6.10}$$

Next, due to (6.8) and the continuity of φ_{ε_i}, we may assume $1 \ge \varphi_{\varepsilon_i} \ge \alpha$ on $B_{r_0}(x')\times[t'-r_0^2, t'+r_0^2]$ without loss of generality. Otherwise, we have $-1 \le \varphi_{\varepsilon_i} \le -\alpha$ and we may argue similarly. In the following, we use

$$W'(s)(s-1) \ge (s-1)^2\kappa \ge c(W)W(s) \tag{6.11}$$

for some $c(W) > 0$ if $s \in [\alpha, 1]$. Multiply the equation (4.2) by $(\varphi_{\varepsilon_i}-1)\phi^2$ and integrate over $Q := \Omega\times[t'-r_0^2, t'+r_0^2]$. By integration by parts, the Cauchy-Schwarz inequality, $|\varphi_{\varepsilon_i}-1| \le 1$ and (6.11), one obtains

$$c(W)\int_Q \phi^2\frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i^2}\,dxdt \le \frac{1}{2}\int_\Omega \phi^2\,dx + \int_Q 2|\nabla\phi|^2 + \frac{1}{2}|u_{\varepsilon_i}|^2\phi^2\,dxdt. \tag{6.12}$$

Since the right-hand side of (6.12) is uniformly bounded, we obtain

$$\lim_{i\to\infty}\int_Q \phi^2\frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i}\,dxdt = 0. \tag{6.13}$$

The estimates (6.10) and (6.13) show that

$$\lim_{i\to\infty}\int_{t'-r_0^2/2}^{t'+r_0^2}\mu_t^{\varepsilon_i}(\phi^2)\,dt = 0. \tag{6.14}$$

By Fatou's lemma, Proposition 5.13 and (6.14), we have

$$\int_{t'-r_0^2/2}^{t'+r_0^2}\mu_t(\phi^2)\,dt = 0. \tag{6.15}$$

This proves that $(x',t') \notin \mathrm{spt}\mu$. □

Corollary 6.1 *Let $U \subset \Omega$ be open. For $0 < t \le T$, there exists c_{19} depending only on $n, c_1, c_2, p, q, T, W, D_0$ with the property that*

$$\mathcal{H}^{n-1}((\mathrm{spt}\mu)_{\mathrm{t}} \cap \mathrm{U}) \le \mathrm{c}_{19} \liminf_{\mathrm{r}\to 0} \mu_{\mathrm{t}-\mathrm{r}^2}(\mathrm{U}) \tag{6.16}$$

and

$$\mathcal{H}^{n-1}(\mathrm{spt}\,\mu_t \cap U) \le c_{19} \liminf_{r\to 0} \mu_{t-r^2}(U). \tag{6.17}$$

Proof We only need to prove the result for every compact set $K \subset U$. Set $X_t = (\mathrm{spt}\mu)_{\mathrm{t}} \cap \mathrm{K}$. For any $(x,t) \in X_t$, by the same argument leading to (6.5), we have

$$2\eta_2 \le \int_\Omega \tilde{\rho}_{(x,t)}(y, t-r^2)\, d\mu_{t-r^2}(y) \tag{6.18}$$

for sufficiently small $r > 0$. For $0 < L < 1/(2r)$, using the upper density ratio bound, we have

$$\begin{aligned} &\int_{\Omega\setminus B_{rL}(x)} \tilde{\rho}_{(x,t)}(y, t-r^2)\, d\mu_{t-r^2}(y) \\ &\quad \le D_1 \omega_{n-1} (\pi)^{-\frac{n-1}{2}} \int_{L^2/4}^{\infty} s^{\frac{n-1}{2}} e^{-s}\, ds. \end{aligned} \tag{6.19}$$

Thus by choosing sufficiently large L depending only on n, D_1 and η_2, (6.18) and (6.19) show

$$\eta_2 \le \int_{B_{rL}(x)} \tilde{\rho}_{(x,t)}(y, t-r^2)\, d\mu_{t-r^2}(y). \tag{6.20}$$

Since $\tilde{\rho}_{(x,t)}(\cdot, t-r^2) \le (4\pi)^{-(n-1)/2} r^{-(n-1)}$, from (6.20) we obtain

$$(4\pi)^{\frac{n-1}{2}} r^{n-1} \eta_2 \le \mu_{t-r^2}(B_{rL}(x)). \tag{6.21}$$

Let $\mathcal{B} = \{\bar{B}_{rL}(x) \subset U \,|\, x \in X_t\}$ which is the covering of X_t by closed balls centered at $x \in X_t$. By the Besicovitch covering theorem, there exist a finite sub-collection $\mathcal{B}_1, \ldots, \mathcal{B}_{B(n)}$ such that each $\mathcal{B}_i$ is a pairwise disjoint family of closed balls and

$$X_t \subset \cup_{i=1}^{B(n)} \cup_{\bar{B}_{rL}(x_j)\in\mathcal{B}_i} \bar{B}_{rL}(x_j). \tag{6.22}$$

Let $\mathcal{H}^{n-1}_\delta$ be defined as in [41], so that $\mathcal{H}^{n-1} = \lim_{\delta\downarrow 0} \mathcal{H}^{n-1}_\delta$. By the definition, (6.21) and (6.22) we obtain

$$
\begin{aligned}
\mathcal{H}^{n-1}_{2rL}(X_t) &\le \sum_{i=1}^{B(n)} \sum_{\bar{B}_{rL}(x_j)\in\mathcal{B}_i} \omega_{n-1}(rL)^{n-1} \\
&\le \sum_{i=1}^{B(n)} \frac{\omega_{n-1}L^{n-1}}{(4\pi)^{\frac{n-1}{2}}\eta_2} \sum_{\bar{B}_{rL}(x_j)\in\mathcal{B}_i} \mu_{t-r^2}(B_{rL}(x_j)) \\
&\le \sum_{i=1}^{B(n)} \frac{\omega_{n-1}L^{n-1}}{(4\pi)^{\frac{n-1}{2}}\eta_2}\mu_{t-r^2}(U) = \frac{\omega_{n-1}L^{n-1}B(n)}{(4\pi)^{\frac{n-1}{2}}\eta_2}\mu_{t-r^2}(U).
\end{aligned}
$$

By setting c_{19} to be the constant above and letting $r \downarrow 0$, we obtain (6.16). The second inequality (6.17) follows immediately from (6.16) and Lemma 5.1. □

Lemma 6.4 *For $1 \le T < \infty$, let η_2 be as in Lemma 6.2 corresponding to T. Define*

$$
Z_T := \left\{ (x,t) \in \operatorname{spt}\mu \ : \ 0 \le t \le T/2,\ \limsup_{s\downarrow t} \int_\Omega \tilde{\rho}_{(y,s)}(x,t)\, d\mu_s(y) \le \eta_2/2 \right\}.
$$

Then we have $\mu(Z_T) = 0$.

Proof For $0 < \tau \le \eta_1$, where η_1 is as in Lemma 6.2, define

$$
Z^\tau := \left\{ (x,t) \in \operatorname{spt}\mu \ : 0 \le t < T/2,\ \int_\Omega \tilde{\rho}_{(y,s)}(x,t)\, d\mu_s(y) < \eta_2,\ \ \forall s \in (t, t+\tau] \right\}.
$$

Note that $Z_T \subset \cup_{m=1}^\infty Z^{\tau_m}$ for some $\{\tau_m\}_{m=1}^\infty$ with $\lim_{m\to\infty}\tau_m = 0$. Hence we only need to prove $\mu(Z^\tau) = 0$. In the following we fix $0 < \tau \le \eta_1$. For $0 \le t \le T/2$ and $x \in \Omega$, set

$$
P_\tau(x,t) := \{(x',t') \ : \ \tau > |t-t'| > \gamma_1^{-2}|x-x'|^2\}, \tag{6.23}
$$

where γ_1 is as in Lemma 6.2. For $(x,t) \in Z^\tau$, we use Lemma 6.2 to prove

$$
P_\tau(x,t) \cap Z^\tau = \emptyset. \tag{6.24}
$$

Suppose for a contradiction that $(x',t') \in P_\tau(x,t) \cap Z^\tau$. Suppose first that $t' > t$. Set $r := \sqrt{t'-t}$ and $s := (t'+t)/2$ so that $t' = s + r^2/2$. Note that we have $|x-x'| < \gamma_1 r$ by $(x',t') \in P_\tau(x,t)$. Since $s - t < \tau \le \eta_1$, we may apply Lemma 6.2 to conclude that $(x,t) \in Z^\tau$ implies $(x',t') \notin \operatorname{spt}\mu$, and in particular, $(x',t') \notin Z^\tau$, which is a contradiction. Next suppose that $t' < t$. We change the role of (x,t) and (x',t') in the previous case, and conclude that $(x',t') \in Z^\tau$ implies $(x,t) \notin Z^\tau$, which is again a contradiction. This proves (6.24). Next, for $(x_0,t_0) \in \Omega \times [\tau/2, T/2]$, define

$$
Z^{\tau,x_0,t_0} = Z^\tau \cap B_{\frac{1}{2}}(x_0) \times (t_0 - \tau/2, t_0 + \tau/2). \tag{6.25}
$$

Then Z^τ can be covered by an at most countable union of Z^{τ,x_j,t_j} with a suitable choice of $\{(x_j,t_j)\}$. Thus we only need to prove $\mu(Z^{\tau,x_0,t_0}) = 0$. With arbitrary $0 < r \le \gamma_1\sqrt{\tau}$, consider a family of closed balls $\{\bar{B}_r(x)\}_{(x,t)\in Z^{\tau,x_0,t_0}}$ and apply the Besicovitch covering theorem. Then we have a finite subfamily $\bar{B}_r(x_1),\ldots,\bar{B}_r(x_N)$ with $(x_j,t_j)\in Z^{\tau,x_0,t_0}$ $(j=1,\ldots,N)$ and

$$\{x\in B_{\frac{1}{2}}(x_0) : (x,t)\in Z^{\tau,x_0,t_0}\} \subset \cup_{j=1}^N \bar{B}_r(x_j), \quad Nr^n \le 2B(n)(1/2)^n. \tag{6.26}$$

Note that for each $j=1,\ldots,N$, by (6.24) and (6.25), we have

$$Z^{\tau,x_0,t_0}\cap \bar{B}_r(x_j)\times(t_0-\tau/2,t_0+\tau/2)\subset \bar{B}_r(x_j)\times(t_0-\tau/2,t_0+\tau/2)\setminus P_\tau(x_j,t_j). \tag{6.27}$$

The inclusions (6.26) and (6.27) show

$$Z^{\tau,x_0,t_0}\subset \cup_{j=1}^N \bar{B}_r(x_j)\times(t_0-\tau/2,t_0+\tau/2)\setminus P_\tau(x_j,t_j). \tag{6.28}$$

Since $\bar{B}_r(x_j)\times(t_0-\tau/2,t_0+\tau/2)\setminus P_\tau(x_j,t_j)\subset \bar{B}_r(x_j)\times[t_j-\gamma_1^{-2}r^2, t_j+\gamma_1^{-2}r^2]$, from (6.28) we obtain

$$Z^{\tau,x_0,t_0}\subset \cup_{j=1}^N \bar{B}_r(x_j)\times[t_j-\gamma_1^{-2}r^2, t_j+\gamma_1^{-2}r^2]. \tag{6.29}$$

Since $d\mu = d\mu_t dt$, (6.29), (4.13) and (6.26) show

$$\begin{aligned}\mu(Z^{\tau,x_0,t_0}) &\le \sum_{j=1}^N \int_{t_j-\gamma_1^{-2}r^2}^{t_j+\gamma_1^{-2}r^2} \mu_t(\bar{B}_r(x_j))\,dt \le 2\omega_{n-1}D_1 r^{n+1}\gamma_1^{-2}N \\ &\le 2^{2-n}\omega_{n-1}B(n)D_1 r\gamma_1^{-2}.\end{aligned} \tag{6.30}$$

Since $0 < r \le \gamma_1\sqrt{\tau}$ is arbitrary, (6.30) shows $\mu(Z^{\tau,x_0,t_0}) = 0$. This concludes the proof. □

6.2 Vanishing of ξ

First we remark the following

Lemma 6.5 *For* $1 \le T < \infty$ *there exists* c_{20} *depending only on* n, c_1, c_2, p, q, T, W, D_0 *with the following property. For any* $(y,s)\in\Omega\times(0,T)$, *we have*

$$\int_{\Omega\times(0,s)} \frac{\tilde{\rho}_{(y,s)}(x,t)}{s-t}\,d|\xi|(x,t) \le c_{20}. \tag{6.31}$$

Proof In (4.90), set $t_0 = 0$ and $t_1 = s-\epsilon$ for $0<\epsilon<s$. We simply let $\varepsilon_i \to 0$ and we set the supremum of the right-hand side of (4.90) (with no ε term) plus D_0 (coming from the left-hand side) to be c_{20}. Then letting $\epsilon\to 0$, we obtain (6.31). □

We are ready to prove Theorem 6.1.

Proof We integrate (6.31) with respect to $d\mu_s ds$ over $\Omega \times (0, T)$ and use Fubini's theorem to obtain

$$\int_{\Omega\times(0,T)} \left(\int_{\Omega\times(t,T)} \frac{\tilde{\rho}_{(y,s)}(x,t)}{s-t}\, d\mu_s(y)ds \right) d|\xi|(x,t) \le c_{20} D_1 T. \tag{6.32}$$

The finiteness of (6.32) shows

$$\int_{\Omega\times(t,T)} \frac{\tilde{\rho}_{(y,s)}(x,t)}{s-t}\, d\mu_s(y)ds < \infty \tag{6.33}$$

for $|\xi|$ a.e. $(x,t) \in \Omega \times (0,T)$. Next, we claim that, whenever (6.33) holds at (x,t), we have

$$\lim_{s\downarrow t} \int_{\Omega} \tilde{\rho}_{(y,s)}(x,t)\, d\mu_s(y) = 0. \tag{6.34}$$

We use the monotonicity formula (4.90) for the proof. Set $\lambda := \log(s-t)$ and

$$h(s) := \int_{\Omega} \tilde{\rho}_{(y,s)}(x,t)\, d\mu_s(y).$$

After the change of variable, (6.33) is equivalent to

$$\int_{-\infty}^{\log(T-t)} h(t+e^{\lambda})\, d\lambda < \infty. \tag{6.35}$$

We fix $\theta \in (0,1]$ in the following. Corresponding to this θ, by (6.35), there exists a decreasing sequence $\{\lambda_i\}_{i=1}^{\infty}$ such that

$$\lambda_i \downarrow -\infty, \quad \lambda_i - \lambda_{i+1} \le \theta, \quad h(t+e^{\lambda_i}) \le \theta. \tag{6.36}$$

For arbitrary $\lambda \in (-\infty, \lambda_1)$, choose i such that $\lambda \in [\lambda_i, \lambda_{i-1})$. Then by (4.90) (with $\varepsilon \to 0$) applied with $t_0 = t + e^{\lambda_i} < t_1 = t + e^{\lambda}$, we have

$$\begin{aligned} h(t+e^{\lambda}) &= \int_{\Omega} \tilde{\rho}_{(y,t+e^{\lambda})}(x,t)\, d\mu_{t+e^{\lambda}}(y) = \int_{\Omega} \tilde{\rho}_{(y,t+2e^{\lambda})}(x,t+e^{\lambda})\, d\mu_{t+e^{\lambda}}(y) \\ &\le \int_{\Omega} \tilde{\rho}_{(y,t+2e^{\lambda})}(x,t+e^{\lambda_i})\, d\mu_{t+e^{\lambda_i}}(y) + o(1) \end{aligned} \tag{6.37}$$

where $\lim_{\theta\to 0} o(1) = 0$. On the other hand, by (6.36) we have

$$\int_{\Omega} \tilde{\rho}_{(y,t+e^{\lambda_i})}(x,t)\, d\mu_{t+e^{\lambda_i}}(y) = h(t+e^{\lambda_i}) \le \theta. \tag{6.38}$$

By direct calculation,

$$\int_{\Omega} \tilde{\rho}_{(y,t+2e^{\lambda})}(x, t + e^{\lambda_i})\, d\mu_{t+e^{\lambda_i}}(y) \\ \leq o(1) + \int_{B_{M\sqrt{2e^{\lambda}-e^{\lambda_i}}}(y)} \tilde{\rho}_{(y,t+2e^{\lambda})}(x, t + e^{\lambda_i})\, d\mu_{t+e^{\lambda_i}}(y) \tag{6.39}$$

where $\lim_{M\to\infty} o(1) = 0$ and the convergence does not depend on θ. For any fixed M, we have

$$\sup_{x\in B_{M\sqrt{2e^{\lambda}-e^{\lambda_i}}}(y)} \tilde{\rho}_{(y,t+2e^{\lambda})}(x, t + e^{\lambda_i})/\tilde{\rho}_{(y,t+e^{\lambda_i})}(x, t) \\ \leq \exp\left(M^2(e^{\lambda-\lambda_i} - 1)/2\right) \leq 1 + o(1) \tag{6.40}$$

where $\lim_{\theta\to 0} o(1) = 0$. The inequalities (6.37)–(6.40) show that $h(t + e^{\lambda})$ is made arbitrarily small for all $\lambda < \lambda_1$ and prove (6.34). Finally define $a(x, t) := \limsup_{s\downarrow t} \int_{\Omega} \tilde{\rho}_{(y,s)}(x, t)\, d\mu_s(y)$ and note that $\Omega \times (0, T)$ may be split into two disjoint sets

$$A \cup B := \{(x, t) : a(x, t) = 0\} \cup \{(x, t) : a(x, t) > 0\}.$$

The claim (6.34) proved $|\xi|(B) = 0$. On the other hand, by Lemma 6.4 we have $\mu(A) = 0$. Since $|\xi| \leq \mu$ by definition, this proves $|\xi|(\Omega \times (0, T)) = 0$. Since $T > 0$ is arbitrary, we have $|\xi|(\Omega \times (0, \infty)) = 0$. □

6.3 Associated varifolds and rectifiability theorem

We have so far obtained μ_t as a limit of Radon measures $\{\mu_t^{\varepsilon_i}\}_{i=1}^{\infty}$. To prove the rectifiability of μ_t for a.e. $t \geq 0$, we now consider a sequence of varifolds which are naturally associated with $\{\mu_t^{\varepsilon_i}\}_{i=1}^{\infty}$.

