J. Eur. Math. Soc. 19, 2997–3051 

DOI 10.4171/JEMS/734

Kelei Wang

# A new proof of Savin's theorem on Allen–Cahn equations

Received November 1, 2014

**Abstract.** In this paper we establish an improvement of the tilt-excess decay estimate for the Allen–Cahn equation, and use this to give a new proof of Savin's theorem on the uniform $C^{1,\alpha}$ regularity of flat level sets. This generalizes Allard's $\varepsilon$-regularity theorem for stationary varifolds to the setting of Allen–Cahn equations. A new proof of Savin's theorem on the one-dimensional symmetry of minimizers in $\mathbb{R}^n$ for $n \le 7$ is also given.

**Keywords.** Allen–Cahn equation, phase transition, improvement of tilt-excess decay, harmonic approximation, De Giorgi conjecture

## Contents



K. Wang: School of Mathematics and Statistics & Computational Science Hubei Key Laboratory, Wuhan University, Wuhan 430072, China; e-mail: wangkelei@whu.edu.cn

*Mathematics Subject Classification (2010):* 35B06, 35B08, 35B25, 35J91

## 1. Introduction

This paper is devoted to generalizing Allard's regularity theory in geometric measure theory to the setting of Allen–Cahn equations and discussing its application to the De Giorgi conjecture.

The Allen–Cahn equation

$$\Delta u = u^3 - u \tag{1.1}$$

is a typical model of phase transition. By now, it has been studied from various aspects. One particular feature of this equation is its close relation to minimal surface theory, through its singularly perturbed version

$$\varepsilon \Delta u_\varepsilon = \frac{1}{\varepsilon}(u_\varepsilon^3 - u_\varepsilon).$$

By this connection and in view of the Bernstein theorem for minimal hypersurfaces [23], De Giorgi made the following conjecture in [8]:

*Let $u \in C^2(\mathbb{R}^{n+1})$ be a solution of (1.1) such that*

$$|u| \le 1, \qquad \frac{\partial u}{\partial x_{n+1}} > 0 \quad \text{in } \mathbb{R}^{n+1}.$$

*Then $u$ depends only on one variable if $n \le 7$.*

This conjecture has been considered by many authors, including Ghoussoub and Gui [14], Ambrosio and Cabré [3] and Savin [20]. Counterexamples in $\mathbb{R}^9$ were also constructed by del Pino, Kowalczyk and Wei [10]. In particular, Savin proved an improvement-of-flatness result for minimizing solutions (i.e. minimizers of the energy functional). This result says that, given any $\theta_0 > 0$, for a minimizer $u$, if in a ball $\mathcal{B}_l$ with $l$ large, its zero level set is trapped in a strip $\{|x_{n+1}| < \theta\}$ with $\theta \ge \theta_0$, which is sufficiently narrow (i.e. $\theta l^{-1}$ is small), then by shrinking the radius of the ball, possibly after a rotation of coordinates, the zero level set of $u$ is trapped in a flatter strip.

By using this estimate, Savin proved

**Theorem 1.1.** *Let $u$ be a minimizing solution of (1.1) defined on the entire space $\mathbb{R}^{n+1}$ where $n \le 6$. Then $u$ is one-dimensional.*

For $n > 6$, if we add some further assumptions on level sets of $u$, e.g. the global Lipschitz regularity of $\{u = 0\}$, it is still possible to prove the one-dimensional symmetry of $u$. This theorem also implies the original De Giorgi conjecture, under an additional assumption that

$$\lim_{x_{n+1} \to \pm\infty} u(x, x_{n+1}) = \pm 1.$$

This type of improvement-of-flatness result appears in the partial regularity theory for various elliptic problems, although sometimes in rather different forms. One main ingredient to establish this improvement of flatness is the *blow up* (or *harmonic approximation*) technique, first introduced by De Giorgi in his work [7] on the almost everywhere regularity of minimal hypersurfaces.

Although the statement of Savin's improvement-of-flatness result bears many similarities with the De Giorgi theorem, the proof in [20] employs some new ideas. Indeed, it is based on Caffarelli–Córdoba's proof of the De Giorgi theorem in [5]. This approach uses the "viscosity" side of the problem, and relies heavily on a Krylov–Safonov type argument. In particular, Savin first obtained a Harnack inequality (hence some kind of uniform Hölder continuity) and then used it to prove that the blow up sequence converges uniformly to a harmonic function.

Savin's approach can be applied to many other problems, even without variational structure—see for example [25, 21, 22]. However, it seems that the maximum principle and Harnack inequality are crucial in this approach. At present it is still not clear how to get this kind of improvement-of-flatness result for elliptic systems, where the Harnack inequality may fail. Thus, in view of the connection between Allen–Cahn equations and minimal hypersurfaces, in this paper we intend to explore the *variational* side of improvement of flatness and establish some results paralleling classical regularity theories in geometric measure theory. As in Allard's regularity theory [2] (see also [16, Section 6.5] for an account), we use the following *excess* (for more details, see Sections 2 and 3)

$$\int_{\mathcal{C}_1(0)} [1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2,$$

where $\mathcal{C}_1(0)$ is the cylinder $B_1(0)\times(-1,1)\subset\mathbb{R}^{n+1}$ and $\nu_\varepsilon=\nabla u_\varepsilon/|\nabla u_\varepsilon|$ is the unit normal vector to level sets of $u_\varepsilon$. This quantity was first used by Hutchinson–Tonegawa [15] to derive the integer multiplicity of the limit varifold arising from general critical points in the Allen–Cahn problem.

This quantity can be used to measure the flatness of level sets of $u_\varepsilon$ (see Lemma 4.6 below). Similar to Allard's $\varepsilon$-regularity theorem, if the excess in a ball is small, then after shrinking the radius of the ball and possibly rotating the vector $e_{n+1}$ a little, the excess becomes smaller. This *improvement of tilt-excess* is the main step in the proof of Allard's $\varepsilon$-regularity theorem, and also in our argument. In contrast to the quantity used in Savin's improvement-of-flatness result, the excess is an energy type quantity. Indeed, if all the level sets $\{u_\varepsilon=t\}$ can be represented by graphs along the $(n+1)$-th direction, in the form $\{x_{n+1}=h(x,t)\}$, then the excess can be written as

$$\int_{\mathcal{C}_1(0)} [1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2 = \int_{-1}^{1}\left(\int_{B_1(0)}\frac{|\nabla_x h(x,t)|^2}{1+|\nabla_x h(x,t)|^2}\varepsilon|\nabla u_\varepsilon(x,h(x,t))|\,dx\right)dt,$$

which is almost a weighted Dirichlet energy, provided $\sup|\nabla_x h(x,t)|$ is small.

Thus the problem can be approximated by harmonic functions (corresponding to critical points of the Dirichlet energy) if $|\nabla_x h(x,t)|$ is small. (We will see that just the smallness of the excess is sufficient for this purpose.) This is exactly the content of *harmonic approximation* technique. Using the excess allows us to work in Sobolev spaces and apply standard compact Sobolev embedding results to get the blow up limit, while in Savin's version the main difficulty lies in the compactness for the blow up sequence where his Harnack inequality enters.

We also note that this type of tilt-excess decay result was known to Tonegawa [24], who he showed that this result implies the uniform $C^{1,\alpha}$ regularity of intermediate transition layers in dimension 2.

However, in this tilt-excess decay estimate we need one more assumption:

$$\int_{\mathcal{C}_1}[1-(\nu_\varepsilon\cdot e)^2]\varepsilon|\nabla u_\varepsilon|^2 \gg \varepsilon^2. \tag{1.2}$$

Compared to Allard's $\varepsilon$-regularity theorem, this condition is not so satisfactory. It prevents us from applying this improvement of decay directly to deduce the uniform $C^{1,\alpha}$ regularity of intermediate transition layers. (This obstruction was also observed by Tonegewa [24].) One reason for the appearance of the condition (1.2) is that the energy, although mostly concentrated on the transition part, is still distributed on a layer of width $\varepsilon$. Note that this phenomenon does not appear in minimal surface theory.

In Savin's version of improvement of flatness, an assumption similar to (1.2) is also needed. Using our terminology, it is equivalent to requiring that the excess is not of the order $o(\varepsilon^2)$. Note that this is weaker than (1.2). This weaker assumption is perhaps due to the fact that in Savin's approach only a single level set is considered, while our improvement of flatness involves a family of level sets.

By exploiting the fact that $u_\varepsilon$ is close to a one-dimensional solution up to $O(\varepsilon)$ scales, an iteration of the improvement of tilt-excess decay estimate gives a Morrey type bound on level sets of $u_\varepsilon$, which then implies that these level sets are graphs. Here, once again due to the obstruction (1.2), this Morrey type bound does not imply the $C^{1,\alpha}$ regularity of $\{u_\varepsilon = 0\}$, but only a Lipschitz one. However, under the condition that $\{u_\varepsilon = 0\}$ is a Lipschitz graph, Caffarelli and Córdoba [6] have shown that transition layers are uniformly bounded in $C^{1,\alpha}$ for some $\alpha \in (0,1)$. Thus we get a full analogue of Allard's $\varepsilon$-regularity theorem in the Allen–Cahn setting (see Theorem 9.1).

In this paper we do not fully avoid the use of the maximum principle. For example, it seems that the Modica inequality is indispensable in our argument, because we need it to derive a monotonicity formula with the correct exponent. We also need to apply the moving plane (or sliding) method (as in Farina [12]) to deduce the one-dimensional symmetry of some entire solutions. There a distance type function is used to control the behavior of $u$ far from the transition part. This function behaves like a distance function, which follows from the Modica inequality. However, we do avoid the use of any Harnack inequality. It may be possible to remove the above mentioned deficiency by strengthening the tilt-excess decay estimate, but as explained above, the current version of Theorem 3.3 is already sufficient for proving Theorem 1.1 (see Section 11).

The above approach was first used by the author in [26], where we consider a De Giorgi type conjecture for the elliptic system

$$\Delta u = uv^2, \quad \Delta v = vu^2, \quad u, v > 0 \quad \text{in } \mathbb{R}^n.$$

For the corresponding singularly perturbed system

$$\begin{cases} \Delta u_\kappa = \kappa u_\kappa v_\kappa^2, \\ \Delta v_\kappa = \kappa v_\kappa u_\kappa^2, \end{cases}$$

an improvement-of-flatness result was established by using the quantity

$$\int_{B_1} |\nabla(u_\kappa - v_\kappa - e \cdot x)|^2 \, dx.$$

These two proofs are similar in spirit. In particular, in order to show that the blow up limit is a harmonic function, we mainly use the *stationary condition* associated to the equation, not the equation itself. This is more apparent in the current setting, because the stationary condition for the singularly perturbed Allen–Cahn equation is directly linked to the corresponding one in the limit problem, the stationary condition for varifolds (in the sense of Allard [2]; see also [15]). Furthermore, since the excess is a kind of $H^1$ norm, to prove strong convergence in $H^1$, we implicitly use a Caccioppoli type inequality, which is again deduced from the stationary condition by choosing a suitable test function (see Remark 4.7 below).

Finally, although in our improvement of tilt-excess decay (Theorem 3.3) and the $\varepsilon$-regularity result (Theorem 9.1), we do not assume that the solution is a minimizer, the multiplicity one property of transition layers is still needed here. (This is associated to the unit density property of the limit varifold.) Thus our results do not remove the no folding assumption in Savin's result. However, we feel that a generalization of our technique to the case with multiple transition layers is possible, which should be of more interest.

The organization of this paper can be seen from the table of contents. Part I is devoted to proving the tilt-excess decay estimate (Theorem 3.3). In Part II, we establish an Allard type $\varepsilon$-regularity theorem, the uniform $C^{1,\alpha}$ regularity of intermediate layers (Theorem 9.1). In the proof of this theorem, a De Giorgi type conjecture for a class of entire solutions (Theorem 11.1) is also obtained, which includes Theorem 1.1 as a special case.

## 2. Settings and notation

We shall work in the following settings. Consider the Allen–Cahn equation in the general form as

$$\Delta u = W'(u), \tag{2.1}$$

where $W$ is a double well potential, that is, $W \in C^3(\mathbb{R})$ satisfying

- $W \geq 0$, $W(\pm 1) = 0$ and $W > 0$ in $(-1, 1)$;
- for some $\gamma \in (0, 1)$, $W' < 0$ on $(\gamma, 1)$ and $W' > 0$ on $(-1, -\gamma)$;
- $W'' \geq \kappa > 0$ for all $|x| \geq \gamma$.

A typical example is $W(u) = (1 - u^2)^2/4$, which gives (1.1).

Through a scaling $u_\varepsilon(X) := u(\varepsilon^{-1}X)$, we get the singularly perturbed version of the Allen–Cahn equation:

$$\varepsilon \Delta u_\varepsilon = \frac{1}{\varepsilon} W'(u_\varepsilon). \tag{2.2}$$

This equation arises as the Euler–Lagrange equation of the energy functional (after adding suitable boundary conditions)

$$E_\varepsilon(u_\varepsilon) = \int \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right). \tag{2.3}$$

We say $u_\varepsilon$ is a *minimizer* (or a *minimizing solution*) if for every ball $\mathcal{B}$ in the definition domain of $u_\varepsilon$,

$$\int_{\mathcal{B}} \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \le \int_{\mathcal{B}} \left( \frac{\varepsilon}{2} |\nabla v|^2 + \frac{1}{\varepsilon} W(v) \right)$$

for any $v \in H^1(\mathcal{B})$ satisfying $v = u_\varepsilon$ on $\partial \mathcal{B}$.

We will always assume $|u_\varepsilon| \le 1$, and that it satisfies the *Modica inequality*

$$\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 \le \frac{1}{\varepsilon} W(u_\varepsilon). \tag{2.4}$$

This inequality (in the exact form as above) may not be essential in the argument, but we prefer to assume it to make the arguments clean. (These estimates can be relaxed, cf. [15].) By standard elliptic estimates, there exists a universal constant $C$ such that

$$\varepsilon |\nabla u_\varepsilon| + \varepsilon^2 |\nabla^2 u_\varepsilon| \le C. \tag{2.5}$$

In particular, $u_\varepsilon$ is a classical solution.

For any smooth vector field $Y$ with compact support, by considering the domain variation in the form

$$u_\varepsilon^t(X) := u_\varepsilon(X + tY(X)) \quad \text{for } |t| \text{ small},$$

from the definition of critical points we get

$$\frac{d}{dt} \int \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \Big|_{t=0} = 0.$$

After some integration by parts, we obtain the *stationary condition* for $u_\varepsilon$:

$$\int \left[ \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \operatorname{div} Y - \varepsilon DY(\nabla u_\varepsilon, \nabla u_\varepsilon) \right] = 0. \tag{2.6}$$

Finally, by the assumption on $W$, there exists a one-dimensional solution $g(t)$ defined on $t \in (-\infty, \infty)$, satisfying

$$g''(t) = W'(g(t)), \qquad \lim_{t \to \pm\infty} g(t) = \pm 1, \tag{2.7}$$

where the convergence rate is exponential.

The first integral for (2.7) can be written as

$$g'(t) = \sqrt{2W(g(t))} > 0. \tag{2.8}$$

For any $\varepsilon > 0$, we denote $g_\varepsilon(t) = g(\varepsilon^{-1}t)$, which satisfies

$$\varepsilon g_\varepsilon''(t) = \frac{1}{\varepsilon} W'(g_\varepsilon(t)). \tag{2.9}$$

Throughout, $\sigma_0$ denotes the constant defined by

$$\sigma_0 := \int_{-\infty}^{\infty} g'(t)^2\, dt = \int_{-\infty}^{\infty} \left(\frac{1}{2} g'(t)^2 + W(g(t))\right) dt. \tag{2.10}$$

In this paper we adopt the following notation.

- A point in $\mathbb{R}^{n+1}$ will be denoted by $X = (x, x_{n+1}) \in \mathbb{R}^n \times \mathbb{R}$.
- $\mathcal{B}_r(X)$ denotes an open ball in $\mathbb{R}^{n+1}$ and $B_r(x)$ an open ball in $\mathbb{R}^n$. If the center is the origin 0, we write $\mathcal{B}_r$ (or $B_r$).
- $\mathcal{C}_r(x) = B_r(x) \times (-1, 1) \subset \mathbb{R}^{n+1}$, the finite cylinder over $B_r(x) \subset \mathbb{R}^n$.
- $e_i$, $1 \le i \le n+1$, is the standard basis in $\mathbb{R}^{n+1}$.
- $P$ denotes a hyperplane in $\mathbb{R}^{n+1}$ and $\Pi_P$ (or simply $P$) the orthogonal projection onto it. If $P = \mathbb{R}^n$, we use $\Pi$.
- $G(n)$ denotes the Grassmann manifold of unoriented $n$-dimensional hyperplanes in $\mathbb{R}^{n+1}$.
- A *varifold* $V$ is a Radon measure on $\mathbb{R}^{n+1} \times G(n)$. We use $\|V\|$ to denote the weighted measure of $V$, that is, for any measurable set $A \subset \mathbb{R}^{n+1}$,

$$\|V\|(A) = V(A \times G(n)).$$

- For a measure $\mu$, spt $\mu$ denotes its support.
- $\nu_\varepsilon(X) = \nabla u_\varepsilon(X)/|\nabla u_\varepsilon(X)|$ if $\nabla u_\varepsilon(X) \neq 0$, otherwise we take it to be an arbitrary unit vector.
- $\mu_\varepsilon := \varepsilon |\nabla u_\varepsilon|^2 dX$.
- $\mathcal{H}^s$ denotes the $s$-dimensional Hausdorff measure.
- $\omega_n$ denotes the volume of the unit ball $B_1$ in $\mathbb{R}^n$.
- $H^1$ is the Sobolev space with the norm $(\int (|\nabla u|^2 + |u|^2))^{1/2}$.
- $\mathrm{dist}_H$ is the Hausdorff distance between sets in $\mathbb{R}^{n+1}$.
- Unless otherwise stated, universal constants $C$, $C_i$ and $K_i$ (large) and $c_i$ (small) depend only on the dimension $n$ and the potential function $W$.

Throughout, $u_\varepsilon$ always denotes a solution of (2.2). We use $\varepsilon$ to denote a sequence of parameters converging to 0, which should be written as $\varepsilon_i$ if we want to be precise.

# Part I. Tilt-excess decay

## 3. Statement

The following quantity will play an important role in our analysis.

**Definition 3.1** (Excess). Let $P$ be an $n$-dimensional hyperplane in $\mathbb{R}^{n+1}$ and $e$ one of its unit normal vectors, $B_r(x) \subset P$ an open ball and $\mathcal{C}_r(x) = B_r(x) \times (-1, 1)$ the cylinder over $B_r(x)$. The *excess* of $u_\varepsilon$ in $\mathcal{C}_r(x)$ with respect to $P$ is

$$E(r; x, u_\varepsilon, P) := r^{-n} \int_{\mathcal{C}_r(x)} [1 - (\nu_\varepsilon \cdot e)^2] \varepsilon |\nabla u_\varepsilon|^2 \, dX. \tag{3.1}$$

If $P = \mathbb{R}^n$ and $e = e_{n+1}$, the excess equals

$$E(r; x, u_\varepsilon) = r^{-n} \int_{\mathcal{C}_r(x)} \varepsilon \sum_{i=1}^{n} \left( \frac{\partial u_\varepsilon}{\partial x_i} \right)^2 dX.$$

**Remark 3.2.** For any unit vectors $\nu$ and $e$, we have

$$|\nu - e| \, |\nu + e| \geq \sqrt{2} \min\{|\nu - e|, |\nu + e|\}.$$

Therefore

$$\begin{aligned} 1 - (\nu \cdot e)^2 = (1 - \nu \cdot e)(1 + \nu \cdot e) &= \tfrac{1}{4} |\nu - e|^2 |\nu + e|^2 \\ &\geq \tfrac{1}{2} \min\{|\nu - e|^2, |\nu + e|^2\}. \end{aligned}$$

By projecting the unit sphere $\mathbb{S}^n$ to the real projective space $\mathbb{RP}^n$ (both with the standard metric), we get

$$1 - (\nu \cdot e)^2 \geq c \, \mathrm{dist}_{\mathbb{RP}^n}(\nu, e)^2,$$

for some universal constant $c$.

Our main objective in Part I is to prove the following decay estimate.

**Theorem 3.3** (Tilt-excess decay). *Given a constant $b \in (0, 1)$, there exist five universal constants $\delta_0, \tau_0, \varepsilon_0 > 0$, $\theta \in (0, 1/4)$ and $K_0$ large such that the following holds. Let $u_\varepsilon$ be a solution of (2.2) with $\varepsilon \leq \varepsilon_0$ in $\mathcal{B}_4$, satisfying the Modica inequality (2.4), $|u_\varepsilon(0)| \leq 1 - b$, and*

$$4^{-n} \int_{\mathcal{B}_4} \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \leq (1 + \tau_0) \sigma_0 \omega_n. \tag{3.2}$$

*Suppose the excess with respect to $\mathbb{R}^n$ satisfies*

$$\delta_\varepsilon^2 := E(2; 0, u_\varepsilon, \mathbb{R}^n) \leq \delta_0^2, \tag{3.3}$$

*where $\delta_\varepsilon \geq K_0 \varepsilon$. Then there exists another plane $P$ such that*

$$E(\theta; 0, u_\varepsilon, P) \leq \frac{\theta}{2} E(2; 0, u_\varepsilon, \mathbb{R}^n). \tag{3.4}$$

*Moreover, there exists a universal constant $C$ such that*

$$\|e - e_{n+1}\| \leq C E(2; 0, u_\varepsilon, \mathbb{R}^n)^{1/2},$$

*where $e$ is the upward pointing unit normal vector to $P$.*

Roughly speaking, this theorem says that if the excess (with respect to some hyperplane) in a ball is small enough, then after shrinking the radius of the ball and perhaps tilting the hyperplane a little, the excess becomes smaller. This decay estimate will be used in Part II to prove the uniform Lipschitz regularity of intermediate layers.

The condition (3.2) says there is only a single transition layer, which corresponds to the unit density assumption in Allard's $\varepsilon$-regularity theorem. In the next section we shall see that (3.3) always holds (with respect to a suitable hyperplane), provided that (3.2) is satisfied with $\tau_0$ sufficiently small (depending on $\delta_0$). However, the assumption that $\delta_\varepsilon \gg \varepsilon$ is crucial here, which is not so satisfactory compared to Allard's and Savin's version.

We shall prove this theorem indirectly. So assume there exists a sequence $\varepsilon_i$ (for simplicity the subscript $i$ will be dropped) and a sequence of solutions $u_\varepsilon$ satisfying all of the assumptions in Theorem 3.3, that is,

- there exists a sequence $\tau_\varepsilon \to 0$ such that

$$4^{-n}\int_{\mathcal{B}_4}\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right)\leq(1+\tau_\varepsilon)\sigma_0\omega_n, \tag{3.5}$$

- the excess satisfies

$$\delta_\varepsilon^2 := E(2;0,u_\varepsilon,\mathbb{R}^n)\to 0, \tag{3.6}$$

  where

$$\delta_\varepsilon/\varepsilon\to\infty \qquad \text{as } \varepsilon\to 0, \tag{3.7}$$

but for any unit vector $e$ with

$$\|e-e_{n+1}\|\leq CE(2;0,u_\varepsilon,\mathbb{R}^n)^{1/2}, \tag{3.8}$$

where the constant $C$ will be determined below (by the constant in (8.1)), we have

$$E(\theta;0,u_\varepsilon,P)\geq\frac{\theta}{2}E(2;0,u_\varepsilon,\mathbb{R}^n). \tag{3.9}$$

Here $P$ is the hyperplane orthogonal to $e$ and $\theta$ is also a constant to be determined later (see (8.11) and (8.16)).

The remaining part, up to and including Section 8, will be devoted to deriving a contradiction from the assumptions (3.5)–(3.9). The proof is divided into four steps:

**Step 1.** It is shown that $\{u_\varepsilon = t\}$ (for $t\in(-1+b,1-b)$) can be represented by Lipschitz graphs over $\mathbb{R}^n$, $x_{n+1}=h_\varepsilon^t(x)$, except a bad set of small measure (controlled by $E(2;0,u_\varepsilon,\mathbb{R}^n)$). This is achieved by the weak $L^1$ estimate for Hardy–Littlewood maximal functions.

**Step 2.** By writing the excess using the $(x,t)$ coordinates ($t$ as in Step 1), $h_\varepsilon^t/\delta_\varepsilon$ are uniformly bounded in $H^1_{loc}(B_1)$. Then we can assume that they converge weakly to a limit $h_\infty$. Here we need the assumption $\delta_\varepsilon \gg \varepsilon$ to guarantee the limit is independent of $t$.

**Step 3.** By choosing the vector field $Y=\varphi\psi e_{n+1}$ in the stationary condition (2.6), where $\varphi\in C_0^\infty(B_1)$ and $\psi\in C_0^\infty((-1,1))$, and then passing to the limit, it is shown that $h_\infty$ is harmonic in $B_1$.

**Step 4.** By choosing the vector field $Y = \varphi\psi x_{n+1} e_{n+1}$ in the stationary condition (2.6) and then passing to the limit, it is shown that (roughly speaking) $h^t_\varepsilon/\delta_\varepsilon$ converges strongly in $H^1_{loc}(B_1)$. The tilt-excess decay estimate then follows from some basic estimates on harmonic functions.

After establishing some preliminary results in the next section, Steps 1–4 will be done in Sections 5–8 respectively.

## 4. Compactness results

In this section, we study the convergence of various quantities associated to $u_\varepsilon$ and establish some preliminary results for the proof of Theorem 3.3.

Recall that we have assumed the Modica inequality (2.4). An important consequence of this inequality is the following monotonicity formula (see for example [19]).

**Proposition 4.1** (Monotonicity formula). *For any $X \in \mathcal{B}_3$,*

$$r^{-n} \int_{\mathcal{B}_r(X)} \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right)$$

*is non-decreasing in $r \in (0, 1)$.*

By combining Proposition 4.1 with (3.5), we get

**Corollary 4.2.** *For any $\mathcal{B}_r(X) \subset \mathcal{B}_3$, we have*

$$\int_{\mathcal{B}_r(X)} \left( \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \right) \le 8^n \sigma_0 \omega_n r^n. \tag{4.1}$$

We use the main result in Hutchinson–Tonegawa [15] to study the convergence of $u_\varepsilon$. Define the varifold $V_\varepsilon$ by

$$\langle V_\varepsilon, \Phi(X, S) \rangle = \int \Phi(X, I - \nu_\varepsilon \otimes \nu_\varepsilon) \varepsilon |\nabla u_\varepsilon|^2 \, dX, \qquad \forall \Phi \in C_0^\infty(\mathcal{C}_2 \times G(n)).$$

Hutchinson and Tonegawa proved:

1. As $\varepsilon \to 0$, $V_\varepsilon$ converges in the sense of varifolds to a stationary, rectifiable varifold $V$ with integer density (modulo division by the constant $\sigma_0$).
2. The measures $\mu_\varepsilon$ converge to $\|V\|$ weakly.
3. The discrepancy quantity satisfies

$$\frac{1}{\varepsilon} W(u_\varepsilon) - \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 \to 0 \quad \text{in } L^1_{loc}. \tag{4.2}$$

4. For any $t \in (-1, 1)$ fixed, $\{u_\varepsilon = t\}$ converges to $\operatorname{spt} \|V\|$ in the Hausdorff distance.

The last statement implies $0 \in \operatorname{spt}\|V\|$, because $0 \in \{|u_\varepsilon| \le 1 - b\}$ ($b$ as in Theorem 3.3).

With the help of the bound (3.5), we can give a description of the limit varifold.

**Proposition 4.3** (Limit varifold). *The limit measure satisfies* $\|V\| = \sigma_0 \mathcal{H}^n \lfloor_{\mathbb{R}^n}$.

*Proof.* By taking the limit in (3.5) and using the integer multiplicity of $V$, we get

$$4^{-n}\|V\|(\mathcal{B}_4) = \sigma_0 \omega_n.$$

On the other hand, the integer multiplicity of $V$ implies

$$\lim_{r \to 0} r^{-n}\|V\| \geq \sigma_0 \omega_n.$$

By the monotonicity formula for stationary varifolds [16, Theorem 6.3.2], we deduce that $V$ is a cone.

Recall that $V$ is a rectifiable, stationary varifold with integer multiplicity. What we have shown says that $V$ has density one at the origin. Hence Allard's $\varepsilon$-regularity theorem implies that spt $\|V\|$ is a smooth hypersurface in a neighborhood of the origin.

Combining the cone property with this smooth regularity, we see that $V$ is the standard varifold associated to a hyperplane with unit density. □

Now we show that away from $\mathbb{R}^n$, $u_\varepsilon$ is exponentially close to $\pm 1$.

**Proposition 4.4.** *For any* $h > 0$*, if* $\varepsilon$ *is sufficiently small, we have*

$$(1 - u_\varepsilon^2) + |\nabla u_\varepsilon| \leq C(h) e^{-\frac{|x_{n+1}|}{C(h)\varepsilon}} \quad \text{in } \mathcal{C}_2 \setminus \{|x_{n+1}| \leq h\}.$$

*In particular,* $\{u_\varepsilon = 0\} \cap \mathcal{C}_2$ *lies in the* $h$*-neighborhood of* $\mathbb{R}^n \cap \mathcal{C}_2$.

*Proof.* By [15], $u_\varepsilon^2$ converges to 1 uniformly on any compact set outside spt $\|V\| = \mathbb{R}^n$. In particular, for all $\varepsilon$ small,

$$u_\varepsilon^2 \geq \gamma \quad \text{in } \mathcal{C}_3 \setminus \{|x_{n+1}| \geq h/2\}.$$

By a direct calculation, there exists a universal constant $c$ such that

$$\Delta(1 - u_\varepsilon^2) \geq \frac{c}{\varepsilon^2}(1 - u_\varepsilon^2) \quad \text{in } \mathcal{C}_3 \setminus \{|x_{n+1}| \geq h/2\}.$$

Hence we can apply Lemma B.1 to get the exponential decay of $1 - u_\varepsilon^2$ in $\{|x_{n+1}| > h\}$. The estimate for $|\nabla u_\varepsilon|$ follows from standard interior gradient estimates. □

**Remark 4.5.** If $u_\varepsilon$ converges to 1 (or $-1$) on both sides of $\mathbb{R}^n$, the multiplicity of $V$ will be greater than 1 (see [15, Theorem 1, (4)]). This contradicts Proposition 4.3.

Thus $u_\varepsilon$ converges to 1 locally uniformly on one side of $\{x_{n+1} = 0\}$, say in $\mathcal{C}_2 \cap \{x_{n+1} > 0\}$, and to $-1$ locally uniformly in $\mathcal{C}_2 \cap \{x_{n+1} < 0\}$. Together with the previous proposition, this implies

$$\mathrm{dist}_H(\{u_\varepsilon = 0\} \cap \mathcal{C}_1, \mathbb{R}^n \cap \mathcal{C}_1) \to 0.$$

The following lemma says that (3.6) is a consequence of (3.5).

**Lemma 4.6.** *Let $u_\varepsilon$ be a sequence of solutions satisfying (3.5) and the Modica inequality (2.4) in $\mathcal{B}_4$. Then the excess with respect to $\mathbb{R}^n$ satisfies*

$$\lim_{\varepsilon\to 0} E(2;0,u_\varepsilon)=0.$$

*Proof.* For any $\eta\in C_0^\infty(\mathcal{C}_2)$, take the vector field $Y=(0,\dots,0,\eta x_{n+1})$ and substitute it into the stationary condition (2.6). This leads to

$$0=\int_{\mathcal{C}_2}\left(\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right)\left(\eta+\frac{\partial\eta}{\partial x_{n+1}}x_{n+1}\right)\right.$$
$$\left.-\eta\nu_{\varepsilon,n+1}^2\varepsilon|\nabla u_\varepsilon|^2-x_{n+1}\sum_{i=1}^{n+1}\frac{\partial\eta}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2\right). \tag{4.3}$$

By (4.2) and our assumptions on $u_\varepsilon$, both the measures

$$\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right)dX,\qquad \nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2dX$$

converge to some measures supported on $\mathbb{R}^n$. Thus

$$\lim_{\varepsilon\to 0}\int_{\mathcal{C}_2}\left(\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right)\frac{\partial\eta}{\partial x_{n+1}}x_{n+1}-x_{n+1}\sum_{i=1}^{n+1}\frac{\partial\eta}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2\right)=0.$$

Inserting this into (4.3) and applying the Modica inequality (2.4) finishes the proof. □

**Remark 4.7.** Although we will not use the Caccioppoli type inequality explicitly, here we show how to use the stationary condition (2.6) to derive it.