Definition 6.1 For $\varphi_{\varepsilon_i}(\cdot, t)$, we define $V_t^{\varepsilon_i} \in \mathbf{V}_{n-1}(\Omega)$ as follows. For $\phi \in C_c(G_{n-1}(\Omega))$,

$$V_t^{\varepsilon_i}(\phi) := \int_{\Omega\cap\{|\nabla\varphi_{\varepsilon_i}(x,t)|\neq 0\}} \phi\left(x, I - \frac{\nabla\varphi_{\varepsilon_i}(x,t)}{|\nabla\varphi_{\varepsilon_i}(x,t)|} \otimes \frac{\nabla\varphi_{\varepsilon_i}(x,t)}{|\nabla\varphi_{\varepsilon_i}(\cdot,t)|}\right) d\mu_t^{\varepsilon_i}(x). \tag{6.41}$$

Lemma 6.6 *For $g = (g_1, \ldots, g_n) \in C_c^1(\Omega; \mathbb{R}^n)$, we have*

$$\delta V_t^{\varepsilon_i}(g) = \int_{\Omega} (g \cdot \nabla\varphi_{\varepsilon_i}) \left(\varepsilon_i \Delta\varphi_{\varepsilon_i} - \frac{W'(\varphi_{\varepsilon_i})}{\varepsilon_i}\right) dx \\ - \int_{\Omega\cap\{|\nabla\varphi_{\varepsilon_i}|=0\}} \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i} I \cdot \nabla g\, dx \\ + \int_{\Omega\cap\{|\nabla\varphi_{\varepsilon_i}|\neq 0\}} \nabla g \cdot \left(\frac{\nabla\varphi_{\varepsilon_i}}{|\nabla\varphi_{\varepsilon_i}|} \otimes \frac{\nabla\varphi_{\varepsilon_i}}{|\nabla\varphi_{\varepsilon_i}|}\right) \xi_{\varepsilon_i}\, dx. \tag{6.42}$$

Proof We omit i in the following. The first variation of V_t^ε with respect to g is

$$\delta V_t^\varepsilon(g) = \int_{G_{n-1}(\Omega)} \nabla g(x) \cdot S \, dV_t^\varepsilon(x, S)$$
$$= \int_{\Omega \cap \{|\nabla \varphi_\varepsilon| \neq 0\}} \nabla g \cdot \left(I - \frac{\nabla \varphi_\varepsilon}{|\nabla \varphi_\varepsilon|} \otimes \frac{\nabla \varphi_\varepsilon}{|\nabla \varphi_\varepsilon|} \right) \left(\frac{\varepsilon}{2} |\nabla \varphi_\varepsilon|^2 + \frac{W}{\varepsilon} \right) dx. \tag{6.43}$$

By repeated integration by parts, we have

$$\int_{\Omega \cap \{|\nabla \varphi_\varepsilon| \neq 0\}} \nabla g \cdot I \, \frac{\varepsilon}{2} |\nabla \varphi_\varepsilon|^2 \, dx = -\varepsilon \int_\Omega \sum_{j,l=1}^n g_j (\varphi_\varepsilon)_{x_j x_l} (\varphi_\varepsilon)_{x_l} \, dx$$
$$= \varepsilon \int_\Omega \nabla g \cdot (\nabla \varphi_\varepsilon \otimes \nabla \varphi_\varepsilon) + (g \cdot \nabla \varphi_\varepsilon) \Delta \varphi_\varepsilon \, dx. \tag{6.44}$$

Also by integration by parts,

$$\int_{\Omega \cap \{|\nabla \varphi_\varepsilon| \neq 0\}} \nabla g \cdot I \, \frac{W}{\varepsilon} \, dx = - \int_{\Omega \cap \{|\nabla \varphi_\varepsilon| = 0\}} \nabla g \cdot I \, \frac{W}{\varepsilon} \, dx - \int_\Omega (g \cdot \nabla \varphi_\varepsilon) \frac{W'}{\varepsilon} \, dx. \tag{6.45}$$

Now substituting (6.44) and (6.45) into (6.43), we obtain (6.42). □

Proposition 6.1 *For a.e.* $t \geq 0$, μ_t *is rectifiable, and any convergent subsequence* $\{V_t^{\varepsilon_{i_j}}\}_{j=1}^\infty$ *with*

$$\liminf_{j \to \infty} \int_\Omega \varepsilon_{i_j} \left(\Delta \varphi_{\varepsilon_{i_j}}(x, t) - \frac{W'(\varphi_{\varepsilon_{i_j}}(x, t))}{\varepsilon_{i_j}^2} \right)^2 dx < \infty \tag{6.46}$$

converges to the unique varifold associated with μ_t.

Proof By Theorem 6.1 and by the dominated convergence theorem, we have

$$\lim_{i \to \infty} \int_\Omega |\xi_{\varepsilon_i}(\cdot, t)| \, dx = 0. \tag{6.47}$$

for full sequence for a.e. $t \geq 0$. By Lemma 4.4, we see that

$$\int_0^T \int_\Omega \varepsilon_i \left(\Delta \varphi_{\varepsilon_i} - \frac{W'}{\varepsilon_i^2} \right)^2 dx dt \leq 2E_0.$$

Thus, by Fatou's lemma, we have

$$\liminf_{i\to\infty}\int_\Omega \varepsilon_i\left(\Delta\varphi_{\varepsilon_i}(x,t)-\frac{W'(\varphi_{\varepsilon_i}(x,t))}{\varepsilon_i^2}\right)^2 dx<\infty \tag{6.48}$$

for a.e. $t\geq 0$. Suppose $t\geq 0$ satisfies both (6.47) and (6.48). Since $\|V_t^{\varepsilon_i}\|(\Omega)\leq \mu_t^{\varepsilon_i}(\Omega)$ is uniformly bounded in i, by the weak compactness theorem for measures, there exists a convergent subsequence $\{V_t^{\varepsilon_{i_j}}\}_{j=1}^\infty$ which satisfies (6.46) and which converges to a varifold V_t. Due to Proposition 5.13 and (6.47), we have

$$\|V_t\|=\mu_t. \tag{6.49}$$

Next, a standard measure theoretic argument (see for example [41, 3.2(2)]) shows

$$\mu_t\left(\left\{x\in \operatorname{spt}\mu_t\,:\,\limsup_{r\downarrow 0}\frac{\mu_t(B_r(x))}{\omega_{n-1}r^{n-1}}\leq s\right\}\right)\leq 2^{n-1}s\mathcal{H}^{n-1}(\operatorname{spt}\mu_t) \tag{6.50}$$

for any $s>0$. By (6.17), $\mathcal{H}^{n-1}(\operatorname{spt}\mu_t)<\infty$, thus (6.50) shows

$$\mu_t(\{x\in\operatorname{spt}\mu_t\,:\,\lim_{r\downarrow 0}r^{1-n}\mu_t(B_r(x))=0\})=0. \tag{6.51}$$

The two equalities (6.49) and (6.51) show that

$$V_t=V_t\lfloor_{\{x\in\Omega\,:\,\limsup_{r\downarrow 0}r^{1-n}\|V_t\|(B_r(x))>0\}\times\mathbf{G}(n,n-1)}. \tag{6.52}$$

Next we use (6.42). For any fixed $g\in C_c^1(\Omega;\mathbb{R}^n)$, (6.47) shows that the limits of the last two terms of (6.42) are both 0. Thus we have

$$\begin{aligned}\lim_{j\to\infty}|\delta V_t^{\varepsilon_{i_j}}(g)|&\leq\liminf_{j\to\infty}\left(\int_\Omega\varepsilon_{i_j}|\nabla\varphi_{\varepsilon_{i_j}}|^2\,dx\right)^{1/2}\\&\quad\times\left(\int_\Omega\varepsilon_{i_j}\left(\Delta\varphi_{\varepsilon_{i_j}}-\frac{W'}{\varepsilon_{i_j}^2}\right)^2dx\right)^{1/2}\end{aligned} \tag{6.53}$$

for g with $\sup|g|\leq 1$. Since the right-hand side of (6.53) does not depend on g and since $\delta V_t^{\varepsilon_{i_j}}(g)\to\delta V_t(g)$, we have

$$\sup_{g\in C_c^1(\Omega;\mathbb{R}^n),\ \sup|g|\leq 1}|\delta V_t(g)|<\infty$$

which shows that the total variation $\|\delta V_t\|$ is a Radon measure. Allard's rectifiability theorem [1] shows that the right-hand side of (6.52) is rectifiable, and hence so is V_t. Once we know that V_t is rectifiable, V_t is determined uniquely by $\|V_t\|=\mu_t$.

In particular, this shows that μ_t is rectifiable. The argument up to this point is valid for any convergent subsequence with (6.46) and (6.47). On the other hand, note that μ_t does not depend on the choice of subsequence $\{V_t^{\varepsilon_{i_j}}\}_{j=1}^{\infty}$. Since μ_t determines V_t uniquely, any converging subsequence of $\{V_t^{\varepsilon_i}\}_{i=1}^{\infty}$ with (6.46) and (6.47) has the same limit V_t. This completes the proof. □

7 Integrality of limit measures

In this section we prove that the density function of μ_t is integer-valued μ_t a.e. modulo division by σ.

7.1 Separating sheets

We prove in this subsection that, if a set of appropriate quantities are controlled, then we have a lower bound on a measure in terms of a sum of densities of vertically aligned points. As the name of the present subsection indicates, what one carries out in essence is to decompose the domain horizontally so that each separated domain contains approximately one sheet of diffused interface. The original idea comes from [1] and it has been first used in the context of the diffused interface problem in [27].

Lemma 7.1 *Suppose*

(1) $N \in \mathbb{N}$, *Y is a finite subset of* $\mathbb{R}^n$, $0 < R < \infty$, $1 < M < \infty$, $0 < a < \infty$, $0 < \varepsilon < 1$, $0 < \varrho < \infty$, $0 < E_0 < \infty$ *and* $-\infty \le l_1 < l_2 \le \infty$.
(2) *Y has no more than $N+1$ elements, and* $Y \subset \{(0,\ldots,0,x_n) : l_1 + a < x_n < l_2 - a\}$. *Moreover* $|x - z| > 3a$ *for* $x, z \in Y$ *with* $x \neq z$.
(3) $(M+1)\mathrm{diam Y} < \mathrm{R}$, *and put* $\tilde{R} := M\mathrm{diam Y}$.
(4) *We have* $\varphi \in C^2(\{y \in \mathbb{R}^n : \mathrm{dist(y, Y)} < \mathrm{R}\})$.
(5) *For all* $x = (0,\ldots,0,x_n) \in Y$,

$$\int_a^R \frac{d\tau}{\tau^n} \int_{B_\tau(x)\cap\{y_n = l_j\}} |e_\varepsilon(y_n - x_n) - \varepsilon\varphi_{x_n}(y - x)\cdot\nabla\varphi|\, d\mathcal{H}^{n-1}(y) \le \varrho \tag{7.1}$$

for $j = 1, 2$, *where e_ε is defined as in* (4.16).
(6) *For all* $x \in Y$ *and* $a \le r \le R$,

$$\int_{B_r(x)} |\xi_\varepsilon| + (1 - (\nu_n)^2)\varepsilon|\nabla\varphi|^2 + \varepsilon|\nabla\varphi|\left|\Delta\varphi - \frac{W'(\varphi)}{\varepsilon^2}\right| dy \le \varrho r^{n-1}, \tag{7.2}$$

where ξ_ε is defined as in (4.16) *and* $\nu = (\nu_1,\ldots,\nu_n) = \frac{\nabla\varphi}{|\nabla\varphi|}$.
(7) *For all* $x \in Y$,

$$\int_a^R \frac{d\tau}{\tau^n} \int_{B_\tau(x)} (\xi_\varepsilon)_+\, dy \le \varrho. \tag{7.3}$$

(8) *For all* $x \in Y$ *and* $a \le r \le R$,

$$\int_{B_r(x)} \varepsilon |\nabla \varphi|^2 \, dy \le E_0 r^{n-1}. \tag{7.4}$$

Then we have the following:

(A) *With* $S := \{x : l_1 < x_n < l_2\}$ *and for all* $x \in Y$ *and* $a \le r < R$,

$$\frac{1}{r^{n-1}} \int_{B_r(x)\cap S} e_\varepsilon \le \frac{1}{R^{n-1}} \int_{B_R(x)\cap S} e_\varepsilon + \varrho(3+R). \tag{7.5}$$

(B) *There exists* $l_3 \in (l_1, l_2)$ *such that* $|x_n - l_3| \ge a$ *and*

$$\begin{aligned} &\int_a^{\tilde{R}} \frac{d\tau}{\tau^n} \int_{B_\tau(x)\cap\{y_n = l_3\}} |e_\varepsilon (y_n - x_n) - \varepsilon \varphi_{x_n} (y-x)\cdot \nabla\varphi| \, d\mathcal{H}^{n-1}(y) \\ &\le 3(N+1)NM\left(\varrho + E_0^{\frac{1}{2}} \varrho^{\frac{1}{2}}\right) \end{aligned} \tag{7.6}$$

for any $x = (0, \cdot, 0, x_n) \in Y$.

(C) *Put*

$$Y_1 := Y \cap \{x : l_1 < x_n < l_3\}, \quad Y_2 := Y \cap \{x : l_3 < x_n < l_2\},$$
$$S_0 := \{x : l_1 < x_n < l_2 \textit{ and } \operatorname{dist}(Y, x) < R\},$$
$$S_1 := \{x : l_1 < x_n < l_3 \textit{ and } \operatorname{dist}(Y_1, x) < \tilde{R}\},$$
$$S_2 := \{x : l_3 < x_n < l_2 \textit{ and } \operatorname{dist}(Y_2, x) < \tilde{R}\}.$$

Then Y_1 *and* Y_2 *are non-empty,*

$$\operatorname{diam} Y_j \le \frac{N-1}{N} \operatorname{diam} Y \quad \textit{for } j = 1, 2 \tag{7.7}$$

and

$$\frac{1}{\tilde{R}^{n-1}} \left(\int_{S_1} e_\varepsilon + \int_{S_2} e_\varepsilon \right) \le \left(1 + \frac{1}{M}\right)^{n-1} \left\{ \frac{1}{R^{n-1}} \int_{S_0} e_\varepsilon + \varrho(3+R) \right\}. \tag{7.8}$$

Proof For any $x \in Y$, after a parallel translation, assume without loss of generality that $x = 0$ for the proof of (A). Let $\zeta_1(y)$ be a smooth approximation of the characteristic function $\chi_{B_r(0)}$, where $a \le r < R$. Let $\zeta_2(y)$ be a smooth approximation to the characteristic function of S which depends only on y_n. Let us denote

$$h_\varepsilon := \Delta\varphi - \frac{W'(\varphi)}{\varepsilon^2}. \tag{7.9}$$

Multiply (7.9) by $(y \cdot \nabla\varphi)\zeta_1\zeta_2$. After integration by parts twice (as in the computation for (6.42)) and letting $\zeta_1 \to \chi_{B_r(0)}$, we obtain

$$\frac{d}{dr}\left\{\frac{1}{r^{n-1}}\int_{B_r} e_\varepsilon\zeta_2\right\} + \frac{1}{r^n}\int_{B_r}(\xi_\varepsilon + \varepsilon h_\varepsilon(y\cdot\nabla\varphi))\zeta_2 - \frac{\varepsilon}{r^{n+1}}\int_{\partial B_r}(y\cdot\nabla\varphi)^2\zeta_2 \\ - \frac{1}{r^n}\int_{B_r}\{e_\varepsilon y_n - \varepsilon\varphi_{x_n}(y\cdot\nabla\varphi)\}\zeta_2' = 0. \tag{7.10}$$

We estimate the integral over $[r, R]$ ($r \geq a$) of the second term in (7.10) first. We let $\zeta_2 \to \chi_S$ and compute

$$\int_r^R \frac{d\tau}{\tau^n}\int_{B_\tau\cap S}(\xi_\varepsilon + \varepsilon h_\varepsilon(y\cdot\nabla\varphi)) \leq \int_r^R \frac{d\tau}{\tau^n}\left(\int_{B_\tau}(\xi_\varepsilon)_+\right) \\ + \int_r^R \frac{d\tau}{\tau^{n-1}}\left(\int_{B_\tau}\varepsilon|h_\varepsilon||\nabla\varphi|\right) \leq (1+R)\varrho \tag{7.11}$$

where (7.2) and (7.3) are used. From (7.10), (7.11) and (7.1), we obtain (7.5), proving (A). Next, choose $\tilde{y}, \tilde{z} \in Y$ such that $\tilde{z}_n - \tilde{y}_n \geq \frac{\mathrm{diam Y}}{N}$ and $Y \cap \{x : \tilde{y}_n < x_n < \tilde{z}_n\} = \emptyset$. Let $\tilde{l}_1 = \tilde{y}_n + \frac{\tilde{z}_n - \tilde{y}_n}{3}$ and $\tilde{l}_2 = \tilde{z}_n - \frac{\tilde{z}_n - \tilde{y}_n}{3}$. To choose an appropriate $l_3 \in (\tilde{l}_1, \tilde{l}_2)$ which satisfies (7.6), we first observe, for $x \in Y$ and $y \in B_r(x)$,

$$\begin{aligned} I &:= |e_\varepsilon(y_n - x_n) - \varepsilon\varphi_{x_n}(y - x)\cdot\nabla\varphi| \\ &= |(-\xi_\varepsilon)(y_n - x_n) + \varepsilon|\nabla\varphi|^2((y_n - x_n) - \nu_n(y - x)\cdot\nu)| \\ &\leq |\xi_\varepsilon| r + \varepsilon|\nabla\varphi|^2 r\left(1 - (\nu_n)^2 + \sqrt{1 - (\nu_n)^2}\right). \end{aligned} \tag{7.12}$$

Thus by Fubini's theorem (7.12), (7.2) and (7.4) we obtain

$$\int_{\tilde{l}_1}^{\tilde{l}_2} dl \int_a^{\tilde{R}} \frac{d\tau}{\tau^n}\int_{B_\tau(x)\cap\{y_n = l\}} I\, d\mathcal{H}^{n-1} \\ = \int_a^{\tilde{R}} \frac{d\tau}{\tau^n}\int_{B_\tau(x)\cap\{\tilde{l}_1 < y_n < \tilde{l}_2\}} I\, dy \leq \tilde{R}\left(\varrho + E_0^{\frac{1}{2}}\varrho^{\frac{1}{2}}\right). \tag{7.13}$$

The inequality (7.13) is satisfied for each $x \in Y$, hence we guarantee that there exists $l_3 \in (\tilde{l}_1, \tilde{l}_2)$ such that

$$\int_a^{\tilde{R}} \frac{d\tau}{\tau^n}\int_{B_\tau(x)\cap\{y_n = l_3\}} I\, d\mathcal{H}^{n-1}(y) \leq \frac{(N+1)\tilde{R}\left(\varrho + E_0^{\frac{1}{2}}\varrho^{\frac{1}{2}}\right)}{\tilde{l}_2 - \tilde{l}_1}$$

for each $x \in Y$. Since $\tilde{l}_2 - \tilde{l}_1 \geq \frac{\mathrm{diam Y}}{3N}$, we have $\frac{\tilde{R}}{\tilde{l}_2 - \tilde{l}_1} \leq 3MN$, and we obtain (B). We have $S_1 \cup S_2 \subset B_{(\tilde{R}+\mathrm{diam Y})}(x) \cap S$ for $x \in Y$ and $S_1 \cap S_2 = \emptyset$. Thus, using also (3) and (7.5) with $r = \tilde{R} + \mathrm{diam Y} < \mathrm{R}$, we have

$$\begin{aligned} \frac{1}{\tilde{R}^{n-1}} \left(\int_{S_1} e_\varepsilon + \int_{S_2} e_\varepsilon \right) &\leq \frac{1}{\tilde{R}^{n-1}} \int_{B_{(\tilde{R}+\mathrm{diam Y})}(x) \cap S} e_\varepsilon \\ &\leq \left(1 + \frac{1}{M}\right)^{n-1} \left\{ \frac{1}{R^{n-1}} \int_{B_R(x) \cap S} e_\varepsilon + \varrho(3+R) \right\}. \end{aligned} \tag{7.14}$$

Since $B_R(x) \cap S \subset S_0$, we obtain (7.8). One can check that $\tilde{z}_n - \tilde{y}_n \geq \frac{\mathrm{diam Y}}{N}$ implies (7.7). This proves (C). □

Proposition 7.1 *Corresponding to* $0 < R < \infty$, $0 < E_0 < \infty$, $0 < s < 1$ *and* $N \in \mathbb{N}$, *there exists* $0 < \varrho < 1$ *with the following property: Assume* $Y \subset \mathbb{R}^n$ *has no more than* $N+1$ *elements and* $Y \subset \{(0, \ldots, 0, x_n) : x_n \in \mathbb{R}\}$. *For some* $0 < a < R$ *and for all* $y, z \in Y$ *with* $y \neq z$, *we have* $|y - z| > 3a$ *and* $\mathrm{diam Y} \leq \varrho \mathrm{R}$. *In addition we assume (4), (6), (7), (8) of Lemma 7.1. Then we have*

$$\sum_{x \in Y} \frac{1}{a^{n-1}} \int_{B_a(x)} e_\varepsilon \leq s + \frac{1+s}{R^{n-1}} \int_{\{x : \mathrm{dist(Y,x)} < \mathrm{R}\}} e_\varepsilon. \tag{7.15}$$

Proof Denote the number of elements in Y by $\#Y$. If $\#Y = 1$, the proof leading to the conclusion (A) of Lemma 7.1 (with $l_1 = -\infty$ and $l_2 = +\infty$) gives (7.15) if $\varrho(1+R) < s$. Note that M is irrelevant in this case since $\mathrm{diam Y} = 0$. If $1 < \#Y \leq N+1$, we use Lemma 7.1 inductively. First, we choose $M > 1$ depending only on s, n, N so that

$$\left(1 + \frac{1}{M}\right)^{(n-1)N} < 1 + s \quad \text{and} \quad \frac{N-1}{N} < \frac{M}{M+1}. \tag{7.16}$$

Suppose $(M+1)\mathrm{diam Y} < \mathrm{R}$. Then all the assumptions of Lemma 7.1 are satisfied, and we obtain Y_1 and Y_2 with the estimates. We apply Lemma 7.1 again to both Y_1 and Y_2 with R there replaced by $\tilde{R} = M\mathrm{diam Y}$. Due to (7.7) and (7.16), we have the assumption (3) satisfied:

$$(M+1)\mathrm{diam Y_j} < \mathrm{M diam Y}$$

for $j = 1, 2$. We have (7.1) with the right-hand side given by the right-hand side of (7.6). For each $j = 1, 2$, if $\#Y_j = 1$, then we obtain (7.5) with $r = a$. Otherwise, we separate Y_j into two non-empty sets. Each time, all the assumptions of Lemma 7.1 are satisfied. Thus, after $(\#Y - 1)$-times, we separate $\mathbb{R}^n$ into $\#Y$ disjoint horizontal stacks, each having one element of Y. With (7.16), (7.8) and (7.5), we may choose a sufficiently small ϱ depending only on s, n, N, R, E_0 so that (7.15) holds. □

7.2 The ε-scale estimate

Next proposition is almost identical to [27] and [45]. It shows that the energy behaves more or less like a 1-D simple ODE solution if certain quantities are controlled.

Proposition 7.2 *Given* $0 < s,\ b,\ \beta < 1$, *and* $1 < c < \infty$, *there exist* $0 < \varrho,\ \epsilon_6 < 1$ *and* $1 < L < \infty$ *(which also depend on n and W) with the following property:*
Assume $0 < \varepsilon < \epsilon_6$, $\varphi \in C^2(B_{4\varepsilon L})$ *and*

$$\sup_{B_{4\varepsilon L}} \varepsilon|\nabla\varphi| \le c, \quad \sup_{x,y\in B_{4\varepsilon L}} \varepsilon^{\frac{3}{2}} \frac{|\nabla\varphi(x) - \nabla\varphi(y)|}{|x-y|^{\frac{1}{2}}} \le c, \ |\varphi(0)| < 1-b, \tag{7.17}$$

$$\int_{B_{4\varepsilon L}} (|\xi_\varepsilon| + (1-(\nu_n)^2)\varepsilon|\nabla\varphi|^2)\, dx \le \varrho(4\varepsilon L)^{n-1} \tag{7.18}$$

and

$$\sup_{B_{4\varepsilon L}} (\xi_\varepsilon)_+ \le \varepsilon^{-\beta}, \tag{7.19}$$

where ν *and* ξ_ε *are as in* (7.2) *and* (4.16)*. Then for* $J := B_{3\varepsilon L} \cap \{x = (0,\ldots,0,x_n)\}$,

$$\inf_{x\in J} \partial_{x_n}\varphi(x) > 0, \quad (\text{or } \sup_{x\in J} \partial_{x_n}\varphi(x) < 0), \text{ and } [-1+b, 1-b] \subset \varphi(J). \tag{7.20}$$

We also have

$$\left|\sigma - \frac{1}{\omega_{n-1}(L\varepsilon)^{n-1}} \int_{B_{\varepsilon L}} e_\varepsilon \right| \le s. \tag{7.21}$$

Proof Rescale the domain by $x \mapsto \frac{x}{\varepsilon}$. The rescaled function defined on B_{4L} is denoted by $\tilde\varphi$. Let $\Psi : \mathbb{R} \to (-1,1)$ be the unique solution of the ODE

$$\begin{cases} \Psi'(t) = \sqrt{2W(\Psi(t))} \text{ for } t \in \mathbb{R}, \\ \Psi(0) = \tilde\varphi(0). \end{cases} \tag{7.22}$$

We have

$$\int_{\mathbb{R}} \frac{1}{2}|\Psi'(t)|^2\, dt = \int_{\mathbb{R}} \sqrt{\frac{W(\Psi(t))}{2}}\Psi'(t)\, dt = \int_{-1}^{1} \sqrt{\frac{W(s)}{2}}\, ds = \frac{\sigma}{2}. \tag{7.23}$$

Define $\hat\Psi(x) = \hat\Psi(x_1, x_2, \ldots, x_n) := \Psi(x_n)$ for $x \in \mathbb{R}^n$. Using (7.23), it is not difficult to check that $\lim_{L\to\infty} \frac{1}{\omega_{n-1}L^{n-1}} \int_{B_L} \left(\frac{|\nabla\hat\Psi|^2}{2} + W(\hat\Psi)\right) = \sigma$. Thus depending only on n, s, b, W, we may choose a sufficiently large $L > 0$ such that

$$\left|\sigma - \frac{1}{\omega_{n-1}L^{n-1}}\int_{B_L}\left(\frac{|\nabla\hat{\Psi}|^2}{2} + W(\hat{\Psi})\right)\right| \le \frac{s}{2} \tag{7.24}$$

whenever $|\hat{\Psi}(0)| = |\tilde{\varphi}(0)| \le 1 - b$. After fixing such L, we next observe that, for a constant $\tilde{c} = \tilde{c}(W)$,

$$\frac{|\nabla\tilde{\varphi}|^2}{2} - \tilde{c}(1 \pm \tilde{\varphi})^2 \le \frac{|\nabla\tilde{\varphi}|^2}{2} - W(\tilde{\varphi}) = \varepsilon(\xi_\varepsilon)_+ \le \varepsilon^{1-\beta} \quad \text{on } B_{4L} \tag{7.25}$$

by (7.19). Some simple ODE argument combined with (7.25) shows that there exist $0 < \tilde{b} < b$ and $0 < \epsilon_6 < 1$ depending only on b, β, L, W such that, whenever $|\tilde{\varphi}(0)| \le 1 - b$ and $\varepsilon < \epsilon_6$, we have $|\tilde{\varphi}| \le 1 - \tilde{b}$ on B_{4L}.

Next, we define $z : B_{4L} \to \mathbb{R}$ by $z(x) = \Psi^{-1}(\tilde{\varphi}(x))$, where Ψ^{-1} is the inverse function of Ψ. By $\Psi' > 0$ and $|\tilde{\varphi}| \le 1 - \tilde{b}$, Ψ^{-1} and z are well-defined and

$$\Psi'(z(x)) \ge \min_{|\tilde{\varphi}| \le 1-\tilde{b}} \sqrt{2W(\tilde{\varphi})} \tag{7.26}$$

for $x \in B_{4L}$. By (7.17), we have $\|\tilde{\varphi}\|_{C^{1,\frac{1}{2}}(B_{4L})} \le 2c$. Since $\|\Psi^{-1}\|_{C^2(\{t\,:\,|t|\le 1-\tilde{b}\})}$ is bounded depending only on b, β, L, W due to (7.26), we have

$$\|z\|_{C^{1,\frac{1}{2}}(B_{4L})} \le C(b, \beta, L, W, c). \tag{7.27}$$

We next note that $\tilde{\varphi} = \Psi \circ z$ and (7.22) give

$$\begin{aligned} \frac{|\nabla\tilde{\varphi}|^2}{2} - W(\tilde{\varphi}) &= \frac{1}{2}(\Psi'(z))^2(|\nabla z|^2 - 1), \\ |\nabla\tilde{\varphi}|^2(1 - (\nu_n)^2) &= (\Psi'(z))^2(|\nabla z|^2 - (\partial_{x_n} z)^2). \end{aligned} \tag{7.28}$$

After rescaling (7.18) and using (7.26) and (7.28), we obtain

$$\int_{B_{4L}} (||\nabla z|^2 - 1| + |\nabla z|^2 - (\partial_{x_n} z)^2)\, dx \le \max_{|t|\le 1-\tilde{b}} W(t)^{-1} \varrho (4L)^{n-1}. \tag{7.29}$$

For a non-negative function $f \in C^{\frac{1}{2}}(B_{4L})$, suppose $\max_{\bar{B}_{3L}} f = f(\hat{x}) > 0$ for $\hat{x} \in \bar{B}_{3L}$. Then it is easy to check that $f(x) \ge f(\hat{x})/2$ as long as $|x - \hat{x}| \le (f(\hat{x}))^2/(2\|f\|_{C^{\frac{1}{2}}(B_{4L})})^2 =: r$. Then we have

$$f(\hat{x}) \le \frac{1}{\omega_n r^n}\int_{B_r(\hat{x})} 2f\, dx \le \frac{2\cdot 4^n \|f\|^{2n}_{C^{\frac{1}{2}}(B_{4L})}}{\omega_n (f(\hat{x}))^{2n}} \int_{B_{4L}} f\, dx$$

and thus we obtain

$$(\max_{\bar{B}_{3L}} f)^{2n+1} \le 2 \cdot 4^n \|f\|^{2n}_{C^{\frac{1}{2}}(B_{4L})} \int_{B_{4L}} f\,dx. \tag{7.30}$$

By (7.27), (7.29) and (7.30), we have

$$\max_{\bar{B}_{3L}} (||\nabla z|^2 - 1| + |\nabla z|^2 - (\partial_{x_n} z)^2) \le C(b, \beta, L, W, c)\varrho^{\frac{1}{2n+1}}. \tag{7.31}$$

Since $\Psi(0) = \tilde{\varphi}(0) = \Psi(z(0))$, we have $z(0) = 0$. Note that (7.31) for sufficiently small ϱ shows that $\nabla z \approx (0, \ldots, 0, \pm 1)$ uniformly on B_{3L}. This shows that $z \approx x_n$ or $-x_n$ in $C^1(B_{3L})$ when ϱ is small, and in particular, we have (7.20). For the former case, we have $\tilde{\varphi}(x) = \Psi(z(x)) \approx \Psi(x_n) = \hat{\Psi}(x)$, and (7.24) gives (7.21) for sufficiently small ϱ with the right dependence. In the case of $-x_n$, we simply note that changing $\hat{\Psi}$ to $\Psi(-x_n)$ does not affect the proof. □

7.3 Estimate on $\{|\varphi_\varepsilon| \ge 1 - b\}$

We need to show some uniform smallness of energy on $\{|\varphi_\varepsilon| \ge 1 - b\}$ for the final step of this section.

Lemma 7.2 *Suppose φ_ε and u_ε are the solutions for (4.2) constructed in Sect. 5. Given $0 < \delta < T$, there exist c_{21} and ϵ_7 depending only on n, c_1, W with the following property. Suppose for $(x_0, t_0) \in \Omega \times (\delta, T)$ and $0 < \lambda \le 2/3$,*

$$\varphi_\varepsilon(x_0, t_0) < 1 - \varepsilon^\lambda \quad (\mathit{or}\ \ \varphi_\varepsilon(x_0, t_0) > -1 + \varepsilon^\lambda), \tag{7.32}$$

where λ additionally satisfies

$$1 \le \tilde{r} := c_{21}\lambda |\log \varepsilon| \le \varepsilon^{-1} \min\{\sqrt{\delta/2}, 1/2\}. \tag{7.33}$$

Then

$$\inf_{B_{\varepsilon\tilde{r}}(x_0)\times(t_0-\varepsilon^2\tilde{r}^2, t_0)} \varphi_\varepsilon < \alpha \quad \left(\mathit{resp.} \sup_{B_{\varepsilon\tilde{r}}(x_0)\times(t_0-\varepsilon^2\tilde{r}^2, t_0)} \varphi_\varepsilon > -\alpha\right)$$

if $\varepsilon \in (0, \epsilon_7)$.

Proof First note that $B_{\varepsilon\tilde{r}}(x_0) \times (t_0 - \varepsilon^2\tilde{r}^2, t_0) \subset \Omega \times (0, T)$ due to (7.33). Rescale the domain by $x \mapsto \frac{x - x_0}{\varepsilon}$ and $t \mapsto \frac{t - t_0}{\varepsilon^2}$, so that we are concerned with the domain $B_{\tilde{r}} \times (-\tilde{r}^2, 0)$. Let $\tilde{\varphi}_\varepsilon(x, t) := \varphi_\varepsilon(\varepsilon x + x_0, \varepsilon^2 t + t_0)$ and $\tilde{u}_\varepsilon(x, t) := u_\varepsilon(\varepsilon x + x_0, \varepsilon^2 t + t_0)$. As a comparison function, we need a function ψ with the following property

$$\begin{cases} \partial_t \psi = \Delta \psi - \frac{\kappa}{2}\psi & \text{on } \mathbb{R}^n \times (-\infty, 0), \\ \psi(x, t) \ge e^{\frac{|x| + |t|}{c_{21}}} & \text{on } \mathbb{R}^n \times (-\infty, 0) \setminus B_1^{n+1}(0, 0), \\ \psi(0, 0) = 1, & \end{cases} \tag{7.34}$$

for some $c_{21} > 0$. To find such a function, solve $\Delta\tilde{\psi} = \kappa\tilde{\psi}/4$ with $\tilde{\psi}(0) = 1$ on $\mathbb{R}^n$ among radially symmetric functions. One can show that $\tilde{\psi}$ grows exponentially as $|x| \to \infty$ and $\tilde{\psi}$ achieves its minimum at the origin, thus $\tilde{\psi} \geq 1$ on $\mathbb{R}^n$ in particular. Then set $\psi(x,t) := e^{-\kappa t/4}\tilde{\psi}(x)$. With a suitably large c_{21} depending only on n and κ, this ψ satisfies (7.34). Next set $\tilde{r} := c_{21}\lambda|\log\varepsilon|$. We choose such $\tilde{r}$ so that

$$1 - \varepsilon^{\lambda} e^{\frac{\tilde{r}}{c_{21}}} = 0. \tag{7.35}$$

Under the assumption of (7.32) which is equivalent to

$$\tilde{\varphi}_\varepsilon(0,0) < 1 - \varepsilon^{\lambda}, \tag{7.36}$$

for a contradiction, assume

$$\inf_{B_{\tilde{r}}\times(-\tilde{r}^2,0)} \tilde{\varphi}_\varepsilon \geq \alpha. \tag{7.37}$$

Define $\phi_\varepsilon := 1 - \varepsilon^{\lambda}\psi$. By (7.34) we have $\partial_t\phi_\varepsilon = \Delta\phi_\varepsilon + \frac{\kappa}{2}(1-\phi_\varepsilon)$ on $\mathbb{R}^n \times (-\infty, 0)$. Furthermore, on the parabolic boundary of $B_{\tilde{r}} \times (-\tilde{r}^2, 0)$, $\psi \geq e^{\frac{\tilde{r}}{c_{21}}}$ by $\tilde{r} \geq 1$ and (7.34), hence

$$\phi_\varepsilon \leq 1 - \varepsilon^{\lambda} e^{\frac{\tilde{r}}{c_{21}}} = 0 < \alpha \leq \tilde{\varphi}_\varepsilon \tag{7.38}$$

where (7.35) and (7.37) are used. On the other hand $\phi_\varepsilon(0,0) = 1 - \varepsilon^{\lambda}\psi(0,0) = 1 - \varepsilon^{\lambda} > \tilde{\varphi}_\varepsilon(0,0)$ by (7.34) and (7.36). Hence a positive maximum value of $\phi_\varepsilon - \tilde{\varphi}_\varepsilon$ is achieved at a parabolic interior point $(x', t') \in B_{\tilde{r}} \times (-\tilde{r}^2, 0]$. We have $\partial_t(\phi_\varepsilon - \tilde{\varphi}_\varepsilon) - \Delta(\phi_\varepsilon - \tilde{\varphi}_\varepsilon) \geq 0$ at (x', t') and $\phi_\varepsilon(x', t') > \tilde{\varphi}_\varepsilon(x', t')$. The latter inequality combined with (7.37) and (3.3) implies $W'(\tilde{\varphi}_\varepsilon) < W'(\phi_\varepsilon)$. By substituting the equations satisfied by ϕ_ε and $\tilde{\varphi}_\varepsilon$ into the former inequality, we obtain

$$\begin{aligned} 0 &\leq \frac{\kappa}{2}(1-\phi_\varepsilon) + \varepsilon\tilde{u}_\varepsilon \cdot \nabla\tilde{\varphi}_\varepsilon + W'(\tilde{\varphi}_\varepsilon) < \frac{\kappa}{2}(1-\phi_\varepsilon) + \varepsilon^{\frac{3}{4}}\|\nabla\tilde{\varphi}_\varepsilon\|_{L^\infty} + W'(\phi_\varepsilon) \\ &\leq -\frac{\kappa}{2}(1-\phi_\varepsilon) + \varepsilon^{\frac{3}{4}}\|\nabla\tilde{\varphi}_\varepsilon\|_{L^\infty} \leq -\frac{\kappa}{2}\varepsilon^{\lambda} + \varepsilon^{\frac{3}{4}}\|\nabla\tilde{\varphi}_\varepsilon\|_{L^\infty}, \end{aligned}$$

where $W'(\phi_\varepsilon) \leq -\kappa(1-\phi_\varepsilon)$ follows from (7.37) and (3.3) and $|\tilde{u}_\varepsilon| \leq \varepsilon^{-\beta} = \varepsilon^{-\frac{1}{4}}$ by (5.7) and (5.5). We also used $\psi \geq \tilde{\psi} \geq 1$ in the last inequality. Since $\|\nabla\tilde{\varphi}_\varepsilon\|_{L^\infty}$ is bounded uniformly in ε (see Lemma 4.1) and $\lambda \leq 2/3 < 3/4$, for sufficiently small ε, this is a contradiction. The other case may be proved similarly. □

Lemma 7.3 *Under the assumptions of Lemma 7.2, there exist c_{22} and ϵ_8 with the following property. For $t_0 \in (\delta, T)$ and $0 < r < 1/2$ define*

$$Z_{r,t_0} := \left\{ x \in \Omega : \inf_{B_r(x)\times(t_0-r^2,t_0)} |\varphi_\varepsilon| < \alpha \right\}. \tag{7.39}$$

If $0 < \varepsilon < \epsilon_8$, *then*

$$\mathcal{L}^n(Z_{r,t_0}) \le c_{22} r. \tag{7.40}$$

Proof For $x_0 \in Z_{r,t_0}$, we claim that there exist positive constants c_{23} and c_{24} such that

$$\mu^{\varepsilon}_{t_0-2r^2}(B_{c_{23}r}(x_0)) \ge c_{24} r^{n-1}. \tag{7.41}$$

Once (7.41) is proved, the Besicovitch covering theorem and (4.40) prove (7.40) with an appropriate choice of c_{22}. To prove (7.41), for each $x_0 \in Z_{r,t_0}$ we have $(x', t') \in B_r(x_0) \times (t_0 - r^2, t_0)$ such that $|\varphi_\varepsilon(x', t')| < \alpha$. Just as in the proof of Lemma 4.5, we have

$$3c_{24} \le \int_\Omega \tilde\rho_{(x',t'+\varepsilon^2)}(x, t')\, d\mu^{\varepsilon}_{t'}(x). \tag{7.42}$$

By (4.90) with t_1 and t_0 there replaced by t' and $t_0 - 2r^2$, and restricting r and ε appropriately depending on constants appearing in the right-hand side of (4.90), we obtain

$$\int_\Omega \tilde\rho_{(x',t'+\varepsilon^2)}(x, t)\, d\mu^{\varepsilon}_{t}(x)\Big|_{t=t_0-2r^2}^{t'} \le c_{24}. \tag{7.43}$$

The inequalities (7.42) and (7.43) show that

$$2c_{24} \le \int_\Omega \tilde\rho_{(x',t'+\varepsilon^2)}(x, t_0 - 2r^2)\, d\mu^{\varepsilon}_{t_0-2r^2}(x). \tag{7.44}$$

Using the estimate (4.13), we may choose a large $c_{23} > 1$ depending only on D_1 so that

$$\int_{\Omega\setminus B_{c_{23}r}(x')} \tilde\rho_{(x',t'+\varepsilon^2)}(x, t_0 - 2r^2)\, d\mu^{\varepsilon}_{t_0-2r^2}(x) \le c_{24}. \tag{7.45}$$

By (7.44) and (7.45) we obtain

$$c_{24} \le \int_{B_{c_{23}r}(x')} \tilde\rho_{(x',t'+\varepsilon^2)}(x, t_0 - 2r^2)\, d\mu^{\varepsilon}_{t_0-2r^2}(x). \tag{7.46}$$

Since $\tilde\rho_{(x',t'+\varepsilon^2)}(x, t_0 - 2r^2) \le r^{1-n}$ and $B_{c_{23}r}(x') \subset B_{(c_{23}+1)r}(x_0)$, by setting $c_{23}+1$ to be again c_{23}, we obtain (7.41). We restricted r to be small, but when r does not satisfy the restriction, we may choose c_{22} large so that (7.40) holds trivially. □

Proposition 7.3 *Suppose φ_ε and u_ε are the solutions for (4.2) constructed in Sect. 5. Given $0 < \delta < T$ and $0 < s < 1$, there exist $0 < b < 1$ and $0 < \epsilon_9 < 1$ such that*

$$\int_{\{x\in\Omega\,:\,|\varphi_\varepsilon(x,t)|\geq 1-b\}} \frac{W(\varphi_\varepsilon(x,t))}{\varepsilon}\,dx \leq s \tag{7.47}$$

for all $t \in (\delta, T)$ if $0 < \varepsilon \leq \epsilon_9$.