Take a $\psi\in C_0^\infty((-1,1))$ satisfying $0\le\psi\le 1$, $\psi\equiv 1$ in $(-1/2,1/2)$, $|\psi'|\le 3$. For any $\phi\in C_0^\infty(B_1)$, take $\eta(x,x_{n+1})=\phi(x)^2\psi(x_{n+1})^2$ and replace $x_{n+1}$ by $x_{n+1}-\lambda$ in (4.3), where $\lambda\in(-1,1)$ is an arbitrary constant. By this choice we get

$$0=\int_{\mathcal{C}_1}\left(\left[\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right][\phi^2\psi^2+2\phi^2\psi\psi'(x_{n+1}-\lambda)]\right.$$
$$-\phi^2\psi^2\nu_{\varepsilon,n+1}^2\varepsilon|\nabla u_\varepsilon|^2-(x_{n+1}-\lambda)\sum_{i=1}^{n}2\phi\psi^2\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2$$
$$\left.-(x_{n+1}-\lambda)2\phi^2\psi\psi'\nu_{\varepsilon,n+1}^2\varepsilon|\nabla u_\varepsilon|^2\right). \tag{4.4}$$

First consider those terms containing $\psi'(x_{n+1})$. Since $\psi'\equiv 0$ in $B_1\times\{|x_{n+1}|<1/2\}$, with the help of Proposition 4.4 we get

$$\int_{\mathcal{C}_1}\left(\left[\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right]2\phi^2\psi\psi'(x_{n+1}-\lambda)-(x_{n+1}-\lambda)2\phi^2\psi\psi'\nu_{\varepsilon,n+1}^2\varepsilon|\nabla u_\varepsilon|^2\right)$$
$$=O(e^{-c\varepsilon^{-1}}).$$

Substituting this into (4.4) leads to

$$\int_{\mathcal{C}_1}\left[\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2-\nu_{\varepsilon,n+1}^2\varepsilon|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right]\phi^2\psi^2 = \int_{\mathcal{C}_1}(x_{n+1}-\lambda)\sum_{i=1}^n 2\phi\psi^2\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2+O(e^{-c/\varepsilon}). \tag{4.5}$$

By the Cauchy inequality,

$$\int_{\mathcal{C}_1}(x_{n+1}-\lambda)\sum_{i=1}^n 2\phi\psi^2\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2 \le \frac{1}{4}\int_{\mathcal{C}_1}\phi^2\psi^2\sum_{i=1}^n\nu_{\varepsilon,i}^2\varepsilon|\nabla u_\varepsilon|^2+64\int_{\mathcal{C}_1}|\nabla\phi|^2\psi^2\,(x_{n+1}-\lambda)^2\varepsilon|\nabla u_\varepsilon|^2.$$

Substituting this into (4.5), by noting that

$$\sum_{i=1}^n\nu_{\varepsilon,i}^2=1-(\nu_\varepsilon\cdot e_{n+1})^2,$$

and

$$\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2-\nu_{\varepsilon,n+1}^2\varepsilon|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\ge[1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2,$$

we obtain the following Caccioppoli type inequality:

$$\int_{\mathcal{C}_1}\phi^2\psi^2[1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2\le 2^8\int_{\mathcal{C}_1}|\nabla\phi|^2\psi^2\,(x_{n+1}-\lambda)^2\varepsilon|\nabla u_\varepsilon|^2+Ce^{-c\varepsilon^{-1}}. \tag{4.6}$$

**Remark 4.8.** Since we only have control on $1-(\nu_\varepsilon\cdot e_{n+1})^2$, in view of Remark 3.2, $\nu_\varepsilon$ may be close to $e_{n+1}$ or $-e_{n+1}$. To exclude one of these two possibilities, we need to use the unit density assumption (3.5). A subtle point here is that, without such an assumption, we cannot say that $1-\nu_\varepsilon\cdot e_{n+1}$ (or $1+\nu_\varepsilon\cdot e_{n+1}$) is small everywhere. This is related to the possible interface foliation (and consequently the higher multiplicity of the limit varifold $V$)—see the examples constructed by del Pino–Kowalczyk–Wei–Yang [9].

## 5. Lipschitz approximation

Let

$$f_\varepsilon(x)=\int_{-1}^1[1-(\nu_\varepsilon(x,x_{n+1})\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon(x,x_{n+1})|^2\,dx_{n+1}.$$

By Lemma 4.6, $f_\varepsilon\to 0$ in $L^1(B_1)$. Consider the Hardy–Littlewood maximal function

$$Mf_\varepsilon(x):=\sup_{r\in(0,1)}r^{-n}\int_{B_r(x)}f_\varepsilon(y)\,dy.$$

For any $l>0$, by the weak $L^1$ estimate, there exists a universal constant $C$ such that

$$\mathcal{H}^n(\{Mf_\varepsilon\ge l\}\cap B_1)\le\frac{C}{l}\|f_\varepsilon\|_{L^1(B_1)}=C\delta_\varepsilon^2/l. \tag{5.1}$$

Denote the set $B_1 \setminus \{Mf_\varepsilon \geq l\}$ by $W_\varepsilon$. (Its dependence on the constant $l$ will not be indicated explicitly.) Note that since the integrand in the definition of $f_\varepsilon$ and hence $f_\varepsilon(x)$ are continuous functions, $W_\varepsilon$ is an open set.

Given $b \in (0,1)$ and $l > 0$, we say a point $X \in \{|u_\varepsilon| < 1-b\} \cap \mathcal{C}_1$ is *good* if

$$\sup_{0<r<1} r^{-n} \int_{\mathcal{B}_r(X)} [1-(\nu_\varepsilon \cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2 < l.$$

The good points form a set $A_\varepsilon$ and we let $B_\varepsilon = (\{|u_\varepsilon| < 1-b\} \cap \mathcal{C}_1) \setminus A_\varepsilon$ be the set of *bad* points. Note that since $[1-(\nu_\varepsilon(x,x_{n+1}) \cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon(x,x_{n+1})|^2$ is continuous, $A_\varepsilon$ is an open set and $B_\varepsilon$ is relatively closed in $\{|u_\varepsilon| < 1-b\} \cap \mathcal{C}_1$. Clearly $W_\varepsilon \subset \Pi(A_\varepsilon)$.

Similar to the weak $L^1$ estimate for the Hardy–Littlewood maximal function, $B_\varepsilon$ is small in the following sense.

**Lemma 5.1.** *There exists a universal constant $C$ such that*

$$\mu_\varepsilon(B_\varepsilon) \leq C\delta_\varepsilon^2/l.$$

*Proof.* For any $X \in B_\varepsilon$, by definition there exists an $r_X \in (0,1)$ satisfying

$$r_X^n \leq \frac{1}{l}\int_{\mathcal{B}_{r_X}(X)} [1-(\nu_\varepsilon \cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2.$$

By the Vitali covering lemma, choose a countable set of $X_i \in B_\varepsilon$ such that $\mathcal{B}_{r_i}(X_i)$ (here $r_i := r_{X_i}$) are disjoint, and

$$B_\varepsilon \subset \bigcup_i \mathcal{B}_{5r_i}(X_i).$$

Then

$$\begin{aligned}\mu_\varepsilon(B_\varepsilon) &\leq \sum_i \mu_\varepsilon(\mathcal{B}_{5r_i}(X_i)) \leq C\sum_i r_i^n \quad \text{(by (4.1))}\\ &\leq Cl^{-1}\int_{\mathcal{B}_{r_i}(X_i)} [1-(\nu_\varepsilon \cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2\,dX \leq Cl^{-1}\delta_\varepsilon^2.\end{aligned}$$

□

Another fact about $B_\varepsilon$ is

**Lemma 5.2.** *There exists a universal constant $C$ such that $\mathcal{H}^n(\Pi(B_\varepsilon)) \leq Cl^{-1}\delta_\varepsilon^2$.*

*Proof.* This is because $\Pi(B_\varepsilon) \subset B_1 \setminus W_\varepsilon$. Hence (5.1) applies. □

Next we show that in $A_\varepsilon$, level sets of $u_\varepsilon$ are essentially Lipschitz graphs.

**Lemma 5.3.** *Given $b \in (0,1)$, if $l$ is small enough, then for any $t \in (-1+b, 1-b)$, the set $\{u_\varepsilon = t\} \cap A_\varepsilon$ can be locally represented by a Lipschtz graph $\{x_{n+1} = h_\varepsilon^t(x)\}$. The Lipschitz constant of $h_\varepsilon^t$ is controlled by a constant $c_0(b,l)$ depending on $b$ and $l$, which satisfies* $\lim_{l\to 0} c_0(b,l) = 0$.

*Proof.* Fix $X_0 \in A_\varepsilon$ with $u_\varepsilon(X_0) = t$. After rescaling

$$v(X) = u_\varepsilon(X_0 + \varepsilon X),$$

we are in the situation that

$$\Delta v = W'(v) \quad \text{in } \mathcal{B}_{\varepsilon^{-1}}, \tag{5.2}$$

$$\int_{\mathcal{B}_R(0)} \left( \frac{1}{2} |\nabla v|^2 + W(v) \right) \leq C R^n, \qquad \forall R \in (0, \varepsilon^{-1}), \tag{5.3}$$

$$\int_{\mathcal{B}_1(0)} \sum_{i=1}^{n} \left( \frac{\partial v}{\partial x_i} \right)^2 \leq l. \tag{5.4}$$

We claim that there exists an $l_0$ small such that for all $l \leq l_0$, there exist two constants $c_1(b, l) \in (0, 1/2)$ and $c_2(b)$ such that

$$\left| \frac{\partial v}{\partial x_{n+1}} \right| \geq (1 - c_1(b, l)) |\nabla v| \geq c_2(b) \quad \text{in } \mathcal{B}_1. \tag{5.5}$$

Assume to the contrary that there exists a sequence of $v_i$ satisfying all the conditions (5.2)–(5.4) with $l$ replaced by $l_i$, which goes to 0 as $i \to 0$. By standard elliptic estimates and the Arzelà–Ascoli theorem, $v_i$ converges to a function $v$ in $C^2_{\text{loc}}(\mathbb{R}^{n+1})$. Now, $v$ is still a solution of (5.2) in $\mathbb{R}^{n+1}$. Because $|v| \leq 1$ and

$$|v(0)| = \lim_{i \to \infty} |v_i(0)| \leq 1 - b,$$

by the strong maximum principle, $|v| < 1$ in $\mathbb{R}^{n+1}$. After passing to the limit in (5.4) (where $l$ is replaced by $l_i$) and (5.3), we see $v(X) \equiv g(x_{n+1} + t)$ for some $t \in \mathbb{R}$. (For more details, see the proof of Lemma B.2.) Then by (2.8),

$$\left| \frac{\partial v}{\partial x_{n+1}}(X) \right| = |\nabla v(X)| = \sqrt{2W(v(X))} \geq c(b) \qquad \text{in } \mathcal{B}_1.$$

Thus for all $i$ large, $v_i$ satisfies (5.5). This also implies that $c_1(b, l)$ converges to 0 as $l \to 0$.

By (5.5), the level set $\{v = v(0)\} \cap \mathcal{B}_1(0)$ is locally a Lipschitz graph of the form $\{x_{n+1} = h(x)\}$, with Lipschitz constant $c_0(b, l) \leq 2c_1(b, l)$. Coming back to $u_\varepsilon$ we finish the proof. □

By Lemma B.2 and [15, Proposition 5.6], for any $L > 0$ and $X = (x, x_{n+1}) \in A_\varepsilon$, if we have chosen $l$ sufficiently small, then

$$\Pi^{-1}(x) \cap \{u_\varepsilon = u_\varepsilon(X)\} \cap \mathcal{B}_{L\varepsilon}(X) = \{X\}. \tag{5.6}$$

The above results only provide a clear picture of $\{u_\varepsilon = t\} \cap A_\varepsilon$ at $O(\varepsilon)$ scales. By the unit density assumption (3.5), we can further claim:

**Lemma 5.4.** *Given $b \in (0, 1)$, for every $t \in (-1+b, 1-b)$ and $x \in \Pi(A_\varepsilon)$, in $\Pi^{-1}(x) \cap \{u_\varepsilon = t\}$ there exists exactly one point, $(x, h^t_\varepsilon(x))$ (as in Lemma 5.3). Moreover, $h^t_\varepsilon$ is Lipschitz on $\Pi(A_\varepsilon)$.*

This lemma is a consequence of the following lemma, provided that we have chosen first $R_0$ large in the following lemma and then $l$ sufficiently small in the definition of $A_\varepsilon$. The proof of Lemma 5.4 will be completed after Remark 5.6.

**Lemma 5.5.** *For any $b \in (0,1)$ and $\delta > 0$, there exist three constants, $R_0$ large and $\tau_1, l_2$ small, such that the following holds. Suppose that $u_\varepsilon$ is a solution of (2.2) in $\mathcal{B}_{R_0}$, where $\varepsilon \le 1$, satisfying $|u_\varepsilon(0)| \le 1-b$, the Modica inequality (2.4), and*

$$R_0^{-n}\int_{\mathcal{B}_{R_0}}\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right)\le(1+\tau_1)\sigma_0\omega_n,$$

$$R_0^{-n}\int_{\mathcal{B}_{R_0}}[1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2\le l_2.$$

*Then $\{u_\varepsilon=u_\varepsilon(0)\}\cap\mathcal{B}_1$ is contained in the $\delta$-neighborhood of $\mathbb{R}^n\cap\mathcal{B}_1$.*

This result can be seen as a quantitative version of the multiplicity one property for the limit varifold $V$.

*Proof of Lemma 5.5.* Assume that there exists a sequence of solutions $u_k$ satisfying the assumptions in this lemma, with $\varepsilon$ replaced by $\varepsilon_k \in (0,1]$,

$$R_0^{-n}\int_{\mathcal{B}_{R_0}}\left(\frac{\varepsilon_k}{2}|\nabla u_k|^2+\frac{1}{\varepsilon_k}W(u_{\varepsilon_k})\right)\le(1+\tau_k)\sigma_0\omega_n, \tag{5.7}$$

where $\tau_k\to 0$, and

$$R_0^{-n}\int_{\mathcal{B}_{R_0}}\varepsilon_k\sum_{i=1}^n\left(\frac{\partial u_k}{\partial x_i}\right)^2\to 0, \tag{5.8}$$

but there exists $X_k=(x_k,x_{n+1,k})\in\{u_k=u_k(0)\}\cap\mathcal{B}_1$ with $|x_{n+1,k}|\ge\delta$. The constant $R_0$ will be determined below. Without loss of generality, assume that $X_k$ converges to some point $X_\infty=(x_\infty,x_{n+1,\infty})\in\mathcal{B}_1$ with $|x_{n+1,\infty}|\ge\delta$.

The proof is divided into two cases.