Proof We restrict $0 < b$ to be small in the following independent of ε. Assume that

$$1-\sqrt{b} > \alpha, \quad c_{21}|\log b| \geq 1. \tag{7.48}$$

Choose $J = J(\varepsilon, b) \in \mathbb{N}$ such that

$$\varepsilon^{\frac{1}{2^{J+1}}} \in (b, \sqrt{b}]. \tag{7.49}$$

Restrict ε so that $\varepsilon \leq \min\{\epsilon_7, \epsilon_8\}$ and $c_{21}|\log\varepsilon| \leq \varepsilon^{-1}\min\{\sqrt{\delta/2}, 1/2\}$. Note that, with this choice of b and J, we have by (7.49) and (7.48) that

$$c_{21}\frac{1}{2^J}|\log\varepsilon| \geq c_{21}|\log b| \geq 1. \tag{7.50}$$

Fix $t_0 \in (\delta, T)$ and we define

$$A_j := \left\{x \in \Omega \,:\, 1-\varepsilon^{\frac{1}{2^{j+1}}} \leq |\varphi_\varepsilon(x,t_0)| \leq 1-\varepsilon^{\frac{1}{2^j}}\right\} \quad \text{for } j = 1, \ldots, J. \tag{7.51}$$

For any point $x_0 \in A_j$, we apply Lemma 7.2 with $\lambda = \frac{1}{2^j}$. Note that the condition (7.33) is satisfied due to (7.50). Thus setting $\tilde{r} := c_{21}|\log\varepsilon|/2^j$, we obtain

$$\inf_{B_{\varepsilon\tilde{r}}(x_0)\times(t_0-\varepsilon^2\tilde{r}^2,t_0)} |\varphi_\varepsilon| < \alpha. \tag{7.52}$$

With the notation of (7.39), (7.52) shows

$$A_j \subset Z_{c_{21}\varepsilon|\log\varepsilon|/2^j,t_0} \tag{7.53}$$

and the application of Lemma 7.3 to (7.53) shows

$$\mathcal{L}^n(A_j) \leq c_{22}c_{21}2^{-j}\varepsilon|\log\varepsilon| \tag{7.54}$$

for all $j = 1, \ldots, J$. On A_j, by $|\varphi_\varepsilon| \geq 1-\varepsilon^{\frac{1}{2^{j+1}}}$, we have

$$\frac{W(\varphi_\varepsilon)}{\varepsilon} \leq \left(\max_{[-1,1]}|W''|\right)\cdot\varepsilon^{-1}\frac{\left(\varepsilon^{\frac{1}{2^{j+1}}}\right)^2}{2} \leq c(W)\varepsilon^{2^{-j}-1}. \tag{7.55}$$

Set $Y := \{x \in \Omega \,:\, 1-b \le |\varphi_\varepsilon(x,t_0)| \le 1-\sqrt{\varepsilon}\}$. By (7.51) and (7.49), we have

$$Y \subset \cup_{j=1}^{J} A_j. \tag{7.56}$$

Combining (7.54)–(7.56) and setting $c_{25} := c(W)c_{22}c_{21}$,

$$\begin{aligned}\int_Y \frac{W(\varphi_\varepsilon)}{\varepsilon} &\le \sum_{j=1}^{J}\int_{A_j}\frac{W(\varphi_\varepsilon)}{\varepsilon} \le c_{25}|\log\varepsilon|\sum_{j=1}^{J}2^{-j}\varepsilon^{2^{-j}}\\ &\le c_{25}|\log\varepsilon|\int_1^{J+1}2^{-t}\varepsilon^{2^{-t}}\,dt = c_{25}\frac{\varepsilon^{\frac{1}{2^{J+1}}}-\sqrt{\varepsilon}}{\log 2} \le \frac{c_{25}\sqrt{b}}{\log 2}\end{aligned} \tag{7.57}$$

where we used the fact that $2^{-x}\varepsilon^{2^{-x}}$ is monotone increasing for $x \in [1, J+1]$ as long as $\log\sqrt{b} \le -1$, and (7.49). We restrict b so that the right-hand side of (7.57) is less than $s/2$. The similar estimate shows

$$\int_{\{1-\sqrt{\varepsilon}\le|\varphi_\varepsilon|\le 1-\varepsilon^{\frac{2}{3}}\}}\frac{W(\varphi_\varepsilon)}{\varepsilon} \le c_{25}\varepsilon|\log\varepsilon|. \tag{7.58}$$

Recalling that $|\varphi_\varepsilon| \le 1$, we have

$$\int_{\{1-\varepsilon^{\frac{2}{3}}\le|\varphi_\varepsilon|\}}\frac{W(\varphi_\varepsilon)}{\varepsilon} \le c(W)(\varepsilon^{\frac{2}{3}})^2\cdot\frac{1}{\varepsilon} \le c(W)\varepsilon^{\frac{1}{3}}. \tag{7.59}$$

By (7.57)-(7.59) we restrict ε depending on s so that we have (7.47). □

7.4 Proof of integrality

Finally we prove the integrality of μ_t.

Theorem 7.1 *For a.e. $t > 0$, $\mu_t = \theta\mathcal{H}^{n-1}\lfloor_{M_t}$, where M_t is countably $(n-1)$-rectifiable and $\theta(x,t) = N(x,t)\sigma$ for some $\mathcal{H}^{n-1}$ measurable integer-valued function, μ_t a.e. $x \in \Omega$.*

Proof By the argument in the proof of Proposition 6.1, for a.e. $t \ge 0$, we may choose a subsequence $\{V_t^{\varepsilon_{i_j}}\}_{j=1}^{\infty}$ such that (6.47) and (with the notation of (7.9))

$$c_h(t) := \sup_j \int_\Omega \varepsilon_{i_j}|h_{\varepsilon_{i_j}}\nabla\varphi_{\varepsilon_{i_j}}|(x,t)\,dx < \infty \tag{7.60}$$

hold while $V_t^{\varepsilon_{i_j}} \to V_t$. Here V_t is the rectifiable varifold uniquely determined by μ_t and recall that $\mu_t = \|V_t\|$. In the following we fix any such t and show the claim of the theorem for μ_t. All functions are evaluated at the same t, and we do not write out the time variable (except for μ_t and V_t with or without ε_i) for simplicity. Moreover, though it is important to note that we are discussing a particular subsequence (or its further subsequence), we denote ε_{i_j} by ε_i for simplicity.

For any $m \in \mathbb{N}$, we define

$$A_{i,m} := \Big\{x \in \Omega \ : \ \int_{B_r(x)} \varepsilon_i |h_{\varepsilon_i} \nabla \varphi_{\varepsilon_i}|\, dx \le m \mu_t^{\varepsilon_i}(B_r(x)) \ \text{ for all } 0 < r < 1/2\Big\}. \tag{7.61}$$

The Besicovitch covering theorem with (7.60) and (7.61) shows that

$$\mu_t^{\varepsilon_i}(\Omega \backslash A_{i,m}) \le \frac{c(n)c_h(t)}{m}. \tag{7.62}$$

We then set

$$A_m := \{x \in \Omega \ : \ \text{there exist } x_i \in A_{i,m} \text{ for infinitely many } i \text{ with } x_i \to x\} \tag{7.63}$$

and

$$A := \cup_{m=1}^{\infty} A_m. \tag{7.64}$$

We claim

$$\mu_t(\Omega \backslash A) = 0. \tag{7.65}$$

Otherwise, we would have a compact set $K \subset \Omega \backslash A$ such that $\mu_t(K) \ge \frac{1}{2}\mu_t(\Omega \backslash A)$. For any $m \in \mathbb{N}$ we have $K \subset \Omega \backslash A_m$ by (7.64). For each point $x \in K$, by (7.63), there exists a neighborhood of x which does not intersect with $A_{i,m}$ for all sufficiently large i. Due to the compactness, thus, there exist i_0 and an open set O_m such that $K \subset O_m$ and $O_m \cap A_{i,m} = \emptyset$ for all $i \ge i_0$. Let $\phi_m \in C_c(O_m; \mathbb{R}^+)$ such that $0 \le \phi_m \le 1$ and $\phi_m = 1$ on K. Then

$$\begin{aligned} \mu_t(K) \le \int_\Omega \phi_m \, d\mu_t &= \lim_{i\to\infty} \int_\Omega \phi_m \, d\mu_t^{\varepsilon_i} = \lim_{i\to\infty} \int_{\Omega \backslash A_{j,m}} \phi_m \, d\mu_t^{\varepsilon_i} \\ &\le \liminf_{i\to\infty} \mu_t^{\varepsilon_i}(\Omega \backslash A_{j,m}) \end{aligned} \tag{7.66}$$

for all $j \ge i_0$. Since the last quantity of (7.66) is less than $c(n)c_h(t)/m$ by (7.62), and since m is arbitrary, we obtain $\mu(K) = 0$. This proves the claim (7.65).

Since μ_t is rectifiable, μ_t a.e. point x has an approximate tangent space. By (7.65), we may also assume that for μ_t a.e. x there exists some $m \in \mathbb{N}$ such that $x \in A_m$. We fix any such point, and after a parallel translation, we may assume that $x = 0$. Furthermore, after a rotation, we may assume that the approximate tangent space is $P := \{x_n = 0\}$. Denote $\theta := \lim_{r\downarrow 0} \frac{\|V_t\|(B_r(x))}{\omega_{n-1} r^{n-1}}$. We will be done if we prove that $\sigma^{-1}\theta \in \mathbb{N}$.

For any sequence $r_i \downarrow 0$, we have $\lim_{i\to\infty} (\Phi_{r_i})_\# V_t = \theta |P|$, where $\Phi_{r_i}(x) = \frac{x}{r_i}$ and $(\Phi_{r_i})_\#$ is the usual push-forward of varifold. $|P|$ is the unit density varifold naturally derived from P. Since $0 \in A_m$, there exists a subsequence (denoted by the same index)

$x_i \in A_{i,m}$ such that $\lim_{i\to\infty} x_i = 0$. After choosing a further subsequence, we may assume that

$$\lim_{i\to\infty} (\Phi_{r_i})_\# V_t^{\varepsilon_i} = \theta |P|, \tag{7.67}$$

$$\lim_{i\to\infty} \frac{x_i}{r_i} = 0 \tag{7.68}$$

and

$$\lim_{i\to\infty} \frac{\varepsilon_i^{\beta\prime-\beta} |\log \varepsilon_i|}{r_i^{n-1}} = 0. \tag{7.69}$$

For a such choice, we also have $\lim_{i\to\infty} \frac{\varepsilon_i}{r_i} = 0$. Rescale the coordinates by $\tilde{x} := \frac{x}{r_i}$ and define $\tilde{\varepsilon}_i := \frac{\varepsilon_i}{r_i} \to 0$. Define $\tilde{\varphi}_{\tilde{\varepsilon}_i}(\tilde{x}) := \varphi_{\varepsilon_i}(r_i \tilde{x})$. We also define $\tilde{\xi}_{\tilde{\varepsilon}_i}$ and $\tilde{h}_{\tilde{\varepsilon}_i}$ as in (4.15) and (7.9) corresponding to $\tilde{\varepsilon}_i$ and $\tilde{\varphi}_{\tilde{\varepsilon}_i}$. Due to (6.47), we may choose a further subsequence so that

$$\lim_{i\to\infty} \int_{B_3} |\tilde{\xi}_{\tilde{\varepsilon}_i}|\, d\tilde{x} = 0. \tag{7.70}$$

Due to Corollary 4.1 and (7.69), for any $y \in B_2$ and $0 < r < 2$, we have

$$\begin{aligned}\int_0^r \frac{d\tilde{\tau}}{\tilde{\tau}^n} \int_{B_{\tilde{\tau}}(y)} (\tilde{\xi}_{\tilde{\varepsilon}_i})_+\, d\tilde{x} &= \frac{1}{r_i^{n-1}} \int_0^{rr_i} \frac{d\tau}{\tau^n} \int_{B_\tau(r_i y)} (\xi_{\varepsilon_i})_+\, dx \\ &\le \frac{2c_{10}\varepsilon_i^{\beta\prime-\beta} |\log \varepsilon_i|}{r_i^{n-1}} \to 0 \end{aligned}\tag{7.71}$$

as $i \to \infty$. For $\tilde{h}_{\tilde{\varepsilon}_i}$, we have

$$\begin{aligned}\tilde{\varepsilon}_i \int_{B_3} |\tilde{h}_{\tilde{\varepsilon}_i} \nabla \tilde{\varphi}_{\tilde{\varepsilon}_i}|\, d\tilde{x} &= \frac{\varepsilon_i}{r_i^{n-2}} \int_{B_{3r_i}} |h_{\varepsilon_i} \nabla \varphi_{\varepsilon_i}|\, dx \le \frac{m}{r_i^{n-2}} \mu_t^{\varepsilon_i}(B_{4r_i}(x_i)) \\ &\le m 4^{n-1} \omega_{n-1} D_1 r_i \to 0 \end{aligned}\tag{7.72}$$

as $i \to \infty$, where we used (7.68), $x_i \in A_{i,m}$, (7.61) and (4.13). If one defines a varifold $\tilde{V}_t^{\tilde{\varepsilon}_i}$ corresponding to $\tilde{\varphi}_{\tilde{\varepsilon}_i}$ as in (6.41), then one can check that $\tilde{V}_t^{\tilde{\varepsilon}_i} = (\Phi_{r_i})_\# V_t^{\varepsilon_i}$. Next we claim

$$\int_{B_3} (1 - (\nu_n)^2) \tilde{\varepsilon}_i |\nabla \tilde{\varphi}_{\tilde{\varepsilon}_i}|^2\, d\tilde{x} \to 0 \tag{7.73}$$

as $i \to \infty$, where $\nu = (\nu_1, \ldots, \nu_n) = \frac{\nabla \tilde{\varphi}_{\tilde{\varepsilon}_i}}{|\nabla \tilde{\varphi}_{\tilde{\varepsilon}_i}|}$. Note first that $G_{n-1}(\mathbb{R}^n) \cong \mathbb{S}^{n-1}/\{\pm 1\}$ and a function defined by $\psi : \pm\nu \in \mathbb{S}^{n-1}/\{\pm 1\} \longmapsto 1 - \nu_n^2$ is continuous. Thus for any $\phi \in C_c(\mathbb{R}^n)$, we have by (7.67)

$$\tilde{V}_t^{\tilde{\varepsilon}_i}(\phi\psi) = \int \phi(\tilde{x})(1-(\nu_n)^2)\,d\|\tilde{V}_t^{\tilde{\varepsilon}_i}\|(\tilde{x}) \to \theta|P|(\phi\psi) \tag{7.74}$$

and since $P = \{x_n = 0\}$,

$$\theta|P|(\phi\psi) = \theta\int_P \phi(\tilde{x})\psi((0,\cdot,0,\pm 1))\,d\mathcal{H}^{n-1}(\tilde{x}) = 0. \tag{7.75}$$

In particular, (7.74) and (7.75) prove (7.73). In the following we fix this subsequence and drop the tilde for simplicity.

Assume that N is the smallest positive integer greater than $\sigma^{-1}\theta$, that is,

$$\theta \in [(N-1)\sigma, N\sigma). \tag{7.76}$$

Let $s > 0$ be arbitrary. By Proposition 7.3 and (7.70), there exists $0 < b < 1$ such that

$$\int_{B_3\cap\{|\varphi_{\varepsilon_i}|\geq 1-b\}} \left(\frac{\varepsilon_i|\nabla\varphi_{\varepsilon_i}|^2}{2} + \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i}\right) \leq s \tag{7.77}$$

for all sufficiently large i. Corresponding to s and b as well as c given by Lemma 4.1, by Proposition 7.2, we choose ϱ and L (with a restriction on ε_i). Then with $R = 2$, by Proposition 7.1, we restrict ϱ further if necessary. We use Proposition 7.1 with $a = L\varepsilon_i$. For all large i we define

$$\begin{aligned} G_i \ := B_2 \cap \{|\varphi_{\varepsilon_i}| \leq 1-b\} \cap \Big\{x \ : \int_{B_r(x)} \varepsilon_i|h_{\varepsilon_i}\nabla\varphi_{\varepsilon_i}| + |\xi_{\varepsilon_i}| \\ +(1-(\nu_n)^2)\varepsilon_i|\nabla\varphi_{\varepsilon_i}|^2 \leq \varrho\,\mu_t^{\varepsilon_i}(B_r(x)) \text{ if } \varepsilon_i L \leq r \leq 1\Big\}. \end{aligned} \tag{7.78}$$

By the Besicovitch covering theorem, we obtain

$$\begin{aligned} \mu_t^{\varepsilon_i}(B_2\cap\{|\varphi_{\varepsilon_i}| \ \leq 1-b\}\backslash G_i) \leq \frac{c(n)}{\varrho}\int_{B_3} \varepsilon_i|h_{\varepsilon_i}\nabla\varphi_{\varepsilon_i}| + |\xi_{\varepsilon_i}| \\ +(1-(\nu_n)^2)\varepsilon_i|\nabla\varphi_{\varepsilon_i}|^2. \end{aligned} \tag{7.79}$$

The right hand side goes to 0 as $i \to \infty$ by (7.72), (7.70), (7.73). Next we claim the following lower bound for all sufficiently large i:

$$\mu_t^{\varepsilon_i}(B_r(x)) \geq (\sigma - 2s)\omega_{n-1}r^{n-1} \tag{7.80}$$

for all $L\varepsilon_i \leq r \leq 1$ and $x \in G_i$. To see this, first note that the assumptions of Proposition 7.2 are all satisfied due to Lemma 4.1, (7.78) and (4.26). This proves the inequality (7.80) with $r = L\varepsilon_i$ and with $2s$ replaced by s. Next the identity (7.10) with $\zeta_2 \equiv 1$, (7.11), (7.71) and (7.78) shows

$$\begin{aligned}\frac{1}{\tau^{n-1}}\mu_t^{\varepsilon_i}(B_\tau(x))\Big|_{\tau=L\varepsilon_i}^{r} &\geq o(1) - \int_{L\varepsilon_i}^{r} \varrho \frac{\mu_t^{\varepsilon_i}(B_\tau(x))}{\tau^{n-1}}\,d\tau \\ &\geq o(1) - \omega_{n-1} D_1 \varrho \end{aligned} \tag{7.81}$$

after integrating over $[L\varepsilon_i, r]$. We may restrict ϱ so that $D_1\varrho < s$. Thus (7.81) gives (7.80) for all sufficiently large i. Since $\mu_t^{\varepsilon_i} = \|V_t^{\varepsilon_i}\| \to \theta\mathcal{H}^{n-1}\lfloor_P$, (7.80) shows that points in G_i converge uniformly to P as $i \to \infty$.

For any $x \in P \cap B_1$ and $|l| \leq 1 - b$, we next prove

$$\#(P^{-1}(x) \cap G_i \cap \{\varphi_{\varepsilon_i} = l\}) \leq N - 1. \tag{7.82}$$

If the claim were not true, we choose N elements and set it to be Y, and apply Proposition 7.1 with $R = 1$, $\varphi = \varphi_{\varepsilon_i}$ and $a = L\varepsilon_i$. The property $|y - z| > 3L\varepsilon_i$ holds due to (7.20), $\mathrm{diam}\, Y \leq \varrho$ due to the uniform convergence of G_i to P, (6), (7) are due respectively to (7.78) and (7.71). Thus all the assumptions of Proposition 7.1 are satisfied and we have

$$\sum_{y\in Y} \frac{1}{(L\varepsilon_i)^{n-1}}\mu_t^{\varepsilon_i}(B_{L\varepsilon_i}(y)) \leq s + (1+s)\mu_t^{\varepsilon_i}(\{z \,:\, \mathrm{dist}\,(Y, z) < 1\}) \tag{7.83}$$

for all sufficiently large i. Since $\lim_{i\to\infty} \mu_t^{\varepsilon_i}(\{z \,:\, \mathrm{dist}\,(Y, z) < 1\}) = \theta\omega_{n-1}, \#Y = N$ and (7.80), we obtain

$$N(\sigma - 2s)\omega_{n-1} \leq s + (1+s)\theta\omega_{n-1}. \tag{7.84}$$

Since $\sigma N > \theta$ by definition, (7.84) is a contradiction for sufficiently small s depending only on σ, θ and n. Thus we proved (7.82).

To conclude the proof, we consider push-forward of

$$\hat{V}_t^{\varepsilon_i} := V_t^{\varepsilon_i}\lfloor_{\{|x_n|\leq 1\}\times \mathbf{G}(n,n-1)}$$

by P, $P_\# \hat{V}_t^{\varepsilon_i}$. For any $\phi(x, S) \in C_c((P \cap B_2) \times \mathbf{G}(n, n-1))$, we have (for all sufficiently large i)

$$P_\# \hat{V}_t^{\varepsilon_i}(\phi) = \int_{\{|x_n|\leq 1\}} \phi(P(x), P)|\Lambda_{n-1} P \circ (I - \nu \otimes \nu)|\, d\mu_t^{\varepsilon_i}. \tag{7.85}$$

Here $\Lambda_{n-1}A$ denotes the Jacobian of $A \in \mathrm{Hom}(\mathbb{R}^n; \mathbb{R}^n)$ ([1]). One can check that $|\Lambda_{n-1} P \circ (I - \nu\otimes\nu)| = |\nu_n| = \frac{|\partial_{x_n}\varphi_{\varepsilon_i}|}{|\nabla\varphi_{\varepsilon_i}|}$. Due to the varifold convergence (7.67), we have $P_\# \hat{V}_t^{\varepsilon_i} \to P_\#(\theta|P|) = \theta|P|$ as $i \to \infty$. In the following we also use

$$\lim_{i\to\infty} \int_{B_3} \left| \frac{\varepsilon_i|\nabla\varphi_{\varepsilon_i}|^2}{2} + \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i} - |\nabla\varphi_{\varepsilon_i}|\sqrt{2W(\varphi_{\varepsilon_i})}\right| dx = 0 \tag{7.86}$$

which follows from (7.70). Now we have

$$\begin{aligned}\omega_{n-1}\theta &= \|\theta|P|\|(B_1) = \lim_{i\to\infty} \|P_\# \hat{V}_t^{\varepsilon_i}\|(B_1) = \lim_{i\to\infty}\int_{B_1} |\nu_n|\, d\mu_t^{\varepsilon_i} \\ &\le \liminf_{i\to\infty} \int_{B_1\cap\{|\varphi_{\varepsilon_i}|\le 1-b\}\cap G_i} |\nu_n|\, d\mu_t^{\varepsilon_i} + 2s \\ &\le \liminf_{i\to\infty} \int_{B_1\cap\{|\varphi_{\varepsilon_i}|\le 1-b\}\cap G_i} |\nu_n||\nabla\varphi_{\varepsilon_i}|\sqrt{2W(\varphi_{\varepsilon_i})}\, dx + 2s \qquad (7.87)\end{aligned}$$

due to (7.77), (7.79) and (7.86). By the co-area formula [41, 10.6], we obtain

$$\begin{aligned}&\int_{B_1\cap\{|\varphi_{\varepsilon_i}|\le 1-b\}\cap G_i} |\nu_n||\nabla\varphi_{\varepsilon_i}|\sqrt{2W(\varphi_{\varepsilon_i})}\, dx \\ &\quad = \int_{-1+b}^{1-b} d\tau \int_{\{\varphi_{\varepsilon_i}=\tau\}\cap B_1\cap G_i} |\nu_n|\sqrt{2W(\tau)}\, d\mathcal{H}^{n-1}. \qquad (7.88)\end{aligned}$$

Then by the area formula [41, 12.4] applied to the map $P : \{\varphi_{\varepsilon_i} = \tau\} \to \{x_n = 0\}$, we have

$$\begin{aligned}&\int_{\{\varphi_{\varepsilon_i}=\tau\}\cap B_1\cap G_i} |\nu_n|\, d\mathcal{H}^{n-1} \\ &\quad = \int_{\{x_n=0\}} \mathcal{H}^0(\{\varphi_{\varepsilon_i} = \tau\}\cap B_1\cap G_i\cap P^{-1}(x))\, d\mathcal{H}^{n-1}(x). \qquad (7.89)\end{aligned}$$

Now the integrand of the right-hand side of (7.89) is $\le N-1$ due to (7.82) for $|x| \le 1$, and 0 otherwise. Combining (7.87)–(7.89), we finally obtain

$$\begin{aligned}\omega_{n-1}\theta &\le 2s + \liminf_{i\to\infty} \omega_{n-1}(N-1)\int_{-1+b}^{1-b}\sqrt{2W(\tau)}\, d\tau \\ &\le 2s + \omega_{n-1}(N-1)\sigma. \qquad (7.90)\end{aligned}$$

Since $s > 0$ is arbitrary, (7.90) shows $\theta \le (N-1)\sigma$. By (7.76), we have $\theta = (N-1)\sigma$. □

8 Proof of the main theorem

We finally define a family of varifolds which will be a generalized solution of (1.2). To remove the multiple of σ, we re-define V_t as follows.

Definition 8.1 For a.e. $t \ge 0$ when μ_t is rectifiable and integral modulo division by σ, let V_t be the uniquely defined integral varifold by $\sigma^{-1}\mu_t$. For any other $t > 0$, define V_t by $V_t(\phi) := \sigma^{-1}\int_U \phi(x, P_0)\, d\mu_t(x)$ for $\phi \in C_c(G_{n-1}(U))$, where $P_0 \in G(n, n-1)$ is an arbitrary fixed element.

With this definition, we have $\|V_t\| = \sigma^{-1}\mu_t$ for all $t \ge 0$, and $V_t \in \mathbf{IV}_{n-1}(\Omega)$ for a.e. $t \ge 0$ by Theorem 7.1. Thus (a) of Definition 2.1 is satisfied. The condition (b) is satisfied due to (4.13). Let us consider (c). The L^2 integrability of u, $\int_0^T \int_\Omega |u|^2 \, d\|V_t\| dt < \infty$, may be proved as in (4.42) and (4.43) once (b) is established. For h, we prove the following.

Proposition 8.1 *For a.e. $t \ge 0$, V_t has a generalized mean curvature $h(V_t)$ and we have*

$$\int_\Omega \phi |h(V_t)|^2 \, d\|V_t\| \le \sigma^{-1} \liminf_{i\to\infty} \int_\Omega \varepsilon_i \phi \left(\Delta\varphi_{\varepsilon_i} - \frac{W'(\varphi_{\varepsilon_i})}{\varepsilon_i^2} \right)^2 dx < \infty \tag{8.1}$$

for any $\phi \in C_c(\Omega\,;\, \mathbb{R}^+)$.

Proof Just as in the proof of Proposition 6.1, for a.e. $t \ge 0$, we may assume (6.47) and (6.48) and there exists a subsequence $\{V_t^{\varepsilon_{i_j}}\}_{j=1}^\infty$ converging to σV_t (note that we re-defined V_t) with (6.46). By arguing as in the proof of Proposition 6.1, for any $g \in C_c^1(\Omega\,;\, \mathbb{R}^n)$, we have

$$|\delta V_t(g)| \le \sigma^{-1} \left(\int_\Omega |g|^2 \, d\mu_t \right)^{1/2} \liminf_{j\to\infty} \left(\int_\Omega \varepsilon_{i_j} \left(\Delta\varphi_{\varepsilon_{i_j}} - \frac{W'}{\varepsilon_{i_j}^2} \right)^2 dx \right)^{1/2}. \tag{8.2}$$

The inequality and (6.46) show that the total variation $\|\delta V_t\|$ of δV_t is absolutely continuous with respect to $\mu_t = \sigma \|V_t\|$. Thus by the Radon-Nikodym theorem there exists a $\|V_t\|$ measurable vector field $h(V_t)$ (generalized mean curvature vector) such that

$$\delta V_t(g) = -\int_\Omega g \cdot h(V_t) \, d\|V_t\|. \tag{8.3}$$

Since V_t is rectifiable, going back to the definition of countably $(n-1)$-rectifiable set, one can show that $C_c^1(\Omega)$ is dense in $L^2(\|V_t\|)$. Then a standard approximation argument shows $h(V_t) \in L^2(\|V_t\|)$ and (8.1) with $\phi = 1$. Next, given $\phi \in C_c(\Omega\,;\, \mathbb{R}^+)$, let $\psi_j \in C_c^1(\Omega\,;\, \mathbb{R}^+)$ be a sequence such that $\lim_{k\to\infty} \|\phi - \psi_k\|_{C^0(\Omega)} = 0$. Using $\psi_k g$ in the proof of Proposition 6.1 and letting $k \to \infty$, we obtain

$$\left| \int_\Omega \phi g \cdot h(V_t) \, d\mu_t \right| \le \left(\int_\Omega \phi |g|^2 \, d\mu_t \right)^{1/2} \liminf_{j\to\infty} \left(\int_\Omega \varepsilon_{i_j} \phi \left(\Delta\varphi_{\varepsilon_{i_j}} - \frac{W'}{\varepsilon_{i_j}^2} \right)^2 dx \right)^{1/2}. \tag{8.4}$$

By approximation, we obtain (8.1) from (8.4). □

Now Proposition 8.1 combined with Lemma 4.4 and Fatou's lemma proves (c). For the proof of (d), one point which we need to be careful about is that we may not have the whole sequence $\{V_t^{\varepsilon_i}\}_{i=1}^{\infty}$ converging to V_t as varifold for a.e. $t \geq 0$ even though $\{\|V_t^{\varepsilon_i}\|\}_{i=1}^{\infty}$ converges to $\sigma\|V_t\| = \mu_t$ for all $t \geq 0$.

Proposition 8.2 *The family of varifolds $\{V_t\}_{t\geq 0}$ defined in Definition 8.1 is a generalized solution of (1.2).*

Proof We prove (2.10) for $\phi \in C_c^2(\Omega \times [0,\infty)\,;\,\mathbb{R}^+)$. For $\phi \in C_c^1$, one can approximate ϕ by a sequence of C_c^2 functions and obtain the same result in the limit. First by modifying (5.11) we obtain (with the notation (7.9))

$$\mu_t^{\varepsilon_i}(\phi(\cdot,t))\Big|_{t=t_1}^{t_2} = \int_{t_1}^{t_2}\int_{\Omega} -\varepsilon_i\phi h_{\varepsilon_i}^2 - \varepsilon_i h_{\varepsilon_i}\nabla\phi\cdot\nabla\varphi_{\varepsilon_i} + \varepsilon_i\phi h_{\varepsilon_i}u_{\varepsilon_i}\cdot\nabla\varphi_{\varepsilon_i} + \varepsilon_i(\nabla\varphi_{\varepsilon_i}\cdot\nabla\phi)(u_{\varepsilon_i}\cdot\nabla\varphi_{\varepsilon_i})\,dxdt + \int_{t_1}^{t_2}\int_{\Omega}\frac{\partial\phi}{\partial t}\,d\mu_t^{\varepsilon_i}\,dt. \tag{8.5}$$

Modulo division by σ, the left-hand side of (8.5) converges to that of (2.10) due to Proposition 5.13. The same is true for the last term of (8.5). Hence we focus on the middle 4 terms. First we approximate u_{ε_i} by a fixed smooth $\tilde{u}$ as follows. Given $\epsilon > 0$, we choose a large j so that $t_2 < T_j$ and

$$\|u - u_{\varepsilon_j}\|_{L^q([0,T_j];W^{1,p}(\Omega))} < \epsilon \ \text{ and } \ \|u_{\varepsilon_j} - u_{\varepsilon_i}\|_{L^q([0,T_j];W^{1,p}(\Omega))} < \epsilon \tag{8.6}$$

for all $i \geq j$. This is possible since u_{ε_i} converges to u in this norm. Set $\tilde{u} := u_{\varepsilon_j}$. Then we have

$$\left|\int_{t_1}^{t_2}\int_{\Omega}\varepsilon_i\phi h_{\varepsilon_i}(u_{\varepsilon_i}-\tilde{u})\cdot\nabla\varphi_{\varepsilon_i} + \varepsilon_i(\nabla\varphi_{\varepsilon_i}\cdot\nabla\phi)((u_{\varepsilon_i}-\tilde{u})\cdot\nabla\varphi_{\varepsilon_i})\right| \leq \left(\int_{t_1}^{t_2}\int_{\Omega}2\varepsilon_i\left(\phi^2h_{\varepsilon_i}^2 + |\nabla\phi|^2|\nabla\varphi_{\varepsilon_i}|^2\right)\right)^{1/2}\left(\int_{t_1}^{t_2}\int_{\Omega}|u_{\varepsilon_i}-\tilde{u}|^2\,d\mu_t^{\varepsilon_i}\,dt\right)^{1/2}. \tag{8.7}$$

As in the proof of Lemma 4.4, and by (4.13) and (8.6), we have

$$\int_{t_1}^{t_2}dt\int_{\Omega}|u_{\varepsilon_i}-\tilde{u}|^2\,d\mu_t^{\varepsilon_i} \leq c(n)D_1(t_2-t_1)^{1-\frac{2}{q}}\|u_{\varepsilon_i} - \tilde{u}\|^2_{L^q([t_1,t_2];W^{1,p}(\Omega))} < c\epsilon^2. \tag{8.8}$$

By (8.7) and (8.8), replacing u_{ε_i} by $\tilde{u}$ in (8.5) produces error of $c\epsilon^2$. Similarly we have

$$\left|\int_{t_1}^{t_2}\int_{\Omega}(-h\phi+\nabla\phi)\cdot((u-\tilde{u})\cdot\nu)\nu\,d\mu_t dt\right| \leq c'\epsilon. \tag{8.9}$$

Thus we will finish the proof if we prove

$$\liminf_{i\to\infty}\int_{t_1}^{t_2}\int_\Omega -\varepsilon_i\phi h_{\varepsilon_i}^2-\varepsilon_i h_{\varepsilon_i}\nabla\phi\cdot\nabla\varphi_{\varepsilon_i}+\varepsilon_i\phi h_{\varepsilon_i}\tilde u\cdot\nabla\varphi_{\varepsilon_i}$$
$$+\varepsilon_i(\nabla\varphi_{\varepsilon_i}\cdot\nabla\phi)(\tilde u\cdot\nabla\varphi_{\varepsilon_i})\,dxdt\le\int_{t_1}^{t_2}\mathcal{B}(\mu_t,\tilde u(\cdot,t),\phi(\cdot,t))\,dt,\quad (8.10)$$

where we denote

$$\mathcal{B}(\mu_t,\tilde u(\cdot,t),\phi(\cdot,t)):=\int_\Omega(\nabla\phi-h\phi)\cdot(h+(\tilde u\cdot\nu)\nu)\,d\mu_t.$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\hat a_i(t)\ &:=\varepsilon_i\int_\Omega-\phi h_{\varepsilon_i}^2-h_{\varepsilon_i}\nabla\phi\cdot\nabla\varphi_{\varepsilon_i}+\phi h_{\varepsilon_i}\tilde u\cdot\nabla\varphi_{\varepsilon_i}+(\nabla\varphi_{\varepsilon_i}\cdot\nabla\phi)(\tilde u\cdot\nabla\varphi_{\varepsilon_i})\\ &\le\int_\Omega\frac{\varepsilon_i}{2}|\nabla\varphi_{\varepsilon_i}|^2\left(\frac{|\nabla\phi|^2}{\phi}+\phi|\tilde u|^2+2|\tilde u||\nabla\phi|\right)\\ &\le\int_\Omega\frac{\varepsilon_i}{2}|\nabla\varphi_{\varepsilon_i}|^2(\hat\phi+\phi|\tilde u|^2+2|\tilde u||\nabla\phi|)=:\hat b_i(t),\end{aligned}\quad (8.11)$$

where $\hat\phi\in C_c(\Omega;\mathbb{R}^+)$ is chosen so that $\frac{|\nabla\phi|^2}{\phi}\le\hat\phi$. This in particular shows $\hat b_i(t)-\hat a_i(t)\ge 0$ for $t_1\le t\le t_2$. Using the general fact that $\liminf_{i\to\infty}(a_i+b_i)\le\limsup_{i\to\infty}a_i+\liminf_{i\to\infty}b_i$ and Fatou's lemma, we have

$$\begin{aligned}\liminf_{i\to\infty}\int_{t_1}^{t_2}\hat a_i(t)\,dt&\le-\liminf_{i\to\infty}\int_{t_1}^{t_2}(\hat b_i(t)-\hat a_i(t))\,dt+\liminf_{i\to\infty}\int_{t_1}^{t_2}\hat b_i(t)\,dt\\ &\le-\int_{t_1}^{t_2}\liminf_{i\to\infty}(\hat b_i(t)-\hat a_i(t))\,dt+\liminf_{i\to\infty}\int_{t_1}^{t_2}\hat b_i(t)\,dt.\end{aligned}\quad (8.12)$$

Since $\hat b_i(t)$ converges to $\frac12\int_\Omega(\hat\phi+\phi|\tilde u|^2+2|\tilde u||\nabla\phi|)\,d\mu_t$ for all $t_1\le t\le t_2$ and bounded uniformly, from (8.12) and the dominated convergence theorem we have

$$\liminf_{i\to\infty}\int_{t_1}^{t_2}\hat a_i(t)\,dt\le-\int_{t_1}^{t_2}\liminf_{i\to\infty}(-\hat a_i(t))\,dt.\quad (8.13)$$