**Case 1:** $\varepsilon_k$ converges to some $\varepsilon_0>0$ (after subtracting a subsequence). By standard elliptic estimates and the Arzelà–Ascoli theorem, $u_k$ converges to a function $u_\infty$ in $C^2(\mathcal{B}_{R_0-1})$. Because $|u_k|<1$, $|u_\infty|\le 1$ in $\mathcal{B}_{R_0-1}$. Further, $u_\infty$ is a solution of (2.2) with $\varepsilon$ replaced by $\varepsilon_0$. Since $|u_\infty(0)|\le 1-b<1$, by the strong maximum principle, $|u_\infty|<1$ strictly in $\mathcal{B}_{R_0-1}$. Passing to the limit in (5.7) leads to

$$\int_{\mathcal{B}_{R_0-1}}\left(\frac{\varepsilon_0}{2}|\nabla u_\infty|^2+\frac{1}{\varepsilon_0}W(u_\infty)\right)\le\sigma_0\omega_n(R_0-1)^n. \tag{5.9}$$

Since $\varepsilon_0\le 1$, we cannot have $u_\infty\equiv u_\infty(0)$, because otherwise

$$\int_{\mathcal{B}_{R_0-1}}\left(\frac{\varepsilon_0}{2}|\nabla u_\infty|^2+\frac{1}{\varepsilon_0}W(u_\infty)\right)\ge\frac{1}{\varepsilon_0}W(u_\infty(0))\omega_{n+1}(R_0-1)^{n+1}>\sigma_0\omega_n(R_0-1)^n,$$

if we choose $R_0$ large to fufill the last inequality. (It depends only on $b$, the dimension $n$ and the potential $W$.)

By passing to the limit in (5.8) we obtain

$$R_0^{-n}\int_{\mathcal{B}_{R_0}}\varepsilon_0\sum_{i=1}^{n}\left(\frac{\partial u_\infty}{\partial x_i}\right)^2=0.$$

Thus $u_\infty(X)\equiv\tilde{u}(x_{n+1})$.

Now, $\tilde{u}$ is a one-dimensional solution. By (5.9), we have

$$\int_{-R_0/2}^{R_0/2}\left(\frac{\varepsilon_0}{2}\left|\frac{d\tilde{u}}{dt}\right|^2+\frac{1}{\varepsilon_0}W(\tilde{u})\right)dt\le C\sigma_0, \tag{5.10}$$

for some universal constant $C$ independent of $R_0$. Since $\varepsilon_0\le 1$, we claim that if $R_0$ is sufficiently large, then

$$\frac{\partial u_\infty}{\partial x_{n+1}}(X)\neq 0\quad \text{for } x_{n+1}\in(-1,1). \tag{5.11}$$

This can be proved by a contradiction argument, using the following fact: Except the heteroclinic solution $g$, all the other solutions of (2.1) in $\mathbb{R}^1$ are periodic, hence their energy on $\mathbb{R}$ is infinite.

By (5.11), $u_\infty\neq u_\infty(0)$ in $\mathcal{B}_1\setminus\mathbb{R}^n$. However, by the convergence of $X_k$ and uniform convergence of $u_k$, $u_\infty(X_\infty)=u_\infty(0)$. Because $X_\infty\in\mathcal{B}_1\setminus\mathbb{R}^n$, this is a a contradiction.

**Case 2:** $\varepsilon_k\to 0$. Let $V_k$ be the varifold associated to $u_k$ as defined in Section 4. For any $\eta\in C_0^\infty(\mathcal{B}_{R_0})$, let

$$\Phi(X,S)=\eta(X)\langle Se_{n+1},e_{n+1}\rangle\in C_0^\infty(\mathcal{B}_{R_0}\times G(n)).$$

By (5.8),

$$\langle V_k,\Phi\rangle=\int_{\mathcal{B}_{R_0}}\eta(X)\varepsilon_k\sum_{i=1}^{n}\left(\frac{\partial u_k}{\partial x_i}\right)^2dX\to 0.$$

Let $V_\infty$ be the limit varifold of $V_k$, which is stationary rectifiable with unit density by the Hutchinson–Tonegawa theorem. Then

$$0=\langle V,\Phi\rangle=\int\eta(X)\langle Te_{n+1},e_{n+1}\rangle\,d\|V_\infty\|,$$

where $T$ is the weak tangent plane of $V$ at $X$. Hence $T=\mathbb{R}^n$ $\|V_\infty\|$-a.e. and $V_\infty=\sigma_0\sum_j i(T_j)$ in $B_{R_0/2}\times(-R_0/2,R_0/2)$, where $T_j=\mathbb{R}^n\times\{(0,t_j)\}$ for some $j$ and $i(T_j)$ is the standard varifold associated to it with unit density. By our assumptions, there are at least two components, say $T_0$ and $T_1$, containing the points 0 and $X_\infty$ respectively.

However, passing to the limit in (5.7) gives

$$\|V\|(\mathcal{B}_{R_0})\le\sigma_0\omega_nR_0^n=\sigma_0\|T_0\|(\mathcal{B}_{R_0}).$$

Thus we cannot have any more components other than $T_0$. This also leads to a contradiction. □

**Remark 5.6.** It will be useful to write the dependence of $\delta$ and $l_2$ reversely as $\delta = c_2(l_2)$. This function is a modulus of continuity, i.e. a non-decreasing function satisfying $\lim_{l_2\to 0} c_2(l_2) = 0$.

For any $X_0 = (x_0, x_{0,n+1}) \in A_\varepsilon$ and $r \in (\varepsilon, 1/R_0)$, applying the previous lemma to

$$\tilde{u}_{\varepsilon,r}(X) = u_\varepsilon(X_0 + rX)$$

gives

$$\{u_\varepsilon = u_\varepsilon(X_0)\} \cap (\mathcal{B}_{1/2}(X_0) \setminus \mathcal{B}_\varepsilon(X_0)) \subset \{|x_{n+1} - x_{0,n+1}| \le c_2(l)|x - x_0|\}. \quad (5.12)$$

Together with (5.6), this implies that for every $t \in (-1+b, 1-b)$ and $x \in \Pi(A_\varepsilon)$, there exists at most one point in $\Pi^{-1}(x) \cap \{u_\varepsilon = t\}$.

On the other hand, by Remark 4.5, for each $x \in B_1$,

$$u_\varepsilon(x, 1) > 1 - b, \qquad u_\varepsilon(x, -1) < -1 + b.$$

Thus, by continuity of $u_\varepsilon$, there must exist one $x_{n+1} \in (-1, 1)$ satisfying $u_\varepsilon(x, x_{n+1}) = t$.

In conclusion, for any $x \in \Pi(A_\varepsilon)$, there exists a unique point $(x, x_{n+1}) \in \Pi^{-1}(x) \cap \{u_\varepsilon = t\}$. Combining (5.6) with (5.12), it can be seen that $h^t_\varepsilon$ is Lipschitz on $\Pi(A_\varepsilon)$. This completes the proof of Lemma 5.4.

Recall that we have assumed $u_\varepsilon > 0$ in $\mathcal{C}_1 \cap \{x_{n+1} > h\}$ and $u_\varepsilon < 0$ in $\mathcal{C}_1 \cap \{x_{n+1} < -h\}$ for some $h > 0$—see Remark 4.5. (This $h$ can be made arbitrarily small as $\varepsilon \to 0$.) Hence for any $r \in (\varepsilon, 1/R_0)$, by continuous dependence on $r$, (5.12) can be improved to

$$\{x_{n+1} - x_{0,n+1} > c_2(l)|x - x_0|\} \cap (\mathcal{B}_{1/2}(X_0) \setminus \mathcal{B}_\varepsilon(X_0)) \subset \{u_\varepsilon > u_\varepsilon(X_0)\}. \quad (5.13)$$

When $r = \varepsilon$, combining this with Lemma 5.6, we obtain

$$\frac{\partial u_\varepsilon}{\partial x_{n+1}}(X) \ge (1 - c_1(b, l))|\nabla u_\varepsilon(X)| \ge \frac{c(b)}{\varepsilon}, \qquad \forall X \in A_\varepsilon. \quad (5.14)$$

In the following, for $t \in (-1+b, 1-b)$, we denote the Lipschitz functions by $h^t_\varepsilon$. By (5.6), the definition domains of $h^t_\varepsilon$ can be made to be a common one, $\Pi(A_\varepsilon)$. By (5.14), $h^t_\varepsilon$ is strictly increasing in $t \in (-1+b, 1-b)$.

In the above construction, $h^t_\varepsilon$ is only defined on a subset of $B_1$, but we can extend it to $B_1$ without increasing its Lipschitz constant by letting (see for example [16, Theoerm 3.1.3])

$$h^t_\varepsilon(x) := \inf_{y\in\Pi(A_\varepsilon)} (h^t_\varepsilon(y) + c_3(b, l)|y - x|), \qquad \forall x \in B_1. \quad (5.15)$$

Here $c_3(b, l) = \max\{c_0(b, l), c_2(l)\}$. This extension preserves the monotonicity of $h^t_\varepsilon$ in $t$.

In Sections 7 and 8, $b$ and hence $l$ may be decreased further. Thus it is worth noting the dependence of these Lipschitz functions on $b$ and $l$.

**Remark 5.7.** If $l$ decreases, the definition domain of $h^t_\varepsilon$ also decreases. But on the common part, these two constructions give the same function. If we define two families by choosing two $0 < b_1 < b_2 < 1$, these two families also coincide on $(-1 + b_2, 1 - b_2)$.

Notation: $D_\varepsilon = \Pi(A_\varepsilon)$.

In the following it will be useful to keep in mind that, on $\{u_\varepsilon = t\} \cap A_\varepsilon$,

$$\frac{\partial u_\varepsilon}{\partial x_{n+1}} = \left(\frac{\partial h_\varepsilon^t}{\partial t}\right)^{-1}, \qquad \frac{\partial u_\varepsilon}{\partial x_i} = -\left(\frac{\partial h_\varepsilon^t}{\partial t}\right)^{-1}\frac{\partial h_\varepsilon^t}{\partial x_i}, \qquad 1 \le i \le n. \tag{5.16}$$

## 6. Estimates on $h_\varepsilon^t$

First we give an $H^1$ bound.

**Lemma 6.1.** *There exists a constant $C(b)$ independent of $\varepsilon$ such that*

$$\int_{-1+b}^{1-b}\int_{B_1} |\nabla h_\varepsilon^t|^2\,dx\,dt \le C(b)\delta_\varepsilon^2.$$

*Proof.* By Lemma 5.2, $\mathcal{H}^n(B_1 \setminus D_\varepsilon) \le C\delta_\varepsilon^2$. The construction in the previous section implies that $|\nabla h_\varepsilon^t| \le c_3(b,l)$ in $B_1$ for all $t \in (-1+b, 1-b)$. Hence

$$\int_{-1+b}^{1-b}\int_{B_1\setminus D_\varepsilon} |\nabla h_\varepsilon^t|^2\,dx\,dt \le C\delta_\varepsilon^2. \tag{6.1}$$

Next, by noting that $\varepsilon|\nabla u_\varepsilon| \ge c(b)$ on $A_\varepsilon$ (see (5.5)), we have

$$\begin{aligned}
\delta_\varepsilon^2 &\ge \int_{A_\varepsilon} [1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2\,dX \\
&= \int_{-1+b}^{1-b}\left[\int_{\{u_\varepsilon=t\}\cap A_\varepsilon} \big(1-(\nu_\varepsilon\cdot e_{n+1})^2\big)\varepsilon|\nabla u_\varepsilon|\,d\mathcal{H}^n\right]dt \qquad \text{(by the coarea formula)} \\
&\ge c(b)\int_{-1+b}^{1-b}\left(\int_{D_\varepsilon}\left[1-\frac{1}{1+|\nabla h_\varepsilon^t|^2}\right]\sqrt{1+|\nabla h_\varepsilon^t|^2}\,dx\right)dt \qquad \text{(by (5.16))} \\
&\ge c(b)\int_{-1+b}^{1-b}\left(\int_{D_\varepsilon}|\nabla h_\varepsilon^t|^2\,dx\right)dt,
\end{aligned}$$

if we choose $l$ so small that the Lipschitz constants of $h_\varepsilon^t$ satisfy $c_3(b,l) \le 1/2$. □

With this lemma in hand, we first choose a $t_\varepsilon \in (-1+b, 1-b)$ such that

$$\int_{B_1} |\nabla h_\varepsilon^{t_\varepsilon}|^2\,dx \le C(b)\delta_\varepsilon^2,$$

and then take a $\lambda_\varepsilon$ so that the function defined by

$$\bar{h}_\varepsilon := \frac{1}{\delta_\varepsilon}h_\varepsilon^{t_\varepsilon} - \lambda_\varepsilon \tag{6.2}$$

satisfies $\int_{B_1}\bar{h}_\varepsilon = 0$. By this choice and the Poincaré inequality,

$$\int_{B_1} \bar{h}_\varepsilon(x)^2\,dx \le C\int_{B_1}|\nabla h_\varepsilon(x)|^2\,dx \le C(b). \tag{6.3}$$

Thus we can assume, after passing to a subsequence of $\varepsilon \to 0$, that $\bar{h}_\varepsilon$ converges to a function $\bar{h}$ weakly in $H^1(B_1)$ and strongly in $L^2(B_1)$.

Let

$$\bar{h}^t_\varepsilon := \frac{1}{\delta_\varepsilon} h^t_\varepsilon - \lambda_\varepsilon.$$

In $D_\varepsilon$,

$$0 \le \frac{\partial h^t_\varepsilon}{\partial t} = \left(\frac{\partial u_\varepsilon}{\partial x_{n+1}}\right)^{-1} \le C(b)\varepsilon, \tag{6.4}$$

with a constant $C(b)$ depending only on $b$. Hence

$$0 \le h^{1-b}_\varepsilon - h^{-1+b}_\varepsilon \le C(b)\varepsilon \quad \text{ in } D_\varepsilon. \tag{6.5}$$

This also holds for $x \in B_1 \setminus D_\varepsilon$ by (5.15).

Hence for any $-1 + b < t_1 < t_2 < 1 - b$,

$$\int_{B_1} (h^{t_1}_\varepsilon - h^{t_2}_\varepsilon)^2 \le C(b)\varepsilon^2. \tag{6.6}$$

Because $\delta_\varepsilon \gg \varepsilon$, for any sequence $\tilde{t}_\varepsilon \in (-1+b, 1-b)$, $\bar{h}^{\tilde{t}_\varepsilon}_\varepsilon$ still converges to $\bar{h}$ in $L^2(B_1)$.

Since $\delta_\varepsilon^{-1}\nabla h^t_\varepsilon$ is uniformly bounded in $L^2(B_1 \times (-1 + b, 1 - b), \mathbb{R}^n)$, it can be assumed to converge weakly to some limit in $L^2(B_1 \times (-1+b, 1-b), \mathbb{R}^n)$. By the above discussion, this limit must be $\nabla\bar{h}$.

By Remark 5.7, $\bar{h}$ is independent of the choice of $b$. Hence we have a universal constant $C$, independent of $b$ and $l$, such that

$$\int_{B_1} (|\nabla\bar{h}|^2 + \bar{h}^2) \le C. \tag{6.7}$$

Concerning the size of $\lambda_\varepsilon$, we have

**Lemma 6.2.** $\lim_{\varepsilon\to 0} |\lambda_\varepsilon\delta_\varepsilon| = 0$.

*Proof.* Note that

$$\lambda_\varepsilon\delta_\varepsilon = \int_{B_1} h^{t_\varepsilon}_\varepsilon. \tag{6.8}$$

By Proposition 4.4,

$$\lim_{\varepsilon\to 0} \sup_{\mathcal{C}_1\cap\{|u_\varepsilon|\le 1-b\}} |x_{n+1}| = 0.$$

Thus

$$\lim_{\varepsilon\to 0} \sup_{t\in(-1+b,1-b)} \sup_{x\in D_\varepsilon} |h^t_\varepsilon(x)| = 0. \tag{6.9}$$

For any $x \in B_1 \setminus D_\varepsilon$, by Lemma 5.2,

$$\text{dist}(x, D_\varepsilon) \le Cl^{-1/n}\delta_\varepsilon^{2/n}.$$

Because the Lipschitz constant of $h^t_\varepsilon$ is smaller than $c_3(b, l) \le 1$, we obtain

$$\sup_{t\in(-1+b,1-b)} \sup_{x\in B_1\setminus D_\varepsilon} |h^t_\varepsilon(x)| \le \sup_{t\in(-1+b,1-b)} \sup_{x\in D_\varepsilon} |h^t_\varepsilon(x)| + C(l)\delta_\varepsilon^{2/n}.$$

Combining this with (6.9) we see

$$\lim_{\varepsilon\to 0}\sup_{t\in(-1+b,1-b)}\sup_{x\in B_1}|h_\varepsilon^t(x)|=0.$$

Substituting this into (6.8) we finish the proof. □

Next we establish a bound for the height excess

$$\int_{\mathcal{C}_{3/4}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2.$$

This can be viewed as a Poincaré inequality on the varifold $V_\varepsilon$. (The Caccioppoli type inequality in Remark 4.7 is a reverse Poincaré inequality.)

**Lemma 6.3.** *There exists a universal constant $C$ such that*

$$\int_{\mathcal{C}_{3/4}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\leq C\delta_\varepsilon^2. \tag{6.10}$$

*Proof.* The proof is divided into two steps. In the following we shall fix two numbers $0<b_2<b_1<1$ so that $W''\geq\kappa$ in $(-1,-1+b_1)\cup(1-b_1,1)$.

**Step 1.** Here we give an estimate in the part $\{|u_\varepsilon|<1-b_2\}\cap\mathcal{C}_1$:

$$\int_{\{|u_\varepsilon|<1-b_2\}\cap\mathcal{C}_1}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\leq C\delta_\varepsilon^2. \tag{6.11}$$

First, by (6.2) and (6.6),

$$\int_{-1+b_2}^{1-b_2}\int_{B_1}(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)^2\,dx\,dt\leq C\delta_\varepsilon^2. \tag{6.12}$$

Then by a change of variable, the gradient bound (2.5) and the Lipschitz bound on $h_\varepsilon^t$, we obtain

$$\begin{aligned}\int_{A_\varepsilon}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2&=\int_{-1+b_2}^{1-b_2}\int_{D_\varepsilon}(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)^2(1+|\nabla h_\varepsilon^t|^2)\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}\,dx\,dt\\&\leq C\int_{-1+b_2}^{1-b_2}\int_{D_\varepsilon}(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)^2\,dx\,dt\leq C\delta_\varepsilon^2.\end{aligned} \tag{6.13}$$

In $B_\varepsilon$, by Lemmas 5.1 and 6.2,

$$\int_{B_\varepsilon}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\leq C\Big[\sup_{\{|u_\varepsilon|<1-b\}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\Big]\mu_\varepsilon(B_\varepsilon)\leq C\delta_\varepsilon^2. \tag{6.14}$$

Combining (6.13) and (6.14) we get (6.11).

**Step 2.** We claim that in $\{|u_\varepsilon|>1-b_2\}\cap\mathcal{C}_{3/4}$,

$$\int_{\{|u_\varepsilon|>1-b_2\}\cap\mathcal{C}_{3/4}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\leq C\delta_\varepsilon^2. \tag{6.15}$$

Choose a function $\zeta \in C^\infty(\mathbb{R})$ satisfying

$$\begin{cases}\zeta(t) \equiv 1 & \text{in } \{|t| > 1 - b_2\},\\ \zeta(t) \equiv 0 & \text{in } \{|t| < 1 - b_1\},\\ |\zeta'| \le \dfrac{2}{b_1 - b_2}, \quad |\zeta''| \le \dfrac{8}{(b_1 - b_2)^2} & \text{in } \{1 - b_1 \le |t| \le 1 - b_2\}.\end{cases}$$

We also fix $\eta \in C_0^\infty(B_1 \times \{|x_{n+1}| < 4/3\})$ such that $0 \le \eta \le 1$ and $\eta \equiv 1$ in $\mathcal{C}_{3/4}$.

It can be directly checked that

$$\Delta(\varepsilon|\nabla u_\varepsilon|^2) \ge \frac{\kappa}{\varepsilon^2}(\varepsilon|\nabla u_\varepsilon|^2) \quad \text{in } \{|u_\varepsilon| > 1 - b_1\}. \tag{6.16}$$

Multiplying this equation by $(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)^2\eta\zeta(u_\varepsilon)$ and integrating by parts, we obtain

$$\begin{aligned}\int_{\mathcal{B}_2} &(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)^2\eta\zeta(u_\varepsilon)\varepsilon|\nabla u_\varepsilon|^2\\ &\le \frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2} \Delta[(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)^2\eta]\zeta(u_\varepsilon)\varepsilon|\nabla u_\varepsilon|^2\\ &\quad + \frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2} 4(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)\frac{\partial u_\varepsilon}{\partial x_{n+1}}\eta\zeta'(u_\varepsilon)\varepsilon|\nabla u_\varepsilon|^2\\ &\quad + \frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2} 2(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)^2(\nabla\eta\cdot\nabla u_\varepsilon)\zeta'(u_\varepsilon)\varepsilon|\nabla u_\varepsilon|^2\\ &\quad + \frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2} \big(\zeta''(u_\varepsilon)|\nabla u_\varepsilon|^2 + \zeta'(u_\varepsilon)\Delta u_\varepsilon\big)(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)^2\eta\varepsilon|\nabla u_\varepsilon|^2.\end{aligned} \tag{6.17}$$

On the right hand side, the first term is bounded by

$$\frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2} \Delta[(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)^2\eta]\zeta(u_\varepsilon)\varepsilon|\nabla u_\varepsilon|^2 \le C\varepsilon^2, \tag{6.18}$$

because both $\Delta[(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)^2\eta]$ and $\zeta(u_\varepsilon)$ are bounded by a universal constant.

Note that the supports of $\zeta'(u_\varepsilon)$ and $\zeta''(u_\varepsilon)$ are contained in $\{|u_\varepsilon| < 1 - b_2\}$. By the Cauchy inequality, the second term is bounded by

$$\begin{aligned}\frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2} &4(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)\frac{\partial u_\varepsilon}{\partial x_{n+1}}\eta\zeta'(u_\varepsilon)\varepsilon|\nabla u_\varepsilon|^2\\ &\le C\varepsilon\left[\int_{\{|u_\varepsilon|<1-b_2\}\cap\mathcal{B}_2} (x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)^2\eta^2\varepsilon|\nabla u_\varepsilon|^2\right]^{1/2}\\ &\quad \times\left[\int_{\{|u_\varepsilon|<1-b_2\}\cap\mathcal{B}_2}\left(\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}\right)^2\varepsilon|\nabla u_\varepsilon|^2\right]^{1/2}\\ &\le C\varepsilon\delta_\varepsilon.\end{aligned} \tag{6.19}$$

Here we have used Proposition 4.4, (6.11) and the fact that $\varepsilon\left|\frac{\partial u_\varepsilon}{\partial x_{n+1}}\right| \le C$.

Similarly, the third term can be controlled as

$$\begin{aligned}\frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2}&(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2(\nabla\eta\cdot\nabla u_\varepsilon)\zeta'(u_\varepsilon)\varepsilon|\nabla u_\varepsilon|^2\\&\leq C(\sup|\nabla\eta|)(\sup\varepsilon|\nabla u_\varepsilon|)\varepsilon\int_{\{|u_\varepsilon|<1-b_2\}\cap(B_1\times(-4/3,4/3))}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\\&\leq C\varepsilon\delta_\varepsilon^2.\end{aligned}\tag{6.20}$$

Finally, in the last term, by employing (2.5), we obtain

$$\begin{aligned}\frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2}&\big(\zeta''(u_\varepsilon)|\nabla u_\varepsilon|^2+\zeta'(u_\varepsilon)\Delta u_\varepsilon\big)(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\eta\varepsilon|\nabla u_\varepsilon|^2\\&\leq C\int_{\{|u_\varepsilon|<1-b_2\}\cap\mathcal{B}_2}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\eta\varepsilon|\nabla u_\varepsilon|^2\leq C\delta_\varepsilon^2.\end{aligned}\tag{6.21}$$

Substituting (6.18)–(6.21) into (6.17), and noting that $\delta_\varepsilon \gg \varepsilon$, we obtain (6.15). Combining (6.11) and (6.15) finishes the proof. □

Once we have this bound, we can further sharpen several estimates in the above proof to show that

**Corollary 6.4.** *For any $\sigma > 0$, there exists a constant $b > 0$ such that*

$$\int_{\{|u_\varepsilon|>1-b\}\cap\mathcal{C}_{3/4}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\leq\sigma\delta_\varepsilon^2+C\varepsilon^2.$$

*Proof.* The starting point is the estimate (6.17), where $\xi$ is now assumed to have its support in $(-1+b, 1-b)$, and satisfies $\xi \equiv 1$ in $(-1+2b, 1-2b)$, and $|\zeta'|^2+|\zeta''| \leq 64b^{-2}$. The constant $b$ will be determined later.

We only need to get a better control in (6.21). Estimates in (6.18)–(6.20) will be kept. By using the Cauchy inequality, they can be bounded by $\sigma\delta_\varepsilon^2 + C\varepsilon^2$.

Replace (6.21) by

$$\begin{aligned}\frac{\varepsilon^2}{\kappa}\int_{\mathcal{B}_2}&\big(\zeta''(u_\varepsilon)|\nabla u_\varepsilon|^2+\zeta'(u_\varepsilon)\Delta u_\varepsilon\big)(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\eta\varepsilon|\nabla u_\varepsilon|^2\\&\leq C\int_{\{1-2b<|u_\varepsilon|<1-b\}\cap\mathcal{B}_2}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\eta\varepsilon|\nabla u_\varepsilon|^2.\end{aligned}\tag{6.22}$$

Here the constant $C$ is independent of $b$. This is because, instead of using the bound (2.5) as in the proof of the previous lemma, we can use

$$\varepsilon^2|\nabla u_\varepsilon|^2\leq 2W(u_\varepsilon),\quad \varepsilon^2|\Delta u_\varepsilon|\leq|W'(u_\varepsilon)|,$$

which follow from the Modica inequality and (2.2). Thus $\varepsilon^2(\zeta''(u_\varepsilon)|\nabla u_\varepsilon|^2+\zeta'(u_\varepsilon)\Delta u_\varepsilon)$ is bounded independent of $b \in (0, 1)$.

In view of this, to complete the proof we only need to prove that, for any $\sigma$, there exists a constant $b \in (0,1)$ such that

$$\int_{\{1-2b<|u_\varepsilon|<1-b\}\cap\mathcal{B}_2}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2 \le \sigma\int_{\{|u_\varepsilon|<1-b\}\cap\mathcal{B}_2}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2. \tag{6.23}$$

To this end, first note that, because (see Proposition 4.4)

$$\lim_{\varepsilon\to 0}\sup_{\{|u_\varepsilon|<1-b\}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2=0,$$

(6.14) can be improved to

$$\int_{B_\varepsilon}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2 \le \frac{\sigma}{2}\delta_\varepsilon^2, \qquad \forall\varepsilon \text{ small.} \tag{6.24}$$

Next, by (6.6), (6.13) can be rewritten as

$$\begin{aligned}
&\int_{A_\varepsilon}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\\
&\qquad=\int_{-1+b}^{1-b}\int_{D_\varepsilon}(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)^2(1+|\nabla h_\varepsilon^t|^2)\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}\,dx\,dt\\
&\qquad=\left[\int_{D_\varepsilon}(h_\varepsilon^{t_\varepsilon}-\lambda_\varepsilon\delta_\varepsilon)^2\right]\left[\int_{-1+b}^{1-b}\int_{D_\varepsilon}(1+|\nabla h_\varepsilon^t|^2)\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}\,dx\,dt\right]+O(\varepsilon^2)\\
&\qquad\ge c\left[\int_{D_\varepsilon}(h_\varepsilon^{t_\varepsilon}-\lambda_\varepsilon\delta_\varepsilon)^2\right]+O(\varepsilon^2).
\end{aligned}\tag{6.25}$$

In the last step we have used the fact that $\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}\ge c$ in $A_\varepsilon\cap\{|u_\varepsilon|<1/2\}$.

Now consider the integral on $(1-2b,1-b)$. By noting that $\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}$ is small in $\{|u_\varepsilon|>1-2b\}$ (using the Modica inequality (2.4)), we obtain

$$\begin{aligned}
&\int_{A_\varepsilon\cap\{1-2b<|u_\varepsilon|<1-b\}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\\
&\qquad=\left[\int_{D_\varepsilon}(h_\varepsilon^{t_\varepsilon}-\lambda_\varepsilon\delta_\varepsilon)^2\right]\left[\int_{1-2b<|t|<1-b}\int_{D_\varepsilon}(1+|\nabla h_\varepsilon^t|^2)\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}\,dx\,dt\right]+O(\varepsilon^2)\\
&\qquad=o_b(1)\left[\int_{D_\varepsilon}(h_\varepsilon^{t_\varepsilon}-\lambda_\varepsilon\delta_\varepsilon)^2\right]+O(\varepsilon^2).
\end{aligned}\tag{6.26}$$

Combining (6.25) and (6.26) we get

$$\int_{A_\varepsilon\cap\{1-2b<|u_\varepsilon|<1-b\}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2=o_b(1)\int_{A_\varepsilon}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2+O(\varepsilon^2).$$

With (6.24) this implies (6.23), if we have chosen $b$ small enough. □

Finally, we give a uniform estimate on $\partial h_\varepsilon^t/\partial t$ in a good set.

**Lemma 6.5.** *There exists a set $E_\varepsilon \subset D_\varepsilon$ with $\mathcal{H}^n(D_\varepsilon \setminus E_\varepsilon) \le C\delta_\varepsilon$ such that for any $t \in (-1+b, 1-b)$ and $X_\varepsilon \in \Pi^{-1}(E_\varepsilon) \cap A_\varepsilon$ with $u_\varepsilon(X_\varepsilon) = t$,*

$$\varepsilon\left[\frac{\partial h_\varepsilon^t}{\partial t}(X_\varepsilon)\right]^{-1} = g'(g^{-1}(t)) + o_\varepsilon(1).$$

*Here $o_\varepsilon(1)$ means a quantity converging to 0 as $\varepsilon \to 0$, independent of $X_\varepsilon$ and $t$.*

*Proof.* Let $E_\varepsilon = D_\varepsilon \cap \{Mf_\varepsilon < \delta_\varepsilon\}$. By (5.1),

$$\mathcal{H}^n(D_\varepsilon \setminus E_\varepsilon) \le \mathcal{H}^n(B_1 \setminus \{Mf_\varepsilon \ge \delta_\varepsilon\}) \le C\delta_\varepsilon \to 0.$$

For any $X_\varepsilon \in \Pi^{-1}(E_\varepsilon) \cap A_\varepsilon$, consider

$$v_\varepsilon(X) := u_\varepsilon(X_\varepsilon + \varepsilon X) \quad \text{for } X \in \mathcal{B}_{\varepsilon^{-1}/2}.$$

Then $v_\varepsilon$ is a solution of (2.1). By definition, $v_\varepsilon(0) = u_\varepsilon(X_\varepsilon) = t \in (-1+b, 1-b)$ because $X_\varepsilon \in A_\varepsilon$. As usual assume $v_\varepsilon$ converges to a function $v_\infty$ in $C^2_{loc}(\mathbb{R}^{n+1})$, which is also a solution of (2.1) on $\mathbb{R}^{n+1}$.

By the definition of the Hardy–Littlewood maximal function and our choice of $E_\varepsilon$,

$$\sup_{0<r<\varepsilon^{-1}/2} r^{-n}\int_{\mathcal{B}_r}\sum_{i=1}^n\left(\frac{\partial v_\varepsilon}{\partial x_i}\right)^2 \le \delta_\varepsilon \to 0.$$

After passing to the limit, we see $v_\infty$ depends only on $x_{n+1}$. Then by (4.1), we have the energy bound

$$\int_{\mathcal{B}_r}\left(\frac{1}{2}|\nabla v_\infty|^2 + W(v_\infty)\right) \le 8^n\sigma_0\omega_n r^n, \qquad \forall r > 0.$$

From this we deduce that $v_\infty(X) \equiv g(x_{n+1}+g^{-1}(t))$ (see again the proof of Lemma B.1).

By definition and the $C^1_{loc}$ convergence of $v_\varepsilon$,

$$\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}(X_\varepsilon) = \frac{\partial v_\varepsilon}{\partial x_{n+1}}(0) \to \frac{\partial v_\infty}{\partial x_{n+1}}(0) = g'(g^{-1}(t)).$$

The claim then follows from (5.16). □

## 7. The blow up limit

This section is devoted to proving

**Proposition 7.1.** *$\bar{h}$ is harmonic in $B_1$.*

Fix a $\psi \in C_0^\infty((-1,1))$ such that $0 \le \psi \le 1$, $\psi \equiv 1$ in $(-1/2, 1/2)$ and $|\psi'| \le 4$. For any $\varphi \in C_0^\infty(B_1)$, let $X(x, x_{n+1}) = \varphi(x)\psi(x_{n+1})e_{n+1}$, which is a smooth vector field with compact support in $\mathcal{C}_1$.