Thus we may finish the proof of (8.10) via (8.13) if we prove

$$-\liminf_{i\to\infty}(-\hat a_i(t))\le\mathcal{B}(\mu_t,\tilde u(\cdot,t),\phi(\cdot,t))\quad (8.14)$$

for a.e. $t \in [t_1, t_2]$. Fix t such that the claim of Proposition 8.1 holds. Let $\{\varepsilon_{i_j}\}_{j=1}^{\infty}$ be a subsequence such that

$$\liminf_{i\to\infty}(-\hat{a}_i(t)) = \lim_{j\to\infty}(-\hat{a}_{i_j}(t)). \tag{8.15}$$

We may choose a further subsequence (denoted by the same index) such that $V_t^{\varepsilon_{i_j}} \to \sigma\tilde{V}_t$ as varifold. By the Cauchy-Schwarz inequality,

$$-\hat{a}_i(t) \geq \int_\Omega \frac{1}{2}\varepsilon_i\phi h_{\varepsilon_i}^2 - \left(\frac{|\nabla\phi|^2}{\phi} + |\tilde{u}|^2 + |\tilde{u}||\nabla\phi|\right)\varepsilon_i|\nabla\varphi_{\varepsilon_i}|^2\,dx \tag{8.16}$$

where the last negative term is bounded uniformly. If $\liminf_{j\to\infty}\int_\Omega \varepsilon_{i_j}\phi h_{\varepsilon_{i_j}}^2\,dx$ is infinity, we have (8.14) with the left-hand side $= -\infty$. Thus we may assume otherwise. At this point, arguing just as in the proof of Proposition 6.1, we may prove that $\tilde{V}_t\lfloor_{\{\phi>0\}}$ is rectifiable and $\tilde{V}_t\lfloor_{\{\phi>0\}} = V_t\lfloor_{\{\phi>0\}}$. Then the argument in the proof of Proposition 8.1 shows (8.1). For the remaining three terms in $\hat{a}_{i_j}(t)$, since $V_t^{\varepsilon_{i_j}}\lfloor_{\{\phi>0\}} \to \sigma V_t\lfloor_{\{\phi>0\}}$ as varifold and by (6.41), we have for any $\tilde{\phi} \in C_c^2(\{\phi > 0\}\,;\,\mathbb{R}^+)$

$$\begin{aligned}
&\lim_{j\to\infty}\varepsilon_{i_j}\int_\Omega h_{\varepsilon_{i_j}}\nabla\tilde{\phi}\cdot\nabla\varphi_{\varepsilon_{i_j}} - \tilde{\phi}h_{\varepsilon_{i_j}}\tilde{u}\cdot\nabla\varphi_{\varepsilon_{i_j}} - (\nabla\varphi_{\varepsilon_{i_j}}\cdot\nabla\tilde{\phi})(\tilde{u}\cdot\nabla\varphi_{\varepsilon_{i_j}})\,dx\\
&= \sigma\delta V_t(\nabla\tilde{\phi} - \tilde{u}\tilde{\phi}) - \int_\Omega (\nabla\tilde{\phi}\cdot\nu)(\tilde{u}\cdot\nu)\,d\mu_t\\
&= \int_\Omega -h\cdot(\nabla\tilde{\phi} - \tilde{u}\tilde{\phi}) - (\nabla\tilde{\phi}\cdot\nu)(\tilde{u}\cdot\nu)\,d\mu_t.
\end{aligned} \tag{8.17}$$

We may construct a sequence of approximation $\{\tilde{\phi}_k\}_{k=1}^{\infty}$ such that $\lim_{k\to\infty}\|\phi - \tilde{\phi}_k\|_{C^2} = 0$, $\phi \geq \tilde{\phi}_k$ and $\operatorname{spt}\tilde{\phi}_k \subset \{\phi > 0\}$. For such approximating sequence,

$$\begin{aligned}
&\left|\int_\Omega \varepsilon_{i_j}h_{\varepsilon_{i_j}}\nabla(\phi - \tilde{\phi}_k)\cdot\nabla\varphi_{\varepsilon_{i_j}}\right|\\
&\leq \left(\int_\Omega \varepsilon_{i_j}h_{\varepsilon_{i_j}}^2\phi\right)^{1/2}\left(\int_\Omega \frac{|\nabla(\phi-\tilde{\phi}_k)|^2}{\phi-\tilde{\phi}_k}\varepsilon_{i_j}|\nabla\varphi_{\varepsilon_{i_j}}|^2\right)^{1/2}\\
&\leq \left(\int_\Omega \varepsilon_{i_j}h_{\varepsilon_{i_j}}^2\phi\right)^{1/2}\left(2\|\phi-\tilde{\phi}_k\|_{C^2}\right)^{1/2}(2\mu_t^{\varepsilon_{i_j}}(\Omega))^{1/2} \to 0
\end{aligned} \tag{8.18}$$

as $k \to \infty$ uniformly in j. The error of replacing $\tilde{\phi} = \tilde{\phi}_k$ in (8.17) by ϕ can be approximated similarly. Thus (8.17) holds also for ϕ instead of $\tilde{\phi}$. Recall that we have taken a subsequence so that (8.15) holds. Combined with (8.1) and (8.17) with $\tilde{\phi} = \phi$, and recalling that $h\cdot\tilde{u} = h\cdot(\tilde{u}\cdot\nu)\nu$ for μ_t a.e. by Brakke's perpendicularity theorem [6, Ch.5], we have proved (8.14). This concludes the proof. □

We next discuss the proof of Theorem 2.2 (2).

Proposition 8.3 *There exists a further subsequence (denoted by the same index) $\{\varphi_{\varepsilon_i}\}_{i=1}^{\infty}$ and a function $\varphi \in BV_{loc}(\Omega \times [0,\infty)) \cap C^{\frac{1}{2}}_{loc}([0,\infty); L^1(\Omega))$ such that for all $t \geq 0$,*

$$w_{\varepsilon_i}(\cdot,t) \to \varphi(\cdot,t) \tag{8.19}$$

strongly in $L^1_{loc}(\Omega)$ and φ satisfies the properties of Theorem 2.2 (2). Here w_{ε_i} is defined by

$$w_{\varepsilon_i} := \Phi \circ \varphi_{\varepsilon_i} \text{ with } \Phi(s) := \sigma^{-1} \int_{-1}^{s} \sqrt{2W(y)}\,dy.$$

Proof Note that $\Phi(1) = 1$ and $\Phi(-1) = 0$. We compute

$$|\nabla w_{\varepsilon_i}| = \sigma^{-1} |\nabla \varphi_{\varepsilon_i}| \sqrt{2W(\varphi_{\varepsilon_i})} \leq \sigma^{-1} \left(\frac{\varepsilon_i |\nabla \varphi_{\varepsilon_i}|^2}{2} + \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i} \right).$$

Fix $T > 0$. For all sufficiently large i, by (4.13) we have

$$\int_\Omega |\nabla w_{\varepsilon_i}(\cdot,t)|\,dx \leq \int_\Omega \sigma^{-1} \left(\frac{\varepsilon_i |\nabla \varphi_{\varepsilon_i}|^2}{2} + \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i} \right) dx \leq \sigma^{-1} D_1 \tag{8.20}$$

for any $t \in [0,T]$. By the similar argument we have

$$\begin{aligned}
\int_0^T \int_\Omega |\partial_t w_{\varepsilon_i}|\,dxdt &\leq \sigma^{-1} \int_0^T \int_\Omega \left(\frac{\varepsilon_i |\partial_t \varphi_{\varepsilon_i}|^2}{2} + \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i} \right) dxdt \\
&\leq \sigma^{-1} \int_0^T \int_\Omega \varepsilon_i \left\{ (u_{\varepsilon_i} \cdot \nabla \varphi_{\varepsilon_i})^2 + \left(\Delta \varphi_{\varepsilon_i} - \frac{W'(\varphi_{\varepsilon_i})}{\varepsilon_i} \right)^2 \right\} dxdt \\
&\quad + \sigma^{-1} \int_0^T \int_\Omega \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i}\,dxdt,
\end{aligned} \tag{8.21}$$

and the last quantity is uniformly bounded due to Lemma 4.4. By (8.20) and (8.21) $\{w_{\varepsilon_i}\}_{i=1}^{\infty}$ is bounded in $BV_{loc}(\Omega \times [0,T])$. By the standard compactness theorem and a diagonal argument, there exists a subsequence (denoted by the same index) $\{w_{\varepsilon_i}\}_{i=1}^{\infty}$ and $w \in BV_{loc}(\Omega \times [0,\infty))$ such that

$$w_{\varepsilon_i} \to w \quad \text{strongly in } L^1_{loc}(\Omega \times [0,\infty)) \tag{8.22}$$

and a.e. pointwise. We set $\varphi := (1 + \Phi^{-1} \circ w)/2$. We have

$$\varphi_{\varepsilon_i} \to 2\varphi - 1 \quad \text{a.e. in } \Omega \times [0,\infty)$$

and by this with $|\varphi_{\varepsilon_i}| \leq 1$ we obtain

$$\varphi_{\varepsilon_i} \to 2\varphi - 1 \quad \text{in } L^1_{loc}(\Omega \times [0,\infty)).$$

Due to the uniform bound on $\int_\Omega \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i}\,dx$, one can prove by Fatou's lemma that $\varphi_{\varepsilon_i} \to \pm 1$ for a.e. (x,t) and hence $\varphi = 1$ or $= 0$ a.e. on $\Omega \times [0,\infty)$. In particular, since $\varphi = 1 \iff w = 1$ and $\varphi = 0 \iff w = 0$, we have $w = \varphi$ on $\Omega \times [0,\infty)$. This in particular proves the $BV_{loc}(\Omega \times [0,\infty))$ property of φ. For a.e. $0 \le t_1 < t_2 \le T$ and any open set $U \subset\subset \Omega$, we have

$$\int_U |\varphi(\cdot,t_1) - \varphi(\cdot,t_2)|\,dx = \lim_{i\to\infty}\int_U |w_{\varepsilon_i}(\cdot,t_1) - w_{\varepsilon_i}(\cdot,t_2)|\,dx$$
$$\le \liminf_{i\to\infty}\int_U\int_{t_1}^{t_2} |\partial_t w_{\varepsilon_i}|\,dtdx$$
$$\le \liminf_{i\to\infty}\sigma^{-1}\int_\Omega\int_{t_1}^{t_2}\left(\frac{\varepsilon_i|\partial_t\varphi_{\varepsilon_i}|^2}{2}\sqrt{t_2-t} + \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i\sqrt{t_2-t}}\right)dtdx.$$

Note that the right-hand side does not depend on U. Thus, by the similar argument to (8.21) we have with $c = c(c_2, n, p, q, D_0, T, W)$

$$\int_\Omega |\varphi(\cdot,t_1) - \varphi(\cdot,t_2)|\,dx \le c\sqrt{t_2 - t_1}. \tag{8.23}$$

Since $(1+\varphi_{\varepsilon_i}(\cdot,0))/2 \to \chi_{\Omega_0}$ by (5.6), we have (2c). We assumed that Ω_0 is a bounded domain, hence, (8.23) shows that $\varphi(\cdot,t) \in L^1(\Omega)$ for a.e. $t \ge 0$. Moreover, we may define $\varphi(\cdot,t)$ as a characteristic function for all $t \ge 0$ so that $\varphi \in C^{\frac{1}{2}}_{loc}([0,\infty); L^1(\Omega))$ due to (8.23). This proves (2a) and $C^{\frac{1}{2}}_{loc}$ property for φ. From (8.22), for a.e. $t \ge 0$, $w_{\varepsilon_i}(\cdot,t) \to \varphi(\cdot,t)$ in $L^1_{loc}(\Omega)$ strongly. Using (8.23), one can show by a simple telescopic argument that the convergence is true for all $t \ge 0$ instead of a.e. t, which proves (8.19). By the standard lower semicontinuity property of BV norm, for any $\phi \in C_c(\Omega; \mathbb{R}^+)$ and $0 \le t < \infty$, we have

$$\int_\Omega \phi\, d\|\nabla\varphi(\cdot,t)\| \le \liminf_{i\to\infty}\int_\Omega \phi|\nabla w_{\varepsilon_i}|\,dx$$
$$\le \lim_{i\to\infty}\sigma^{-1}\int_\Omega\left(\frac{\varepsilon_i|\nabla\varphi_{\varepsilon_i}|^2}{2} + \frac{W(\varphi_{\varepsilon_i})}{\varepsilon_i}\right)\phi\,dx = \int_\Omega \phi\, d\|V_t\|.$$

This proves (2b).

To prove (2d), we consider the a.e. $t \ge 0$ for which we have proved the integrality of V_t. Writing $\|V_t\| = \theta\mathcal{H}^{n-1}\lfloor_{M_t}$, we already know that θ is integer-valued $\|V_t\|$ a.e. and that M_t is countably $(n-1)$-rectifiable. In addition, by (2.8), we have $1 \le \theta \le N(t)$, $\mathcal{H}^{n-1}$ a.e. on M_t for some integer $N(t)$. The latter shows in particular that

$$\mathcal{H}^{n-1}\lfloor_{M_t} \le \|V_t\| \le N(t)\mathcal{H}^{n-1}\lfloor_{M_t}. \tag{8.24}$$

By (2a) and (2b), we know that $\|\nabla\varphi(\cdot,t)\| = \mathcal{H}^{n-1}\lfloor_{\tilde{M}_t}$ for some countably $(n-1)$-rectifiable set by De Giorgi's theorem (see [24, 4.4]). To prove (2.16), assume the contrary. Then by the standard argument (see [41, 3.5]), there would be a point $x \in$

$\tilde{M}_t \setminus M_t$ with $\lim_{r\downarrow 0} \mathcal{H}^{n-1}(B_r(x) \cap \tilde{M}_t)/\omega_{n-1}r^{n-1} = 1$ while $\lim_{r\downarrow 0} \mathcal{H}^{n-1}(B_r(x) \cap M_t)/\omega_{n-1}r^{n-1} = 0$. Then, using also (8.24), one would then have a contradiction to Theorem 2.2 (2b). Thus we have (2.16).

To prove (2.17), we closely follow the proof of integrality again. We already know that for $\|V_t\|$ a.e. x, we have the properties described in the proof of Theorem 7.1. By the well-known property of set of finite perimeter ([24, 3.8]), for $\mathcal{H}^{n-1}$ a.e. $x \in \tilde{M}_t$, the blow-up limit of φ centered at x is supported by a half-space. For $\mathcal{H}^{n-1}$ a.e. $x \in \Omega\setminus\tilde{M}_t$ (in particular on $M_t\setminus\tilde{M}_t$), the blow-up limit centered at x is a constant function with value either 0 or 1. By (8.19), up to $\mathcal{H}^{n-1}$ null set, we may assume in addition to the properties of $\{V_t^{\varepsilon_i}\}_{i=1}^{\infty}$ in the proof of Theorem 7.1 that $\tilde{w}_{\varepsilon_i}(\tilde{x}) := w_{\varepsilon_i}(r_i\tilde{x})$ converges strongly in $L^1_{loc}(\mathbb{R}^n)$ and pointwise $\mathcal{L}^n$ a.e. to $\chi_{\{x_n\geq 0\}}$ (or $\chi_{\{x_n\leq 0\}}$) if $x = 0$ is in $\tilde{M}_t$, or to 1 (or 0) if $x = 0$ is in $M_t\setminus\tilde{M}_t$. Since the proof for other cases is similar, we only discuss the case of $\tilde{M}_t$ and $\lim_{i\to\infty}\tilde{w}_{\varepsilon_i} = \chi_{\{x_n\geq 0\}}$ in the following. In terms of φ_{ε_i} (which is the relabeling of $\tilde{\varphi}_{\varepsilon_i}$), note that this means that φ_{ε_i} converges a.e. to $\chi_{\{x_n\geq 0\}} - \chi_{\{x_n<0\}}$.

As one follows the proof of Theorem 7.1, the difference occurs at (7.76), where we already know that θ is an integer multiple of σ. So let $N-1 := \sigma^{-1}\theta(\geq 1)$. We want to conclude that N is an even integer. We follow the proof until (7.89), and at this point, define for $i \in \mathbb{N}$ (and writing $Y(\tau, x) := \{\varphi_{\varepsilon_i} = \tau\} \cap B_1 \cap G_i \cap P^{-1}(x)$)

$$\begin{aligned}\tilde{A}_i &:= \{x \in B_1^{n-1} \,:\, \forall\tau \in (-1+b, 1-b) \Rightarrow \mathcal{H}^0(Y(\tau,x)) \leq N-2\},\\ A_i &:= \{x \in B_1^{n-1} \,:\, \exists\tau \in (-1+b, 1-b) \Rightarrow \mathcal{H}^0(Y(\tau,x)) = N-1\}.\end{aligned} \tag{8.25}$$

We know from (7.82) that $\mathcal{H}^0(Y(\tau,x))$ has to be $\leq N-1$, thus, $B_1^{n-1} = \tilde{A}_i \cup A_i$ and

$$\mathcal{H}^{n-1}(\tilde{A}_i) = \omega_{n-1} - \mathcal{H}^{n-1}(A_i) \tag{8.26}$$

for all sufficiently large i. In (7.90), we have

$$\begin{aligned}\omega_{n-1}\sigma(N-1) &\leq 2s + \liminf_{i\to\infty}\int_{-1+b}^{1-b}\sqrt{2W(\tau)}\{(N-2)\mathcal{H}^{n-1}(\tilde{A}_i)\\ &\quad +(N-1)\mathcal{H}^{n-1}(A_i)\}\,d\tau \leq 2s + (N-2)\sigma\omega_{n-1} + \sigma\liminf_{i\to\infty}\mathcal{H}^{n-1}(A_i)\end{aligned} \tag{8.27}$$

where we used (8.26). Thus we have from (8.27)

$$\omega_{n-1} - 2\sigma^{-1}s \leq \liminf_{i\to\infty}\mathcal{H}^{n-1}(A_i). \tag{8.28}$$

By (7.20), for all sufficiently large i and any point $x \in A_i$, the image $\varphi_{\varepsilon_i}(B_1 \cap P^{-1}(x))$ covers $[-1+b, 1-b]$ at least $N-1$ times. The each covering is monotone, thus we know that $\varphi_{\varepsilon_i}(y)$ as y moves from $P^{-1}(x) \cap \{x_n = -s\}$ to $P^{-1}(x) \cap \{x_n = s\}$ along $P^{-1}(x)$ has to go up and down between $-1+b$ and $1-b$ at least $N-1$ times. Next, since φ_{ε_i} converges a.e. pointwise to $\chi_{\{x_n\geq 0\}} - \chi_{\{x_n<0\}}$, by Egoroff's Theorem and then Fubini's Theorem, there exists $s_1 \in [s, 2s]$, $s_2 \in [-2s, -s]$, $C_1 \subset B_1^{n-1}$

and $C_2 \subset B_1^{n-1}$ such that φ_{ε_i} converges uniformly to 1 on $C_1 \times \{s_1\}$ and to -1 on $C_2 \times \{s_2\}$ while

$$\mathcal{H}^{n-1}(C_i) \geq \omega_{n-1} - s \quad \text{for } i = 1, 2. \tag{8.29}$$

Set $C_3 = C_1 \cap C_2$ so that, by (8.29),

$$\mathcal{H}^{n-1}(C_3) \geq \omega_{n-1} - 2s. \tag{8.30}$$

Now, for a contradiction, assume that N is odd. For $x \in A_i \cap C_3$, consider the image of φ_{ε_i} on $\{(x, x_n) : x_n \in [s_2, s_1]\}$. By the uniform convergence and $x \in C_3$, for sufficiently large i, $\varphi_{\varepsilon_i}(x, s_2) < -1 + b$ and $\varphi_{\varepsilon_i}(x, s_1) > 1 - b$. Since φ_{ε_i} is continuous, image of φ_{ε_i} having at least even $N - 1$ covering of $[-1 + b, 1 - b]$ implies that there has to be at least another covering of $[-1+b, 1-b]$. Thus, for each $\tau \in [-1 + b, 1 - b]$ and $x \in A_i \cap C_3$, we have

$$\mathcal{H}^0(\{x_n \in [s_2, s_1] : \varphi_{\varepsilon_i}(x, x_n) = \tau\}) \geq N. \tag{8.31}$$

Then by the coarea formula and (8.31), we have

$$\begin{aligned}
&\int_{s_2}^{s_1} \sqrt{2W(\varphi_{\varepsilon_i}(x, x_n))}|\partial_{x_n}\varphi_{\varepsilon_i}(x, x_n)|\, dx_n \\
&\quad = \int_{-1}^{1} \sqrt{2W(\tau)}\mathcal{H}^0(\{x_n \in [s_2, s_1] : \varphi_{\varepsilon_i}(x, x_n) = \tau\})\, d\tau \\
&\quad \geq N \int_{-1+b}^{1-b} \sqrt{2W(\tau)}\, d\tau.
\end{aligned} \tag{8.32}$$

Note that by (8.28) and (8.30), we have for sufficiently large i

$$\mathcal{H}^{n-1}(A_i \cap C_3) \geq \omega_{n-1} - (3 + 2\sigma^{-1})s. \tag{8.33}$$

Integrating (8.32) over $A_i \cap C_3$ and (8.33) give

$$\begin{aligned}
\int_{B_1} \sqrt{2W(\varphi_{\varepsilon_i})}|\nabla\varphi_{\varepsilon_i}| &\geq \int_{(A_i \cap C_3)\times[s_2,s_1]} \sqrt{2W(\varphi_{\varepsilon_i})}|\partial_{x_n}\varphi_{\varepsilon_i}| \\
&\geq (\omega_{n-1} - (3 + 2\sigma^{-1})s)N \int_{-1+b}^{1-b} \sqrt{2W(\tau)}\, d\tau.
\end{aligned} \tag{8.34}$$

We may choose b so that $\int_{-1+b}^{1-b} \sqrt{2W(\tau)}\, d\tau \geq \sigma - s$. On the other hand, by (7.67), we have

$$\int_{B_1} \sqrt{2W(\varphi_{\varepsilon_i})}|\nabla\varphi_{\varepsilon_i}|\, dx \leq \int_{B_1} \frac{\varepsilon_i|\nabla\varphi_{\varepsilon_i}|^2}{2} + \frac{W}{\varepsilon_i}\, dx \to \omega_{n-1}(N - 1)\sigma. \tag{8.35}$$

For sufficiently small s depending only on n, N and σ, (8.34) and (8.35) lead to a contradiction. This proves N has to be even. As we mentioned, other cases of φ being constant (either 0 or 1) can be similarly proved. This concludes the proof of (2.17) and (2d). □

We next verify

Proposition 8.4 *The function u satisfies the property of Theorem 2.2 (3).*

Proof Consider the case $p < n$ and fix $T > 0$. Since $\lim_{i\to\infty} \|u_{\varepsilon_i} - u\|_{L^q([0,T];(W^{1,p})^n)} = 0$, $\{u_{\varepsilon_i}\}$ is a Cauchy sequence in this norm. By (2.11) with $s = \frac{p(n-1)}{n-p}$, we have

$$\int_0^T dt \left(\int_\Omega |u_{\varepsilon_i} - u_{\varepsilon_j}|^s \, d\|V_t\|\right)^{\frac{q}{s}} \le c(n, p, q, D_1)\|u_{\varepsilon_i} - u_{\varepsilon_j}\|^q_{L^q([0,T];(W^{1,p}(\Omega))^n)}. \tag{8.36}$$

By a standard argument, we may subtract a subsequence $\{u_{\varepsilon_{i_j}}\}_{j=1}^\infty$ which converges pointwise $\|V_t\| \times dt$ a.e. on $\Omega \times [0, T]$ to an element of $L^q([0, T]; (L^s(\|V_t\|))^n)$. This limit function is uniquely determined by u independent of the approximate sequence and (2.18) holds. For $p = n$, we apply the same argument locally for $p' < n$ which gives (2.18) with any $2 \le s < \infty$. For $p > n$, the standard Sobolev inequality and the Hölder inequality prove the claim immediately. □

To conclude the proof of Theorem 2.2 we prove

Proposition 8.5 *We have $T_1 > 0$ with the property described in Theorem 2.2 (4).*

Proof By integrality, we already know that $\|V_t\| = \theta\mathcal{H}^{n-1}\lfloor_{M_t}$ for a.e. $t \ge 0$, where θ is integer-valued $\mathcal{H}^{n-1}$ a.e. on M_t. Thus we should prove that $\mathcal{H}^{n-1}(\{\theta(\cdot, t) \ge 2\}) = 0$ for a.e. $0 < t < T_1$ for some $T_1 > 0$. We will determine the lower bound of T_1 in the following. Assume there exist $0 < \hat{t} < T_1$ and $\hat{x} \in M_{\hat{t}}$ such that $M_{\hat{t}}$ has the approximate tangent space at $\hat{x}$ and the density $\theta(\hat{x}, \hat{t}) \ge 2$. Then it is not difficult to check that

$$\lim_{r\to 0} \int_\Omega \tilde{\rho}_{(\hat{x},\hat{t}+r^2)} \, d\|V_{\hat{t}}\| = \theta(\hat{x}, \hat{t}) \ge 2. \tag{8.37}$$

Since $\|V_0\| = \mathcal{H}^{n-1}\lfloor_{M_0}$ and M_0 is C^1, we have

$$\int_\Omega \tilde{\rho}_{(x,t)} \, d\|V_0\| \le 3/2 \tag{8.38}$$

for any $(x, t) \in \Omega \times (0, T_1]$, where T_1 depends only on M_0. We then use (4.90) with $\varepsilon \to 0$. We then have

$$\lim_{r\to 0} \int_\Omega \tilde{\rho}_{(\hat{x},\hat{t}+r^2)} \, d\|V_t\| \Big|_{t=0}^{\hat{t}} \le c_{14}c_2^2\hat{t}^{\hat{p}} D_1 + c_3 e^{-\frac{1}{128\hat{t}}} \hat{t} D_1, \tag{8.39}$$

and the right hand side of (8.39) may be made smaller than $1/2$ by restricting T_1. Then we would have a contradiction since the left-hand side is $\geq 1/2$ due to (8.37) and (8.38). This proves the first part of (4). We next prove $\|\nabla\varphi(\cdot,t)\| = \|V_t\|$ a.e. $t \in [0, T_1]$. With the notation of (2d), for a.e. $t \in [0, T_1]$, we have $\|V_t\| = \mathcal{H}^{n-1}\lfloor_{M_t}$ since $\theta = 1$ a.e. from the first part. But then, by (2.17), $\mathcal{H}^{n-1}(M_t\setminus\tilde{M}_t) = 0$ since $\theta = 1$ and odd. Thus combined with (2.16), $\tilde{M}_t = M_t$ modulo null set, and this shows the claim. We may take T_1 to be $\sup\{t > 0 : V_t$ is unit density for a.e.$t \in [0, t]\}$. □

As for the proof of Theorem 2.3, (1) and (3) follow from [30] and [46], respectively, which give criterion for partial $C^{1,\zeta}$ and $C^{2,\alpha}$ regularity. For (1), we check that [30, Sect. 3.1 (A1)–(A4)] are all satisfied. Namely, (A1) asks V_t to be unit density for a.e. t, (A2) is on the uniform density ratio upper bound which follows from (2.8), (A3) is on the integrability of u which is given by (2.18) and (A4) is the flow equation which is (2.10). If $p < n$, the exponent of integrability of u in (2.18) has to satisfy $\zeta := 1 - (n-1)/s - 2/q = 2 - n/p - 2/q > 0$, and this follows from (2.14). If $p \geq n$, we may choose any $s > (n-1)q/(q-2)$ in (2.18) so that we have $0 < \zeta$, and we may take sufficiently large s so that $0 < \zeta < 1 - 2/q$ can be arbitrarily close to $1 - 2/q$. This proves (1). The conclusion for $C^{2,\alpha}$ is precisely the claim of [46]. Thus we only need to prove (2) and (4).

Proposition 8.6 *The family of varifolds $\{V_t\}_{t\geq 0}$ satisfies the property of Theorem 2.3 (2) and (4)*

Proof For a.e. $0 \leq t < T_1$, we have proved that V_t has unit density property, thus we may use results in [30] for $\{V_t\}_{0\leq t<T_1}$. We first claim that there exists $0 < T_3 \leq T_1$ depending only on $D_1, n, p, q, \|u\|_{L^q([0,T_1];(W^{1,p}(\Omega))^n)}$ (D_1 corresponding to T_1) and $c_{26} = c_{26}(D_1, n)$ such that

$$\mathrm{dist}\,(\mathrm{spt}\,\|V_t\|, M_0) \leq c_{26}\sqrt{t} \tag{8.40}$$

for a.e. $0 \leq t \leq T_3$. For the proof, we use [30, Proposition6.2]. Citing the result for the convenience of the reader, we have for $x \in \Omega$ and $0 < r < 1$

$$\int_{B_r(x)} \hat{\rho}_{(x,t+\epsilon)}(\cdot,t)\,d\|V_t\| - \int_{B_r(x)} \hat{\rho}_{(x,t+\epsilon)}(\cdot,0)\,d\|V_0\| \\ \leq c(n,s,q)\|u\|^2_{L^{s,q}} D_1^{1-\frac{2}{s}} t^{\zeta} + c(n)D_1 r^{-2} t, \tag{8.41}$$

where $s := \frac{p(n-1)}{n-p}$ if $p < n$ and any $\frac{(n-1)q}{q-2} < s < \infty$ if $p \geq n$, $\zeta = 1-(n-1)/s-2/q$ and $\|u\|_{L^{s,q}} := (\int_0^t (\int_{B_r(x)} |u|^s\,d\|V_\lambda\|)^{q/s}\,d\lambda)^{1/q}$. $\hat{\rho}_{(x,t+\epsilon)}$ is $\rho_{(x,t+\epsilon)}$ times a radially symmetric cut-off function with support in $B_{14r/15}(x)$ and $= 1$ near x. Note that $\|u\|_{L^{s,q}}$ may be bounded in terms of D_1 and $\|u\|_{L^q([0,T_1];(W^{1,p}(\Omega))^n)}$ as was done for the proof of (2.18). By restricting T_3 small, we may conclude from (8.41) that

$$\int_{B_r(x)} \hat{\rho}_{(x,t+\epsilon)}(\cdot,t)\,d\|V_t\| - \int_{B_r(x)} \hat{\rho}_{(x,t+\epsilon)}(\cdot,0)\,d\|V_0\| \leq \frac{1}{2} + c(n)D_1 r^{-2} t. \tag{8.42}$$

Let c_{26} be a constant to be fixed shortly and assume that there exists $x \in \operatorname{spt}\|V_t\|$ such that $\operatorname{dist}(x, M_0) > c_{26}\sqrt{t}$ and $0 < t \le T_3$. We may assume that V_t is unit density and has approximate tangent space with multiplicity 1 at x, since such time and point are generic. In particular, one can check that $\lim_{\epsilon\to 0+}\int_{B_r(x)} \hat\rho_{(x,t+\epsilon)}(\cdot,t)\,d\|V_t\| = 1$ and (8.42) thus shows

$$\frac{1}{2} - \int_{B_r(x)} \hat\rho_{(x,t)}(\cdot,0)\,d\|V_0\| \le c(n)D_1 r^{-2} t. \tag{8.43}$$

We now choose $r = c_{26}\sqrt{t}/2$. Since $B_r(x)\cap M_0 = \emptyset$, the integral in (8.43) is 0. Hence we obtain $\frac{1}{2} \le 4c(n)D_1 c_{26}^{-2}$. If we choose a sufficiently large c_{26} depending only on n and D_1, we obtain a contradiction. This proves (8.40).

Next, since $\operatorname{spt}\|\nabla\varphi(\cdot,t)\| \subset \operatorname{spt}\|V_t\|$ by Theorem 2.2 (2b), (8.40) shows that $\varphi(\cdot,t)$ is a constant function on each connected component of $\Omega\setminus\{x : \operatorname{dist}(x,M_0) \le c_{26}\sqrt{t}\}$ for a.e. $0 \le t \le T_3$. Since $\varphi(\cdot,t)$ is a characteristic function and is continuous in L^1 norm with respect to time, one sees that

$$\begin{aligned}\varphi(\cdot,t) = 1 \quad &\text{on } \{x \in \Omega_0 : \operatorname{dist}(x,M_0) > c_{26}\sqrt{t}\},\\ \varphi(\cdot,t) = 0 \quad &\text{on } \{x \notin \Omega_0 : \operatorname{dist}(x,M_0) > c_{26}\sqrt{t}\}\end{aligned} \tag{8.44}$$

for all $0 \le t \le T_3$. We now estimate the location of $\operatorname{spt}\|V_t\|$ during the short initial time. Since M_0 is assumed to be C^1, there exists $r_1 > 0$ such that, for each $x \in M_0$ (we may assume that x is the origin and $T_x M_0 = \mathbb{R}^{n-1}\times\{0\}$ after parallel translation and orthogonal rotation), M_0 is locally represented as a C^1 graph $g : B^{n-1}_{r_1} \to \mathbb{R}$ on $B^{n-1}_{r_1}\times(-r_1, r_1)$. We take the coordinate system so that Ω_0 is located on the upper side, above the graph of g. We may also restrict r_1 (uniformly on M_0) so that for all $r \le r_1$, we have

$$\sup_{x\in B^{n-1}_r} |g(x)| \le \frac{r}{10}. \tag{8.45}$$

For $t \in [0, (10c_{26})^{-2}r^2]$, (8.44) and (8.45) show that

$$\begin{aligned}\varphi(\cdot,t) &= 1 \quad \text{on } B^{n-1}_{9r/10}\times[r/5, r_1),\\ \varphi(\cdot,t) &= 0 \quad \text{on } B^{n-1}_{9r/10}\times(-r_1, -r/5].\end{aligned} \tag{8.46}$$

Next we use [30, Theorem 8.7]. Using the notation there, corresponding to $1 \le E_1 < \infty$, $0 < \nu < 1$, p, q with $1-(n-1)/p - 2/q > 0$, there exist 4 constants ($\varepsilon_6, \sigma, \Lambda_3, c_{19}$ in [30]) with the stated properties. Here, we use $E_1 = D_1$, $\nu = 1/2$, $p = s$ above and the same q. The condition $1-(n-1)/p-2/q > 0$ is then satisfied. To avoid confusion in the following, we denote the constants in [30] corresponding to these choices by $\varepsilon_{6,KT}, \sigma_{KT}, \Lambda_{3,KT}, c_{19,KT}$. In the following, let $P \in \mathbf{G}(n, n-1)$ be the projection $\mathbb{R}^n \to \mathbb{R}^{n-1}\times\{0\}$ and $P^\perp$ be its orthogonal complement. We then use [30, Proposition 6.5] with

$$\Lambda = \Lambda_{3,KT}/18 \tag{8.47}$$

to obtain $c_{6,KT}$ with the property that

$$\frac{1}{r^{n+1}}\int_{B_r} |P^{\perp}(x)|^2\, d\|V_t\| \le \exp(1/(4\Lambda))\frac{1}{r^{n+1}}\int_{B_{Lr}} |P^{\perp}(x)|^2\, d\|V_0\| \\ + c_{6,KT}\{(r^{2\zeta}\|u\|^2_{L^{s,q}} + r^{\zeta}\|u\|_{L^{s,q}})L^2 + L^{n+1}\exp(-(L-1)^2/(8\Lambda))\} \tag{8.48}$$

for all $t \in [0, \Lambda r^2]$ provided $2 \le L < \infty$ and $rL \le 1$. Here $c_{6,KT}$ depends only on $s, q, D_1, \Lambda_{3,KT}$ but not on L. Given $1 > \varepsilon > 0$, we may choose $L \ge 2$ so that

$$c_{6,KT} L^{n+1} \exp(-(L-1)^2/(8\Lambda)) < \varepsilon \tag{8.49}$$

and then choose $r_2 \le L^{-1}$ uniformly on M_0 so that (using M_0 is C^1)

$$\exp(1/(4\Lambda)) \sup_{0<r\le r_2} \frac{1}{r^{n+1}}\int_{B_{Lr}} |P^{\perp}(x)|^2\, d\|V_0\| < \varepsilon, \\ c_{6,KT}\left(r_2^{2\zeta}\|u\|^2_{L^{s,q}} + r_2^{\zeta}\|u\|_{L^{s,q}}\right) L^2 < \varepsilon. \tag{8.50}$$

The inequalities (8.48)–(8.50) gives for $r \le r_2$ and $t \in [0, \Lambda r^2]$

$$\frac{1}{r^{n+1}}\int_{B_r} |P^{\perp}(x)|^2\, d\|V_t\| \le 3\varepsilon. \tag{8.51}$$

We next use [30, Proposition 6.4] on $B_r \times [0, \Lambda r^2]$ with a slight modification. Instead of obtaining result on the time interval $[R^2/5, \Lambda]$ as in [30], we modify the proof so that we obtain the similar estimate on the time interval $[(10c_{26})^{-2}r^2, \Lambda r^2]$. This is achieved by a simple replacement of the cut-off function. We have a different constants which depends also on c_{26}. Citing the result from [30, Proposition6.4], we obtain

$$\operatorname{spt}\|V_t\| \cap B_{4r/5} \subset \{|P^{\perp}(x)| \le \mu r\} \text{ for } t \in [(10c_{26})^{-2}r^2, \Lambda r^2], \tag{8.52}$$

where

$$\mu^2 := \frac{c_{5,KT}}{r^{n+3}}\int_0^{\Lambda r^2}\int_{B_r} |P^{\perp}(x)|^2\, d\|V_t\| dt + c_{2,KT}\|u\|^2_{L^{s,q}} D_1^{1-\frac{2}{s}} \Lambda^{\zeta} r^{2\zeta}(2+\Lambda) \tag{8.53}$$

and where $c_{5,KT}$ and $c_{2,KT}$ depend only on n, s, q and c_{26}. If we restrict r_2 further so that the second term of (8.53) is smaller than ε, (8.51)-(8.53) with sufficiently small ε gives

$$\operatorname{spt}\|V_t\| \cap B_{4r/5} \subset \{|P^{\perp}(x)| \le r/5\} \quad \text{for } t \in [(10c_{26})^{-2}r^2, \Lambda r^2]. \tag{8.54}$$

Combining (8.46) and (8.54), and using the L^1 continuity of $\varphi(\cdot,t)$, we obtain

$$\begin{aligned}&\varphi(\cdot,t)=1 \ \text{ on } B_{4r/5}\cap\{P^{\perp}(x)\ge r/5\},\\&\varphi(\cdot,t)=0 \ \text{ on } B_{4r/5}\cap\{P^{\perp}(x)\le -r/5\}\end{aligned}\tag{8.55}$$

for $t\in[0,\Lambda r^2]$. Since $B^{n-1}_{r/2}\times[-r/2,r/2]\subset B_{4r/5}$, (8.55) shows

$$\begin{aligned}&\varphi(\cdot,t)=1 \ \text{ on } B^{n-1}_{r/2}\times[r/5,r/2],\\&\varphi(\cdot,t)=0 \ \text{ on } B^{n-1}_{r/2}\times[-r/2,-r/5],\\&\mathrm{spt}\,\|V_t\|\cap(B^{n-1}_{r/2}\times[-r/2,r/2])\subset B^{n-1}_{r/2}\times[-r/5,r/5]\end{aligned}\tag{8.56}$$

for $t\in[0,\Lambda r^2]$ and $r\le r_2$. At this point, because of the third claim of (8.56), by setting $V_t=0$ on $B^{n-1}_{r/2}\times(\mathbb{R}\setminus[-r/2,r/2])$, we may assume that $\{V_t\}_{0\le t\le\Lambda r^2}$ satisfies (2.10) on $(B^{n-1}_{r/2}\times\mathbb{R})\times[0,\Lambda r^2]$. We next want to apply [30, Theorem 8.7] with $R:=r/6$. For the application, we need to check the conditions (8.83)–(8.86) of [30]. The first condition (8.83), the smallness of space-time L^2-height may be achieved due to (8.51), (8.56) and by restricting ε depending on $\varepsilon_{6,KT}$ and $\Lambda_{3,KT}$. The second condition (8.84), the smallness of $\|u\|$, may be achieved by simply restricting r_2. Thus we need to check the last two conditions, (8.85) and (8.86) of [30]. Let $\phi_{P,R}$ and $\mathbf{c}$ be defined as in [30, Definition 5.1]. We need to show that (recall that we have set $\nu=1/2$)

$$\exists t_1\in(3R^2/2,2R^2)\ \ :\ \ R^{-(n-1)}\|V_{t_1}\|(\phi^2_{P,R})<\frac{3}{2}\mathbf{c}\tag{8.57}$$

and

$$\exists t_2\in((2\Lambda_{3,KT}-2)R^2,(2\Lambda_{3,KT}-3/2)R^2)\ \ :\ \ R^{-(n-1)}\|V_{t_2}\|(\phi^2_{P,R})>\frac{1}{2}\mathbf{c}.\tag{8.58}$$