To prove Proposition 7.1, we substitute this vector field into the stationary condition (2.6). Roughly speaking, if we view the level set of $u_\varepsilon$ as the graph of a function $h$, because $h$ almost satisfies an elliptic equation, this procedure amounts to multiplying the equation of $h$ by a $C_0^\infty$ function and then integrating by parts, which of course is a standard method in elliptic equation theory.

Note that

$$\begin{aligned}&DX(x,x_{n+1})=\psi(x_{n+1})\nabla\varphi(x)\otimes e_{n+1}+\varphi(x)\psi'(x_{n+1})e_{n+1}\otimes e_{n+1},\\&\operatorname{div}X(x,x_{n+1})=\varphi(x)\psi'(x_{n+1}).\end{aligned}$$

Since $\operatorname{div} X$ vanishes in $B_1\times(-1/2,1/2)$, by Proposition 4.4,

$$\int_{\mathcal{C}_1}\left[\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right]\operatorname{div}X=O(e^{-c/\varepsilon}). \tag{7.1}$$

Similarly,

$$\int_{\mathcal{C}_1}\varphi(x)\psi'(x_{n+1})\varepsilon\left(\frac{\partial u_\varepsilon}{\partial x_{n+1}}\right)^2=O(e^{-c/\varepsilon}). \tag{7.2}$$

Thus from the stationary condition (2.6) we deduce that

$$\int_{\mathcal{C}_1}\varepsilon\psi\left(\sum_{i=1}^n\frac{\partial u_\varepsilon}{\partial x_i}\frac{\partial\varphi}{\partial x_i}\right)\frac{\partial u_\varepsilon}{\partial x_{n+1}}=O(e^{-c/\varepsilon})=o(\delta_\varepsilon), \tag{7.3}$$

where in the last equality we have used the assumption (3.7). First note that

$$\begin{aligned}&\int_{\mathcal{C}_1\cap\{|u_\varepsilon|\geq 1-b\}}\varepsilon\psi\left(\sum_{i=1}^n\frac{\partial u_\varepsilon}{\partial x_i}\frac{\partial\varphi}{\partial x_i}\right)\frac{\partial u_\varepsilon}{\partial x_{n+1}}\,dx\,dx_{n+1}\\&\quad\leq C\Big(\sup_{B_1}|\nabla\varphi|\Big)\left[\int_{\mathcal{C}_1\cap\{|u_\varepsilon|\geq 1-b\}}\varepsilon\left(\sum_{i=1}^n\frac{\partial u_\varepsilon}{\partial x_i}\right)^2\right]^{1/2}\left[\int_{\mathcal{C}_1\cap\{|u_\varepsilon|\geq 1-b\}}\varepsilon\left(\frac{\partial u_\varepsilon}{\partial x_{n+1}}\right)^2\right]^{1/2}\\&\quad\leq C(\varphi)o_b(1)\delta_\varepsilon,\end{aligned} \tag{7.4}$$

where $o_b(1)$ converges to 0 as $b\to 0$ (by Lemma B.3).

Next in $B_\varepsilon$,

$$\begin{aligned}&\int_{B_\varepsilon}\varepsilon\psi\left(\sum_{i=1}^n\frac{\partial u_\varepsilon}{\partial x_i}\frac{\partial\varphi}{\partial x_i}\right)\frac{\partial u_\varepsilon}{\partial x_{n+1}}\,dx\,dx_{n+1}\\&\qquad\leq C\Big(\sup_{B_1}|\nabla\varphi|\Big)\left[\int_{B_\varepsilon}\varepsilon\left(\sum_{i=1}^n\frac{\partial u_\varepsilon}{\partial x_i}\right)^2\right]^{1/2}\left[\int_{B_\varepsilon}\varepsilon\left(\frac{\partial u_\varepsilon}{\partial x_{n+1}}\right)^2\right]^{1/2}\\&\qquad\leq C(\varphi)\delta_\varepsilon\mu_\varepsilon(B_\varepsilon)^{1/2}\leq C(\varphi)\delta_\varepsilon^2\quad\text{(by Lemma 5.1).}\end{aligned} \tag{7.5}$$

Substituting (7.4) and (7.5) into (7.3) and noting that $\psi\equiv 1$ on $A_\varepsilon$ (recall that $A_\varepsilon\subset B_1\times\{|x_{n+1}|<1/2\}$), we see that

$$\int_{A_\varepsilon}\varepsilon\left(\sum_{i=1}^n\frac{\partial u_\varepsilon}{\partial x_i}\frac{\partial\varphi}{\partial x_i}\right)\frac{\partial u_\varepsilon}{\partial x_{n+1}}\,dx\,dx_{n+1}=o(\delta_\varepsilon)+o_b(1)\delta_\varepsilon. \tag{7.6}$$

By using the transformation $(x, x_{n+1}) = (x, h^t_\varepsilon(x))$ and (5.16), this integral can be transformed into

$$\int_{-1+b}^{1-b}\int_{D_\varepsilon} \varepsilon\left(\frac{\partial h^t_\varepsilon}{\partial t}\right)^{-1}\sum_{i=1}^n \frac{\partial h^t_\varepsilon}{\partial x_i}\frac{\partial \varphi}{\partial x_i}\,dx\,dt = o(\delta_\varepsilon) + o_b(1)\delta_\varepsilon. \tag{7.7}$$

We need to further divide $D_\varepsilon$ into two parts, using the set $E_\varepsilon$ introduced in Lemma 6.5.

In the first part $D_\varepsilon \setminus E_\varepsilon$, by (2.5), (5.16), and Lemmas 6.1 and 6.5,

$$\begin{aligned}
&\int_{-1+b}^{1-b}\int_{D_\varepsilon\setminus E_\varepsilon} \varepsilon\left(\frac{\partial h^t_\varepsilon}{\partial x_{n+1}}\right)^{-1}\sum_{i=1}^n \frac{\partial h^t_\varepsilon}{\partial x_i}\frac{\partial \varphi}{\partial x_i}\,dx\,dt\\
&\quad ] \le \left(\sup_{A_\varepsilon}\left|\varepsilon\left(\frac{\partial h^t_\varepsilon}{\partial t}\right)^{-1}\right|\right)\left(\sup_{B_1}|\nabla\varphi|\right)\\
&\qquad \times\left[\int_{-1+b}^{1-b}\int_{D_\varepsilon\setminus E_\varepsilon}\sum_{i=1}^n\left(\frac{\partial h^t_\varepsilon}{\partial x_i}\right)^2 dx\,dt\right]^{1/2}\mathcal{H}^n(D_\varepsilon\setminus E_\varepsilon)^{1/2}\\
&\quad \le C(\varphi)\delta_\varepsilon^{3/2} = o(\delta_\varepsilon).
\end{aligned}$$

In $E_\varepsilon$,

$$\begin{aligned}
&\int_{-1+b}^{1-b}\int_{E_\varepsilon} \varepsilon\left(\frac{\partial h^t_\varepsilon}{\partial t}\right)^{-1}\sum_{i=1}^n \frac{\partial h^t_\varepsilon}{\partial x_i}\frac{\partial \varphi}{\partial x_i}\,dx\,dt\\
&\qquad = \int_{-1+b}^{1-b}\int_{E_\varepsilon} g'(g^{-1}(t))\sum_{i=1}^n \frac{\partial h^t_\varepsilon}{\partial x_i}\frac{\partial \varphi}{\partial x_i}\,dx\,dt + o(\delta_\varepsilon),
\end{aligned}$$

where in the last equality we have used

$$\sup_{E_\varepsilon}\left|\varepsilon\left(\frac{\partial h^t_\varepsilon}{\partial t}\right)^{-1} - g'(g^{-1}(t))\right| = o_\varepsilon(1) \to 0,$$

and the bound

$$\begin{aligned}
&\int_{-1+b}^{1-b}\int_{E_\varepsilon}\sum_{i=1}^n \frac{\partial h^t_\varepsilon}{\partial x_i}\frac{\partial \varphi}{\partial x_i}\,dx\,dt\\
&\qquad \le \left[\int_{-1+b}^{1-b}\int_{B_1}|\nabla h^t_\varepsilon|^2\,dx\,dt\right]^{1/2}\left[\int_{-1+b}^{1-b}\int_{B_1}|\nabla\varphi|^2\,dx\,dt\right]^{1/2}\\
&\qquad \le C(\varphi)\delta_\varepsilon \quad \text{(by Lemma 6.1)}.
\end{aligned}$$

By the Cauchy inequality we also have

$$\begin{aligned}
&\int_{-1+b}^{1-b}\int_{B_1\setminus E_\varepsilon} g'(g^{-1}(t))\sum_{i=1}^n \frac{\partial h^t_\varepsilon}{\partial x_i}\frac{\partial \varphi}{\partial x_i}\,dx\,dt\\
&\quad \le C\left(\sup_{B_1}|\nabla\varphi|\right)\left[\int_{-1+b}^{1-b}\int_{B_1}|\nabla h^t_\varepsilon|^2\,dx\,dt\right]^{1/2}\mathcal{H}^n(B_1\setminus E_\varepsilon)^{1/2} \le C(\varphi)\delta_\varepsilon^{3/2} = o(\delta_\varepsilon),
\end{aligned}$$

where we have used Lemmas 6.1, 5.2 and 6.5.

Putting these three integrals together, by (7.7) we get

$$\int_{-1+b}^{1-b}\int_{B_1} g'(g^{-1}(t))\sum_{i=1}^{n}\frac{\partial h_\varepsilon^t}{\partial x_i}\frac{\partial\varphi}{\partial x_i}\,dx\,dt = o_b(1)\delta_\varepsilon + o(\delta_\varepsilon).$$

By the weak convergence of $\delta_\varepsilon^{-1}\nabla h_\varepsilon^t$ to $\nabla\bar{h}$ in $L^2(B_1\times(-1+b,1-b))$, we can let $\varepsilon\to 0$ to obtain

$$\left[\int_{-1+b}^{1-b} g'(g^{-1}(t))\,dt\right]\left[\int_{B_1}\sum_{i=1}^{n}\frac{\partial\bar{h}}{\partial x_i}\frac{\partial\varphi}{\partial x_i}\,dx\right] = o_b(1). \tag{7.8}$$

For $b\in(0,1/2)$,

$$\int_{-1+b}^{1-b} g'(g^{-1}(t))\,dt = \int_{g^{-1}(-1+b)}^{g^{-1}(1-b)} g'(s)^2\,ds \geq c\sigma_0.$$

At the first step, we can choose a smaller $\tilde{b}$ and get another family $\tilde{h}_\varepsilon^t$ for $t\in(-1+\tilde{b},1-\tilde{b})$. Assume its limit is $\tilde{h}$. By Remark 5.7, $\tilde{h}_\varepsilon^t = h_\varepsilon^t$ for $t\in(-1+b,1-b)$. Then by (6.6), $\tilde{h}=\bar{h}$. In other words, the limit $\bar{h}$ does not depend on $b$.

After taking $b\to 0$ in (7.8), we get

$$\int_{B_1}\sum_{i=1}^{n}\frac{\partial\bar{h}}{\partial x_i}\frac{\partial\varphi}{\partial x_i}\,dx = 0.$$

Since $\varphi\in C_0^\infty(B_1)$ can be arbitrary and $\bar{h}\in H^1(B_1)$, standard harmonic function theory implies that $\bar{h}$ is harmonic in $B_1$, and the proof of Proposition 7.1 is finished.

## 8. Proof of the tilt-excess decay

Recall that $\bar{h}$ is a harmonic function satisfying (see (6.7))

$$\int_{B_1}(|\nabla\bar{h}|^2+\bar{h}^2)\leq C.$$

By standard interior gradient estimates for harmonic functions we get

$$|\nabla\bar{h}(0)|\leq C,\qquad \sup_{B_r}|\nabla\bar{h}-\nabla\bar{h}(0)|\leq Cr,\qquad \forall r\in(0,1/2). \tag{8.1}$$

Thus

$$\int_{B_r}|\nabla\bar{h}-\nabla\bar{h}(0)|^2\leq Cr^{n+2},\qquad \forall r\in(0,1/2). \tag{8.2}$$

In this section we complete the proof of Theorem 3.3. We first consider the special case when $\nabla\bar{h}(0)=0$, and then reduce the general case to this one.

*8.1. The case* $\nabla\bar{h}(0) = 0$

Take a $\psi \in C_0^\infty((-1,1))$ satisfying $0 \le \psi \le 1$, $\psi \equiv 1$ in $(-1/2, 1/2)$, $|\psi'| \le 3$. For any $r \in (0, 1/4)$, choose a $\phi \in C_0^\infty(B_{2r})$ such that $0 \le \phi \le 1$, $\phi \equiv 1$ in $B_r$. In the stationary condition (2.6), take the vector field

$$Y = \phi(x)^2\psi(x_{n+1})^2(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)e_{n+1},$$

where $\lambda_\varepsilon$ is the constant appearing in (6.2).

As in Section 7, by viewing the level set of $u_\varepsilon$ as the graph of a function $h$, because $h$ almost satisfies an elliptic equation, taking such a vector field as a test function corresponds to the procedure of multiplying the equation of $h$ by $h\phi^2$ and then integrating by parts, which is again a standard method in elliptic equation theory. (It is used to derive the Caccioppoli inequality.)

By this choice of $Y$ we get

$$\begin{aligned} 0 = \int_{\mathcal{C}_1} \Bigg( & \left[\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon}W(u_\varepsilon)\right][\phi^2\psi^2 + 2\phi^2\psi\psi'(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)] \\ & - \phi^2\psi^2\nu_{\varepsilon,n+1}^2\varepsilon|\nabla u_\varepsilon|^2 - (x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^n 2\phi\psi^2\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2 \\ & - (x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)2\phi^2\psi\psi'\nu_{\varepsilon,n+1}^2\varepsilon|\nabla u_\varepsilon|^2\Bigg). \end{aligned} \tag{8.3}$$

As in the proof of the Caccioppoli inequality (4.6), those terms containing $\psi'$ are bounded by $O(e^{-1/(C\varepsilon)})$. By the Modica inequality (2.4), (8.3) can be transformed to

$$\begin{aligned} & \int_{\mathcal{C}_1} \phi^2\psi^2[1 - (\nu_\varepsilon \cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2 \\ & \qquad \le \int_{\mathcal{C}_1} 2\phi\psi^2(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^n \frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2 + O(e^{-1/(C\varepsilon)}). \end{aligned}$$

Since $1 - \psi^2 \equiv 0$ in $\{|x_{n+1}| \le 1/2\}$, as before we have

$$\int_{\mathcal{C}_1} \phi^2(1 - \psi^2)[1 - (\nu_\varepsilon \cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2 = O(e^{-1/(C\varepsilon)}).$$

Thus we obtain

$$\begin{aligned} & \int_{\mathcal{C}_1} \phi^2[1 - (\nu_\varepsilon \cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2 \\ & \qquad \le \int_{\mathcal{C}_1} 2\phi\psi^2(x_{n+1} - \lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^n \frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2 + O(e^{-1/(C\varepsilon)}). \end{aligned} \tag{8.4}$$

Now we consider the convergence of the integral on the right hand side of (8.4).