First we show (8.57). Since M_0 is C^1, we may restrict r_2 uniformly in x so that for all $R=r/6\le r_2/6$, we have

$$R^{-(n-1)}\|V_0\|(\phi^2_{P,R})\le R^{-(n-1)}\int_P\phi^2_{P,R}\,d\mathcal{H}^{n-1}+\frac{1}{10}\mathbf{c}=\frac{11}{10}\mathbf{c}.\tag{8.59}$$

By (2.10), we have for $t_1\in(3R^2/2,2R^2)$

$$\|V_t\|(\phi^2_{P,R})\Big|_{t=0}^{t_1}\le\int_0^{t_1}\int(-h\phi^2_{P,R}+\nabla\phi^2_{P,R})\cdot(h+(u\cdot\nu)\nu)\,d\|V_t\|dt.\tag{8.60}$$

By (2.5) and (2.6), we may replace $\nabla\phi^2_{P,R}$ by $S^{\perp}(\nabla\phi^2_{P,R})$ for $\|V_t\|$ a.e., where S is the approximate tangent space at the point. Since $\nabla\phi_{P,R}=P(\nabla\phi_{P,R})$ (note $\phi_{P,R}(x)=\phi_{P,R}(P(x))$ by definition), we have

$$S^{\perp}(\nabla\phi_{P,R}^2) = (I - S)\circ(P(\nabla\phi_{P,R}^2)) = (P - S)\circ(P(\nabla\phi_{P,R}^2)). \tag{8.61}$$

Thus, by using the Cauchy-Schwarz inequality to (8.60) and by (8.61), we obtain

$$\begin{aligned}\|V_t\|(\phi_{P,R}^2)\Big|_{t=0}^{t_1} \le &\int_0^{t_1}\int -\frac{1}{2}|h|^2\phi_{P,R}^2 + 2|u|^2\phi_{P,R}^2 \\ &+8\|S-P\|^2|\nabla\phi_{P,R}|^2\, dV_t(\cdot,S)dt.\end{aligned} \tag{8.62}$$

The first term on the right-hand side of (8.62) can be dropped. The second term can be estimated using the Hölder inequality as

$$\begin{aligned}\int_0^{t_1}\int 2|u|^2\phi_{P,R}^2\, d\|V_t\|dt &\le \int_0^{t_1}\left(\int |u|^s\, d\|V_t\|\right)^{\frac{2}{s}} dt \cdot \sup_{t\in[0,t_1]}\|V_t\|(\phi_{P,R}^2)^{1-\frac{2}{s}} \\ &\le \|u\|_{L^{s,q}}^2 t_1^{1-\frac{2}{q}} \cdot \sup_{t\in[0,t_1]}\|V_t\|(\phi_{P,R}^2)^{1-\frac{2}{s}}.\end{aligned} \tag{8.63}$$

Due to the third claim of (8.56), $\mathrm{spt}\,\|V_t\| \cap \mathrm{spt}\,\phi_{P,R} \subset B_{3R}$, for example. Thus we have $\|V_t\|(\phi_{P,R}^2) \le D_1\omega_{n-1}(3R)^{n-1}$. Since $t_1 \le 2R^2$, we obtain from (8.63)

$$\int_0^{t_1}\int 2|u|^2\phi_{P,R}^2\, d\|V_t\|dt \le c(D_1,n,s,q)\|u\|_{L^{s,q}}^2 R^{n-1+2\zeta}. \tag{8.64}$$

For the third term of (8.62), we use [30, Lemma 11.2] (or [1, 8.13]), namely, for $\phi = \phi_{P,R}$

$$\begin{aligned}\int \|S-P\|^2\, |\nabla\phi|^2\, dV_t(\cdot,S) \le 16\int |P^{\perp}(x)|^2|\nabla|\nabla\phi||^2\, d\|V_t\| \\ +4\Big(\int |h|^2|\nabla\phi|^2\, d\|V_t\|\Big)^{\frac{1}{2}}\Big(\int |P^{\perp}(x)|^2|\nabla\phi|^2\, d\|V_t\|\Big)^{\frac{1}{2}}.\end{aligned} \tag{8.65}$$

By repeating a similar argument leading to (8.62) with slightly larger test function which is 1 on $\mathrm{spt}\,\phi_{P,R}$, one can obtain

$$\int_0^{t_1}\int |h|^2|\nabla\phi_{P,R}|^2\, d\|V_t\| \le c(n)R^{n-3}. \tag{8.66}$$

Since we have $\mathrm{spt}\,\|V_t\| \cap \mathrm{spt}\,\phi_{P,R} \subset B_{3R}$ and by (8.51), we obtain

$$\int_0^{t_1}\int |P^{\perp}(x)|^2|\nabla\phi_{P,R}|^2\, d\|V_t\|dt \le 3\varepsilon(3R)^{n+1}t_1 \sup|\nabla\phi_{P,R}|^2 \le c(n)\varepsilon R^{n+1}. \tag{8.67}$$

Thus, by (8.66) and (8.67) and similarly estimating the last term, we obtain from (8.65) that

$$\int_0^{t_1}\int \|S-P\|^2|\nabla\phi_{P,R}|^2\,dV_t(\cdot,S)dt \le c(n)(\sqrt{\varepsilon}+\varepsilon)R^{n-1}. \tag{8.68}$$

Combining (8.59), (8.62), (8.64) and (8.68), we obtain

$$R^{-(n-1)}\|V_{t_1}\|(\phi^2_{P,R}) \le \frac{11}{10}\mathbf{c} + c(D_1,n,s,q)\|u\|^2_{L^{s,q}}R^{2\zeta} + c(n)(\sqrt{\varepsilon}+\varepsilon). \tag{8.69}$$

Thus, by restricting $r < r_2$ and ε in (8.69), we can guarantee that (8.57) holds. To see (8.58) holds, we use the first two claims of (8.56). Due to the unit density property, recall that for a.e. t, we have $\|V_t\| = \|\nabla\{\varphi(\cdot,t)=1\}\| = \mathcal{H}^{n-1}\lfloor_{\partial^*\{\varphi(\cdot,t)=1\}}$, where $\partial^* A$ denotes the reduced boundary of A (see [24]). Let ν_n be the x_n component of the inward pointing unit normal vector of $\partial^*\{\varphi(\cdot,t)=1\}$. We apply the generalized divergence theorem valid for sets of finite perimeter, in this case, $\{\varphi(\cdot,t)=1\}\cap\{x_n \le r/3\}$. Then we have for a.e. $t\in[0,\Lambda r^2]$

$$\begin{aligned}\int \phi^2_{P,R}\,d\|V_t\| &\ge \int_{\partial^*\{\varphi(\cdot,t)=1\}} \nu_n\phi^2_{P,R}\,d\mathcal{H}^{n-1}\\ &= -\int_{\{\varphi(\cdot,t)=1\}\cap\{x_n\le r/3\}} \partial_{x_n}\phi^2_{P,R}\,dx + \int_{\{x_n=r/3\}}\phi^2_{P,R}\,d\mathcal{H}^{n-1} = R^{n-1}\mathbf{c}\end{aligned} \tag{8.70}$$

since $\phi^2_{P,R}$ does not depend on x_n and by the definition of $\mathbf{c}$. In particular, we have proved (8.58). Now we are ready to apply [30, Theorem 8.7]. For all sufficiently small $\varepsilon > 0$, we have seen that we may choose r_2 independent of $x \in M_0$ such that all the assumptions of [30, Theorem 8.7] hold on $(B_{r/2}\times\mathbb{R})\times[0,\Lambda r^2]$ for all $r \le r_2$. The conclusion is that in $B^{n-1}_{\sigma_{KT}R}\times\mathbb{R}$ and for $t \in ((\Lambda_{3,KT}-1/4)R^2, (\Lambda_{3,KT}+1/4)R^2)$, spt $\|V_t\|$ is represented as a graph $F(\cdot,t)$ of $C^{1,\zeta}$ function and it is $C^{(1+\zeta)/2}$ in time, with $|\nabla F| + R^{-1}|F|$ bounded by a constant multiple of ε (see (8.89) of [30]). The argument up to this point can be carried out for each point on $x \in M_0$ uniformly and spt $\|V_t\|$ can be covered by such graphs. This shows that for all small $t > 0$, spt $\|V_t\|$ is $C^{1,\zeta}$ everywhere. We have the local graph representation as claimed in (4) and $t^{-1/2}$dist (spt $\|V_t\|, M_0) \to 0$ as $t \to 0$. It is possible that spt $\|V_t\|$ remains $C^{1,\zeta}$ for some more time, and let T_2 be the maximal time without non-$C^{1,\zeta}$ regular point. In case that u is α-Hölder continuous, the regularity criterion are the same (see [46, Theorem 3.6]) except that the constant corresponding to $\varepsilon_{6,KT}$ may need to be smaller there. Thus, in this case, there is a short initial time interval such that spt $\|V_t\|$ is a $C^{2,\alpha}$ hypersurface. This ends the proof of (2) and (4). □

9 Final remarks

9.1 Non-uniqueness

The solution may be non-unique without having singularities of M_t, as a simple example demonstrates. An example such as $M_0 = \{x_2 = 0\} \subset \mathbb{T}^2$ and $u(x_1,x_2) =$

$(0, \sqrt{|x_2|}) \in (W^{1,p}(\mathbb{T}^2))^2$ $(p < 2)$ has an obvious ODE-level non-uniqueness. Thus, on top of the non-uniqueness issues generally associated with singularity occurrences of the MCF, one has far richer source of possible non-uniqueness with irregular u, even though we have a local regularity theory. It is interesting to investigate how generic the uniqueness may hold for the flow in this paper with respect to the initial data and the transport term. We mention that there is a nice generic property for the MCF besides the existence of unique viscosity solution. If M_0 is C^2 and d_0 is the signed distance function to M_0, then the viscosity solution for the MCF starting from $\{d_0 = s\}$ in the sense of [11,16] is a unit density Brakke MCF for a.e. $s \in (-r_0, r_0)$, where $r_0 > 0$ is some small number depending on M_0 [18]. For such level set, a phenomena called fattening does not occur in particular. It is interesting to see if there is some generalization of this type to the setting of this paper.

9.2 Structure of singularities

There have been intensive effort to understand the nature of singularities for the MCF in recent years. A particular emphasis has been placed on the mean convex flow and we mention names of Andrews, Huisken, Sinestrari and White who analyzed structure of singularities in depth. We mention a recent work by Haslhofer and Kleiner [25] for a streamlined treatment of the regularity theory of mean convex flows as well as up-to-date references. Note that many of the techniques used by White such as the dimension reducing and stratification of singularities [47] may be used for the flow in this paper. While there may be some limitation compared to the mean convex flow, it is interesting and challenging problem to investigate the singularities in the setting of the present paper.

References

1. Allard, W.: On the first variation of a varifold. Ann. Math. **95**(2), 417–491 (1972)
2. Allen, S.M., Cahn, J.W.: A macroscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. Acta. Metal. **27**, 1085–1095 (1979)
3. Almgren, F.J., Taylor, J.E., Wang, L.-H.: Curvature-driven flows: a variational approach. SIAM J. Control Optim. **31**(2), 387–438 (1993)
4. Ambrosio, L.: Geometric Evolution Problems, Distance Function and Viscosity Solutions. Calculus of Variations and Partial Differential Equations (Pisa, 1996), pp. 5–93. Springer, Berlin (1996)
5. Bellettini, G.: Lecture notes on mean curvature flow: barriers and singular perturbations. Edizioni Della Normale, 12, Scuola Normale Supriore, Pisa (2014)
6. Brakke, K.: The Motion of a Surface by Its Mean Curvature. Mathematical Notes 20. Princeton University Press, Princeton (1978)
7. Bronsard, L., Kohn, R.V.: Motion by mean curvature as the singular limit of Ginzburg-Landau dynamics. J. Differ. Equ. **90**(2), 211–237 (1991)
8. Bronsard, L., Stoth, B.: On the existence of high multiplicity interfaces. Math. Res. Lett. **3**(1), 41–50 (1996)
9. Chen, X.: Generation and propagation of interface in reaction-diffusion equations. J. Differ. Equ. **96**(1), 116–141 (1992)
10. Chen, X.-Y.: Dynamics of interfaces in reaction diffusion systems. Hiroshima Math. J. **21**(1), 47–83 (1991)
11. Chen, Y.-G., Giga, Y., Goto, S.: Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. J. Differ. Geom. **33**(3), 749–786 (1991)

12. Colding, T.H., Minicozzi II, W.P.: Minimal surfaces and mean curvature flow. Surveys in geometric analysis and relativity. Adv. Lect. Math. **20**, 73–143 (2011)
13. de Mottoni, P., Schatzman, M.: Evolution géométric d'interfaces. C. R. Acad. Paris Sci. Sér. I Math. **309**(7), 453–4558 (1989)
14. Ecker, K.: Regularity theory for mean curvature flow. In: Progress in Nonlinear Differential Equations and their Applications 57. Birkhäuser Boston Inc, Boston (2004)
15. Evans, L.C., Soner, H.M., Souganidis, P.E.: Phase transitions and generalized motion by mean curvature. Commun. Pure Appl. Math. **45**(9), 1097–1123 (1992)
16. Evans, L.C., Spruck, J.: Motion of level sets by mean curvature I. J. Differ. Geom. **33**(3), 635–681 (1991)
17. Evans, L.E., Spruck, J.: Motion of level sets by mean curvature II. Trans. Am. Math. Soc. **330**(1), 321–332 (1992)
18. Evans, L.C., Spruck, J.: Motion of level sets by mean curvature IV. J. Geom. Anal. **5**(1), 77–114 (1995)
19. Fife, P.C.: Dynamics of internal layers and diffusive interfaces. In: CBMS-NSF Reg. Conf. Ser. Applied Math., vol. 53. SIAM, Philadelphia (1988)
20. Giga, Y.: Surface Evolution Equations. A Level Set Approach. Monographs in Mathematics 99. Birkhäuser Verlag, Basel (2006)
21. Giga, Y., Goto, S.: Geometric evolution of phase-boundaries, On the evolution of phase boundaries (Minneapolis, MN, 1990–91). IMA Vol. Math. Appl. **43**, 51–65 (1992)
22. Giga, Y., Goto, S., Ishii, H., Sato, M.-H.: Comparison principle and convexity preserving properties for singular degenerate parabolic equations on unbounded domains. Indiana Univ. Math. J. **40**(2), 443–470 (1991)
23. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order, 2nd edn. Springer, Berlin (1983)
24. Giusti, E.: Minimal Surfaces and Functions of Bounded Variation. Monographs in Mathematics, vol. 80. Birkhäuser, Basel (1984)
25. Haslhofer, R., Kleiner, B.: Mean curvature flow of mean convex hypersurfaces. arXiv:1304.0926
26. Huisken, G.: Asymptotic behavior for singularities of the mean curvature flow. J. Differ. Geom. **31**(1), 285–299 (1990)
27. Hutchinson, J.E., Tonegawa, Y.: Convergence of phase interfaces in the van der Waals-Cahn-Hilliard theory. Calc. Var. PDE **10**(1), 49–84 (2000)
28. Ilmanen, T.: Convergence of the Allen-Cahn equation to Brakke's motion by mean curvature. J. Differ. Geom. **38**(2), 417–461 (1993)
29. Ilmanen, T.: Elliptic regularization and partial regularity for motion by mean curvature. Mem. Am. Math. Soc. **108**, 520 (1994)
30. Kasai, K., Tonegawa, Y.: A general regularity theory for weak mean curvature flow. Calc. Var. PDE **50**(1), 1–68 (2014)
31. Kim, Y., Lai, M.-C., Peskin, C.: Numerical simulations of two-dimensional foam by the immersed boundary method. J. Comput. Phys. **229**(13), 5194–5207 (2010)
32. Ladyženskaja, O.A., Solonnikov, V.A., Ural'ceva, N.N.: Linear and quasi-linear equations of parabolic type. Am. Math. Soc. (1968, Translated from Russian)
33. Liu, C., Sato, N., Tonegawa, Y.: On the existence of mean curvature flow with transport term. Interf. Free Bound **12**(2), 251–277 (2010)
34. Luckhaus, S., Sturzenhecker, T.: Implicit time discretization for the mean curvature flow equation. Calc. Var. PDE **3**(2), 253–271 (1995)
35. Mantegazza, C.: Lecture Notes on Mean Curvature Flow. Progress in Mathematics, vol. 290. Birkhäuser/Springer Basel AG, Basel (2011)
36. Meyers, N.G., Ziemmer, W.P.: Integral inequalities of Poincaré and Wirtinger type for BV functions. Am. J. Math. **99**(6), 1345–1360 (1977)
37. Mugnai, L., Röger, M.: Convergence of perturbed Allen-Cahn equations to forced mean curvature flow. Indiana Univ. Math. J. **60**(1), 41–75 (2011)
38. Röger, M., Schätzle, R.: On a modified conjecture of De Giorgi. Math. Z. **254**(4), 675–714 (2006)
39. Rubinstein, J., Sternberg, P., Keller, J.B.: Fast reaction, slow diffusion and curve shortening. SIAM J. Appl. Math. **49**(1), 116–133 (1989)
40. Sato, N.: A simple proof of convergence of the Allen-Cahn equation to Brakke's motion by mean curvature. Indiana Univ. Math. J. **57**(4), 1743–1751 (2008)

41. Simon, L.: Lectures on Geometric Measure Theory. Proc. Centre Math. Anal. Austral. Nat. Univ. 3 (1983)
42. Soner, H.M.: Convergence of the phase-field equations to the Mullins-Sekerka problem with kinetic undercooling. Arch. Rational Mech. Anal. **131**(2), 139–197 (1995)
43. Soner, H.M.: Ginzburg-Landau equation and motion by mean curvature. Converg. J. Geom. Anal. I. **7**(3), 437–475 (1997)
44. Takasao, K.: Gradient estimates and existence of mean curvature flow with transport term. Differ. Integral Equ. **26**(1–2), 141–154 (2013)
45. Tonegawa, Y.: Integrality of varifolds in the singular limit of reaction-diffusion equations. Hiroshima Math. J. **33**(3), 323–341 (2003)
46. Tonegawa, Y.: A second derivative Hölder estimate for weak mean curvature flow. Adv. Calc. Var. **7**(1), 91–138 (2014)
47. White, B.: Stratification of minimal surfaces, mean curvature flows, and harmonic maps. J. Reine Angew. Math. **488**, 1–35 (1997)
48. White, B.: Evolution of curves and surfaces by mean curvature. In: Proceedings of the ICM, 1 (Beijing, 2002), pp. 525–538. Higher Ed. Press, Beijing (2002)
49. Ziemer, W.P.: Weakly Differentiable Functions. Springer, Berlin (1989)