**Lemma 8.1.** *We have*

$$\lim_{\varepsilon\to 0}\delta_\varepsilon^{-2}\int_{\mathcal{C}_1}2\phi\psi^2\,(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^n\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2 = \left[\int_{-1}^1 g'(g^{-1}(t))\,dt\right]\left[\int_{B_1}\phi^2|\nabla\bar{h}(x)|^2\,dx\right].$$

*Proof.* In $\{|u_\varepsilon|\geq 1-b\}$,

$$\begin{aligned}&\left|\int_{\{|u_\varepsilon|\geq 1-b\}\cap\mathcal{C}_1}2\phi\psi^2\,(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^n\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2\right|\\&\quad\leq C\Big(\sup_{B_1}|\phi\psi^2\nabla\phi|\Big)\left[\int_{\{|u_\varepsilon|\geq 1-b\}}\sum_{i=1}^n\nu_{\varepsilon,i}^2\varepsilon|\nabla u_\varepsilon|^2\right]^{1/2}\\&\qquad\times\left[\int_{\{|u_\varepsilon|\geq 1-b\}}(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)^2\varepsilon|\nabla u_\varepsilon|^2\right]^{1/2}\\&\quad=o_b(1)\delta_\varepsilon^2\qquad\text{(by the definition of }\delta_\varepsilon\text{ and Corollary 6.4).}\end{aligned}$$

In $B_\varepsilon$,

$$\begin{aligned}&\left|\int_{B_\varepsilon}2\phi\psi^2\,(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^n\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2\right|\\&\quad\leq C\Big(\sup_{\{|u_\varepsilon|\leq 1-b\}}|x_{n+1}-\lambda_\varepsilon\delta_\varepsilon|\Big)\Big(\sup_{B_1}|\phi\psi^2\nabla\phi|\Big)\left[\int_{B_\varepsilon}\sum_{i=1}^n\nu_{\varepsilon,i}^2\varepsilon|\nabla u_\varepsilon|^2\right]^{1/2}\\&\qquad\times\left[\int_{B_\varepsilon}\varepsilon|\nabla u_\varepsilon|^2\right]^{1/2}\\&\quad=o(\delta_\varepsilon^2),\end{aligned}$$

where we have used the definition of excess, Lemma 5.1 and the fact that $\{|u_\varepsilon|\leq 1-b\}$ is contained in a small neighborhood of $\{x_{n+1}=0\}$ (see Proposition 4.4), which together with Lemma 6.2 implies that

$$\lim_{\varepsilon\to 0}\sup_{\{|u_\varepsilon|\leq 1-b\}}|x_{n+1}-\lambda_\varepsilon\delta_\varepsilon|=0. \tag{8.5}$$

Because $A_\varepsilon\subset\{|x_{n+1}|\leq 1/2\}$, we have $\psi(x_{n+1})\equiv 1$ in $A_\varepsilon$. Hence, by using the $(x,t)$ coordinates,

$$\begin{aligned}&\int_{A_\varepsilon}2\phi\psi^2\,(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^n\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2\\&\quad=-\int_{-1+b}^{1-b}\int_{D_\varepsilon}2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)\varepsilon\left(\frac{\partial h_\varepsilon^t}{\partial t}\right)^{-1}dx\,dt.\end{aligned}\tag{8.6}$$

In $A_\varepsilon$, by (5.16) and (2.5),

$$\varepsilon\left(\frac{\partial h_\varepsilon^t}{\partial t}\right)^{-1}=\varepsilon\frac{\partial u_\varepsilon}{\partial x_{n+1}}\leq C. \tag{8.7}$$

Let $E_\varepsilon$ be the set defined in Lemma 6.5. By the Cauchy inequality, Lemma 6.1, (8.7), (6.2), (6.6) and the Sobolev inequality,

$$\begin{aligned}
&\int_{-1+b}^{1-b}\int_{D_\varepsilon\setminus E_\varepsilon} 2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)\varepsilon\left(\frac{\partial h_\varepsilon^t}{\partial t}\right)^{-1}dx\,dt\\
&\quad\le C\left[\int_{-1+b}^{1-b}\int_{D_\varepsilon\setminus E_\varepsilon}(\nabla\phi\cdot\nabla h_\varepsilon^t)^2\,dx\,dt\right]^{1/2}\left[\int_{-1+b}^{1-b}\int_{D_\varepsilon\setminus E_\varepsilon}(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)^2\phi^2\,dx\,dt\right]^{1/2}\\
&\quad\le C\delta_\varepsilon\mathcal{H}^n(D_\varepsilon\setminus E_\varepsilon)^{\frac{p-1}{2p}}\left[\int_{-1+b}^{1-b}\left(\int_{B_1}(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)^{2p}\phi^{2p}dx\right)^{1/p}dt\right]^{1/2}\\
&\quad\le C\delta_\varepsilon\mathcal{H}^n(D_\varepsilon\setminus E_\varepsilon)^{\frac{p-1}{2p}}\left[\int_{-1+b}^{1-b}\int_{B_1}|\nabla(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)\phi|^2\,dx\,dt+O(\varepsilon^2)\right]^{1/2}\\
&\quad\le C\mathcal{H}^n(D_\varepsilon\setminus E_\varepsilon)^{\frac{p-1}{2p}}\delta_\varepsilon^2=o(\delta_\varepsilon^2).
\end{aligned}$$

In the above, $p>1$ is a constant depending only on the dimension $n$. This estimate gives

$$\int_{-1+b}^{1-b}\int_{D_\varepsilon\setminus E_\varepsilon} 2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)\varepsilon\left(\frac{\partial h_\varepsilon^t}{\partial t}\right)^{-1}dx\,dt=o(\delta_\varepsilon^2). \tag{8.8}$$

Hence by (5.16),

$$\begin{aligned}
&\delta_\varepsilon^{-2}\int_{\mathcal{C}_1}2\phi\psi^2\,(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^n\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2\\
&\quad=-\delta_\varepsilon^{-2}\int_{-1+b}^{1-b}\int_{E_\varepsilon}2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)\varepsilon\left(\frac{\partial h_\varepsilon^t}{\partial t}\right)^{-1}dx\,dt+o_b(1)+o_\varepsilon(1).
\end{aligned}\tag{8.9}$$

In $E_\varepsilon$, by Lemma 6.1, (6.2)–(6.6) and the Cauchy inequality, we have

$$\left|\int_{-1+b}^{1-b}\int_{E_\varepsilon}2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)\,dx\,dt\right|\le C\delta_\varepsilon^2.$$

Then by Lemma 6.5,

$$\begin{aligned}
&\int_{-1+b}^{1-b}\int_{E_\varepsilon}2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)\varepsilon\left(\frac{\partial h_\varepsilon^t}{\partial t}\right)^{-1}dx\,dt\\
&\qquad=\int_{-1+b}^{1-b}\int_{E_\varepsilon}2\phi(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)g'(g^{-1}(t))\,dx\,dt+o(\delta_\varepsilon^2).
\end{aligned}$$

Finally, similar to (8.8), we have

$$\int_{-1+b}^{1-b}\int_{B_1\setminus E_\varepsilon}2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)g'(g^{-1}(t))\,dx\,dt=o(\delta_\varepsilon^2). \tag{8.10}$$

This, combined with Lemma 6.5, implies that

$$\delta_\varepsilon^{-2}\int_{\mathcal{C}_1} 2\phi\psi^2\,(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)(\nabla\phi\cdot\nu_\varepsilon)\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2$$
$$= -\delta_\varepsilon^{-2}\int_{-1+b}^{1-b}\int_{B_1} 2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)(h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon)g'(g^{-1}(t))\,dx\,dt + o_b(1)+o_\varepsilon(1).$$

By the Rellich compactness embedding theorem, Lemma 6.1 and (6.2)–(6.6), it can be directly checked that

$$\lim_{\varepsilon\to 0}\delta_\varepsilon^{-2}\int_{-1+b}^{1-b}\int_{B_1} 2\phi\,(\nabla\phi\cdot\nabla h_\varepsilon^t)[h_\varepsilon^t-\lambda_\varepsilon\delta_\varepsilon]g'(g^{-1}(t))\,dx\,dt$$
$$=\left[\int_{-1+b}^{1-b} g'(g^{-1}(t))\,dt\right]\left[\int_{B_1} 2\phi\,(\nabla\phi\cdot\nabla\bar{h})\bar{h}\,dx\right].$$

Since $\bar{h}$ is a harmonic function (see Proposition 7.1), an integration by parts gives

$$\int_{B_1} 2\phi\,(\nabla\phi\cdot\nabla\bar{h})\bar{h}\,dx = -\int_{B_1}\phi^2|\nabla\bar{h}|^2\,dx.$$

Now we have proved that

$$\lim_{\varepsilon\to 0}\delta_\varepsilon^{-2}\int_{\mathcal{C}_1} 2\phi\psi^2\,(x_{n+1}-\lambda_\varepsilon\delta_\varepsilon)\sum_{i=1}^{n}\frac{\partial\phi}{\partial x_i}\nu_{\varepsilon,i}\nu_{\varepsilon,n+1}\varepsilon|\nabla u_\varepsilon|^2$$
$$=\left[\int_{-1+b}^{1-b} g'(g^{-1}(t))\,dt\right]\left[\int_{B_1}\phi^2|\nabla\bar{h}|^2\,dx\right]+o_b(1).$$

As in the proof of Proposition 7.1, we can let $b\to 0$ to finish the proof. □

Note that

$$\int_{-1}^{1} g'(g^{-1}(t))\,dt = \int_{-\infty}^{\infty} g'(s)^2\,ds = \sigma_0.$$

By (8.2), we can choose a $\theta\in(0,1/2)$ such that

$$\theta^{-n}\int_{B_{2\theta}}|\nabla\bar{h}|^2 \le C\theta^2 \le \frac{\theta}{4\max\{\sigma_0,1\}}. \tag{8.11}$$

Then by choosing $r=2\theta$ in the definition of $\phi$, (8.4) and Lemma 8.1 give, for all $\varepsilon$ small,

$$\theta^{-n}\int_{\mathcal{C}_\theta}[1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2 \le \frac{\theta}{3}\delta_\varepsilon^2,$$

which contradicts the initial assumption (3.9). This completes the proof of Theorem 3.3 in the special case $\nabla\bar{h}(0)=0$.

### *8.2. The general case*

In general $\nabla\bar{h}(0)$ may not be 0, and we only have an estimate as in (8.1). Here we show how to reduce this problem to the special case treated in the previous subsection.

For each $\varepsilon$, take a rotation $T_\varepsilon \in \mathrm{SO}(n+1)$ such that

$$T_\varepsilon e_{n+1} = e_\varepsilon := \frac{e_{n+1} + \delta_\varepsilon \nabla \bar{h}(0)}{(1+\delta_\varepsilon^2 |\nabla \bar{h}(0)|^2)^{1/2}}. \tag{8.12}$$

Next define $\tilde{u}_\varepsilon(X) := u_\varepsilon(T_\varepsilon X)$, which is still a solution of (2.2) in $\mathcal{B}_4$.

By (8.1),

$$|e_\varepsilon - e_{n+1}| \le C\delta_\varepsilon. \tag{8.13}$$

We can also choose $T_\varepsilon$ so that it satisfies the following estimates.

**Lemma 8.2.**

$$\|T_\varepsilon - I\| \le C\delta_\varepsilon, \qquad \|\Pi \circ T_\varepsilon - I_{\mathbb{R}^n}\| \le C\delta_\varepsilon^2. \tag{8.14}$$

*Proof.* Choose a basis in $\mathbb{R}^n$ so that $\nabla \bar{h}(0) = |\nabla \bar{h}(0)| e_n$. We have defined $T_\varepsilon e_{n+1}$. Now take

$$T_\varepsilon e_i = e_i \quad \text{for } 1 \le i \le n-1, \qquad T_\varepsilon e_n = \frac{e_n - \delta_\varepsilon |\nabla \bar{h}(0)| e_{n+1}}{(1+\delta_\varepsilon^2 |\nabla \bar{h}(0)|^2)^{1/2}}.$$

In particular, $T_\varepsilon$ is only a rotation in the $(e_n, e_{n+1})$-plane.

Since $\delta_\varepsilon |\nabla \bar{h}(0)| \le 1/2$ (recall that $\delta_\varepsilon$ converges to 0 and we have a universal bound on $|\nabla \bar{h}(0)|$), the first inequality in (8.14) can be verified directly. For the second, we have

$$\begin{aligned} |\Pi \circ T_\varepsilon e_n - e_n| &= \left| \frac{e_n}{(1+\delta_\varepsilon^2 |\nabla \bar{h}(0)|^2)^{1/2}} - e_n \right| = 1 - \frac{1}{(1+\delta_\varepsilon^2 |\nabla \bar{h}(0)|^2)^{1/2}} \\ &\le C\delta_\varepsilon^2 |\nabla \bar{h}(0)|^2, \end{aligned}$$

and $\Pi \circ T_\varepsilon e_i = e_i$ for $1 \le i \le n-1$. This finishes the proof. □

Similar to $\nu_\varepsilon$, define the unit normal vector $\tilde{\nu}_\varepsilon$ associated to $\tilde{u}_\varepsilon$ as in Section 2.

**Lemma 8.3.** *There exists a universal constant $C$ such that*

$$\int_{\mathcal{C}_{3/4}} [1 - (\tilde{\nu}_\varepsilon \cdot e_{n+1})^2] \varepsilon |\nabla \tilde{u}_\varepsilon|^2 \le C\delta_\varepsilon^2.$$

*Proof.* First by noting (8.14) and a change of variables, we have

$$\begin{aligned} &\int_{\mathcal{C}_{3/4}} [1 - (\tilde{\nu}_\varepsilon \cdot e_{n+1})^2] \varepsilon |\nabla \tilde{u}_\varepsilon|^2 \\ &\quad = \int_{T_\varepsilon^{-1}(B_{3/4} \times \{|x_{n+1}| < 1/2\})} [1 - (\nu_\varepsilon \cdot e_\varepsilon)^2] \varepsilon |\nabla u_\varepsilon|^2 + O(e^{-c/\varepsilon}) \\ &\quad \le \int_{\mathcal{C}_1} [1 - (\nu_\varepsilon \cdot e_\varepsilon)^2] \varepsilon |\nabla u_\varepsilon|^2 + O(e^{-c/\varepsilon}), \end{aligned} \tag{8.15}$$

where $O(e^{-c/\varepsilon})$ represents the contribution from the part near $B_1 \times \{\pm 1\}$ where Proposition 4.4 applies.

By (8.12),

$$1-(\nu_\varepsilon\cdot e_\varepsilon)^2\le 1-(\nu_\varepsilon\cdot e_{n+1})^2+2(\nu_\varepsilon\cdot e_{n+1})^2\left(1-\frac{1}{(1+\delta_\varepsilon^2|\nabla\bar{h}(0)|^2)^{1/2}}\right)$$
$$+2\delta_\varepsilon|\nu_\varepsilon\cdot e_{n+1}|\,|\nu_\varepsilon\cdot\nabla\bar{h}(0)|.$$

By definition,

$$\int_{\mathcal{C}_1}[1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2=\delta_\varepsilon^2.$$

Next, by (8.1),

$$2(\nu_\varepsilon\cdot e_{n+1})^2\left(1-\frac{1}{(1+\delta_\varepsilon^2|\nabla\bar{h}(0)|^2)^{1/2}}\right)\le C\delta_\varepsilon^2.$$

Finally, by noting that

$$|\nu_\varepsilon\cdot\nabla\bar{h}(0)|\le|\nabla\bar{h}(0)|\Big(\sum_{i=1}^n\nu_{\varepsilon,i}^2\Big)^{1/2}\le C[1-(\nu_\varepsilon\cdot e_{n+1})^2]^{1/2},$$

we can use the Cauchy inequality to derive that

$$\delta_\varepsilon\int_{\mathcal{C}_1}|\nu_\varepsilon\cdot e_{n+1}|\,|\nu_\varepsilon\cdot\nabla\bar{h}(0)|\varepsilon|\nabla u_\varepsilon|^2$$
$$\le C\delta_\varepsilon\left(\int_{\mathcal{C}_1}|\nu_\varepsilon\cdot e_{n+1}|^2\varepsilon|\nabla u_\varepsilon|^2\right)^{1/2}\left(\int_{\mathcal{C}_1}[1-(\nu_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla u_\varepsilon|^2\right)^{1/2}\le C\delta_\varepsilon^2.$$

Putting these together we get

$$\int_{\mathcal{C}_1}[1-(\nu_\varepsilon\cdot e_\varepsilon)^2]\varepsilon|\nabla u_\varepsilon|^2\le C\delta_\varepsilon^2.$$

Substituting this into (8.15) and noting (3.7) finishes the proof. □

With this lemma in hand, we can proceed as before to construct the Lipschitz functions $\tilde{h}_\varepsilon^t$, and prove that $\delta_\varepsilon^{-1}(\tilde{h}_\varepsilon^t-\tilde{\lambda}_\varepsilon\delta_\varepsilon)$ converge to a harmonic function $\tilde{h}$ (the constant $\tilde{\lambda}_\varepsilon$ is defined as $\lambda_\varepsilon$), weakly in $H^1(B_{3/4})$ and strongly in $L^2(B_{3/4})$.

However by the definition of $\tilde{u}_\varepsilon$, the graph of $\tilde{h}_\varepsilon^t$ is only a rotation of the one of $h_\varepsilon^t$. More precisely, for any $x\in B_{3/4}$ and $t\in(-1+b,1-b)$,

$$\frac{\tilde{h}_\varepsilon^t(x)+\delta_\varepsilon\nabla\bar{h}(0)\cdot x}{(1+\delta_\varepsilon^2|\nabla\bar{h}(0)|^2)^{1/2}}=h_\varepsilon^t(\Pi\circ T_\varepsilon(x,\tilde{h}_\varepsilon^t(x))).$$

In fact, because $\tilde{u}_\varepsilon(X)=t$ if and only if $T_\varepsilon X\in u_\varepsilon^{-1}(t)$, $x_{n+1}=\tilde{h}_\varepsilon^t(x)$ if and only if

$$(T_\varepsilon X)_{n+1}=h_\varepsilon^t(\Pi\circ T_\varepsilon X),$$

which can be written as

$$\frac{x_{n+1}+\delta_\varepsilon\nabla\bar{h}(0)\cdot x}{(1+\delta_\varepsilon^2|\nabla\bar{h}(0)|^2)^{1/2}}=h_\varepsilon^t\left(x_1,\dots,x_{n-1},\frac{x_n-\delta_\varepsilon|\nabla\bar{h}(0)|x_{n+1}}{(1+\delta_\varepsilon^2|\nabla\bar{h}(0)|^2)^{1/2}}\right).$$

From this we deduce that

$$\begin{aligned}\tilde{h}^t_\varepsilon(x) &= h^t_\varepsilon\big(x_1,\dots,x_{n-1},x_n-\delta_\varepsilon|\nabla\bar{h}(0)|\tilde{h}^t_\varepsilon(x)+O(\delta_\varepsilon^2)\big)-\delta_\varepsilon\nabla\bar{h}(0)\cdot x+O(\delta_\varepsilon^2)\\ &= h^t_\varepsilon(x)-\delta_\varepsilon\nabla\bar{h}(0)\cdot x+o(\delta_\varepsilon).\end{aligned}$$

Here we have used the facts that the Lipschitz constant of $h^t_\varepsilon$ is smaller than $1/2$ (by its construction), and the sup bound of $\tilde{h}^t_\varepsilon$ goes to 0 as $\varepsilon\to 0$ (by Proposition 4.4).

Hence $\tilde{\lambda}_\varepsilon-\lambda_\varepsilon=o_\varepsilon(1)$, and

$$\tilde{h}(x)=\lim_{\varepsilon\to 0}[\tilde{h}^t_\varepsilon/\delta_\varepsilon-\tilde{\lambda}_\varepsilon]=\lim_{\varepsilon\to 0}[h^t_\varepsilon/\delta_\varepsilon-\lambda_\varepsilon-\nabla\bar{h}(0)\cdot x]=\bar{h}(x)-\nabla\bar{h}(0)\cdot x.$$

Combined with Proposition 7.1, this implies that $\tilde{h}$ is a harmonic function in $B_{3/4}$ satisfying $\nabla\tilde{h}(0)=0$. Then we can proceed as in the previous subsection. By choosing a smaller $\theta$ to incorporate the constant $C$ appearing in Lemma 8.3, for all $\varepsilon$ small,

$$\theta^{-n}\int_{\mathcal{C}_\theta}[1-(\tilde{\nu}_\varepsilon\cdot e_{n+1})^2]\varepsilon|\nabla\tilde{u}_\varepsilon|^2\le\frac{\theta}{2C}\delta_\varepsilon^2. \tag{8.16}$$

Here $C$ is the constant appearing in Lemma 8.3, due to a change of variable associated to the rotation $T_\varepsilon$. After rotating back, this contradicts (3.9) and finishes the proof of Theorem 3.3.

# Part II. Uniform $C^{1,\alpha}$ regularity of intermediate layers

## 9. Statement

In this part we prove the following local uniform $C^{1,\alpha}$ regularity for intermediate layers. This parallels Allard's $\varepsilon$-regularity theorem for stationary varifolds.

**Theorem 9.1.** *For any $b\in(0,1)$, there exist five universal constants $\varepsilon_A,\tau_A,\alpha_A\in(0,1)$ and $R_A,K_A$ such that the following holds. Let $u_\varepsilon$ be a solution of (2.2) with $\varepsilon\le\varepsilon_A$, defined in $\mathcal{B}_{R_A}$, satisfying $|u_\varepsilon(0)|\le 1-b$ and*

$$R_A^{-n}\int_{\mathcal{B}_{R_A}}\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right)\le(1+\tau_A)\omega_n\sigma_0. \tag{9.1}$$

*Then there exists a hyperplane, say $\mathbb{R}^n$ (after a suitable rotation), such that for any $t\in(-1+b,1-b)$, the set $\{u_\varepsilon=t\}\cap\mathcal{C}_1$ is a $C^{1,\alpha_A}$ hypersurface represented by the graph of a function $x_{n+1}=h^t_\varepsilon(x)$ with*

$$\|h^t_\varepsilon\|_{C^{1,\alpha_A}(B_1)}\le K_A.$$

Assume the limit varifold $V$ of $u_\varepsilon$ satisfies the assumptions in Allard's $\varepsilon$-regularity theorem at the origin 0. Hence it is a smooth minimal hypersurface with unit density near 0.

By enlarging this minimal hypersurface around 0, the assumptions in this theorem are fulfilled and this theorem applies, so that, in a neighborhood of 0, intermediate layers of $u_\varepsilon$ are hypersurfaces with uniformly $C^{1,\alpha_A}$ bound and they converge to the minimal hypersurface in a $C^{1,\alpha_A}$ manner.

To prove this theorem, we first use Theorem 3.3 to obtain a Morrey type bound. As explained in Section 1, due to the assumption $\delta_\varepsilon \geq K_0\varepsilon$ in Theorem 3.3, this Morrey type bound does not give the required $C^{1,\alpha_A}$ regularity. It only says that at every scale up to $O(\varepsilon)$, $\{u_\varepsilon = t\}$ is close to a *fixed* hyperplane, i.e. a kind of Lipschitz regularity for $\{u_\varepsilon = t\}$ up to $O(\varepsilon)$ scales. This is already sufficient for the proof of Theorem 1.1, which is given in Section 11. The proof of Theorem 9.1 will be completed in Section 12, and it uses the intermediate results established in Section 11.

## 10. A Morrey type bound

In this section, $u_\varepsilon$ denotes a fixed solution satisfying all the assumptions in Theorem 9.1. Here we prove

**Lemma 10.1.** *There exist two universal constants $K_1$ and $K_2$ such that for any $X_0 \in \{|u_\varepsilon| \leq 1-b\} \cap \mathcal{B}_1$ and any ball $\mathcal{B}_r(X_0)$ with $r \in (K_1\varepsilon, \theta)$, we can find a unit vector $e_r(X_0)$ such that*

$$r^{-n}\int_{\mathcal{B}_r(X_0)}[1-(\nu_\varepsilon\cdot e_r(X_0))^2]\varepsilon|\nabla u_\varepsilon|^2 \leq K_2^2\max\{\varepsilon^2 r^{-2}, \delta_0^2 r^\alpha\}. \tag{10.1}$$

*Here $\alpha = |\log(\theta/2)|/|\log\theta| \in (1,2)$.*

For convenience, we shall replace the cylinders $\mathcal{C}_2$ and $\mathcal{C}_\theta$ in Theorem 3.3 by balls $\mathcal{B}_1$ and $\mathcal{B}_\theta$ respectively. This may change the constants in that theorem by a factor, which however only depends on the dimension $n$ and does not affect our argument too much.

By the monotonicity formula (Proposition 4.1) and (9.1), if $R_A$ is sufficiently large, then for any $X \in \mathcal{B}_1$ and $r \in (0, R_A - 1)$,

$$r^{-n}\int_{\mathcal{B}_r(X)}\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon}W(u_\varepsilon)\right) \leq (1+2\tau_A)\omega_n\sigma_0. \tag{10.2}$$

If $\tau_A$ is sufficiently small, we can apply Proposition 4.4 to $u_\varepsilon(rX)$, which gives

**Lemma 10.2.** *For any $\delta > 0$, there exists a $K(\delta)$ such that for any $X \in \{|u_\varepsilon| \leq 1-b\}\cap\mathcal{B}_1$ and $r \in (K(\delta)\varepsilon, 1)$, there exists a hyperplane $P_r(X)$ such that*

$$\mathrm{dist}_H(\{u_\varepsilon = u_\varepsilon(X)\}\cap\mathcal{B}_r(X), P_r(X)\cap\mathcal{B}_r(X)) \leq \delta r.$$

By Lemma 4.6, if $r \geq K_1\varepsilon$ (with $K_1$ the constant determined by Lemma 4.6) and $\tau_A$ is sufficiently small, the excess with respect to $P_r(X)$ (with unit normal vector $e_r(X)$) satisfies

$$E(r; X, u_\varepsilon, P_r(X)) \leq \delta_0^2, \tag{10.3}$$

with $\delta_0$ as in Theorem 3.3. Note that in (10.3) it is integrated on $\mathcal{B}_r(X)$, not on a cylinder.

Now Theorem 3.3 applies. In the current setting it reads:

**Lemma 10.3.** *If $E(r; X, u_\varepsilon, P_r(X)) \geq K_0^2 r^{-2}\varepsilon^2$, then there exists another hyperplane $\tilde{P}_r(X)$ such that*

$$E(\theta r; X, u_\varepsilon, \tilde{P}_r(X)) \leq \frac{\theta}{2} E(r; X, u_\varepsilon, P_r(X)).$$

*Here $\tilde{e}_r(X)$, one of the unit normal vectors to $\tilde{P}_r(X)$, satisfies*

$$\|\tilde{e}_r(X) - e_r(X)\| \leq C E(r; X, u_\varepsilon, P_r(X))^{1/2}.$$

The constant $\theta$ may be different from the one in Theorem 3.3, but we still have $\theta < 1$.

With this lemma in hand we can prove Lemma 10.1. The following proof is similar to the one of [26, Theorem 2.3].

*Proof of Lemma 10.1.* Assume $X_0 = 0$. For $k \geq 0$, let $r_k = \theta^k$. Define

$$E_k := \min_{e \in \mathbb{S}^n} \varepsilon^{-2} r_k^{2-n} \int_{\mathcal{B}_{r_k}} [1 - (\nu_\varepsilon \cdot e)^2] \varepsilon |\nabla u_\varepsilon|^2.$$

Take a unit vector $\bar{e}_k$ attaining this minimum.

As in (10.3), for all $r_k \geq K_1 \varepsilon$,

$$E_k \leq \delta_0^2 \varepsilon^{-2} r_k^2. \tag{10.4}$$

Lemma 10.3 implies that, once $E_k \geq K_0^2$, then

$$E_{k+1} \leq \frac{\theta^3}{2} E_k. \tag{10.5}$$

Moreover, by the definition of $E_k$, we always have

$$E_{k+1} \leq \theta^{2-n} E_k. \tag{10.6}$$

Let $k_1 \in \mathbb{N}$ be the unique number satisfying $\theta^{k_1} \in [K_1\varepsilon, K_1\theta^{-1}\varepsilon)$.

Now we derive the claimed bound on $E_k$ from (10.4)–(10.6), for $k \leq k_1$. Let $k_0$ be the smallest number such that, for all $k > k_0$,

$$E_k \leq K_0^2 \theta^{2-n}. \tag{10.7}$$

As we will see below, this is well defined.

If $k_0 = 0$, then for any $0 \leq k \leq k_1$,

$$\int_{\mathcal{B}_{r_k}} [1 - (\nu_\varepsilon \cdot \bar{e}_k)^2] \varepsilon |\nabla u_\varepsilon| \leq K_0^2 \theta^{2-n} \varepsilon^2 r_k^{n-2}. \tag{10.8}$$

This can be extended to those $r \in [K_1\varepsilon, \theta)$ by choosing a (unique) $k$ so that $r \in [r_{k+1}, r_k)$.

Next we assume there exists a $\tilde{k} > 0$ such that $E_{\tilde{k}} \geq K_0^2 \theta^{2-n}$. By (10.6), $E_{\tilde{k}-1} \geq K_0^2$. Then (10.5) applies, which says

$$E_{\tilde{k}-1} \geq \frac{2}{\theta^3} E_{\tilde{k}}.$$

In particular,

$$E_{\tilde{k}-1} \geq E_{\tilde{k}} \geq K_0 \theta^{2-n}.$$

With this estimate we can repeat the above procedure to deduce that, for all $i \in [0, \tilde{k})$,

$$E_i \geq \frac{2}{\theta^3} E_{i+1} \geq K_0^2 \theta^{2-n}.$$

From this we see $k_0$ is well defined.

The above decay estimate implies that, for all $i \leq k_0$,

$$E_i \leq \left(\frac{\theta^3}{2}\right)^i E_0,$$

in other words,

$$\int_{\mathcal{B}_{r_i}} [1 - (\nu_\varepsilon \cdot \bar{e}_i)^2] \varepsilon |\nabla u_\varepsilon| \leq \delta_0^2 r_i^{n+\alpha}. \tag{10.9}$$

This estimate can also be extended to those $r \in [r_{k_0}, \theta)$ by choosing an $i$ so that $r \in [r_{i+1}, r_i)$.

In conclusion, for $r \in [r_{k_0}, \theta)$, we have the estimate (10.9), and for $r \in (K_1\varepsilon, r_{k_0})$ (10.8) applies. By choosing a suitable universal constant $K_2$, (10.1) follows from these two estimates. □

Next we show that $e_r(X_0)$ can be replaced by a fixed unit vector (independent of $r$).

**Lemma 10.4.** *For any $\sigma > 0$, there exist constants $K_3 := K_3(\sigma)$ and $K_4$ (with $K_4$ universal, independent of $\sigma$) such that for any $X_0 \in \{|u_\varepsilon| \leq 1 - b\} \cap \mathcal{B}_1$ and any ball $\mathcal{B}_r(X_0)$ with $r \in (K_3\varepsilon, \theta)$, there exists a unit vector $e(X_0)$ such that*

$$r^{-n} \int_{\mathcal{B}_r(X_0)} [1 - (\nu_\varepsilon \cdot e(X_0))^2] \varepsilon |\nabla u_\varepsilon|^2 \leq \sigma + K_4 \delta_0 r^{\alpha/2}. \tag{10.10}$$

*Here $e(X_0)$ is independent of $r \in (K_3\varepsilon, \theta)$.*

*Proof.* Keep notation as in the proof of Lemma 10.1. For any $r \in (K_1\varepsilon, \theta)$, combining Remark 3.2 and Lemma B.5, we get

$$\begin{aligned}
&\int_{\mathcal{B}_{2r}(X_0)} [1 - (\nu_\varepsilon \cdot e_{2r}(X_0))^2] \varepsilon |\nabla u_\varepsilon|^2 + \int_{\mathcal{B}_r(X_0)} [1 - (\nu_\varepsilon \cdot e_r(X_0))^2] \varepsilon |\nabla u_\varepsilon|^2 \\
&\quad \geq c \int_{\mathcal{B}_r(X_0)} [\mathrm{dist}_{\mathbb{RP}^n}(\nu_\varepsilon, e_r(X_0))^2 + \mathrm{dist}_{\mathbb{RP}^n}(\nu_\varepsilon, e_{2r}(X_0))^2] \varepsilon |\nabla u_\varepsilon|^2 \\
&\quad \geq c\, \mathrm{dist}_{\mathbb{RP}^n}(e_{2r}(X_0), e_r(X_0))^2 \int_{\mathcal{B}_r(X_0)} \varepsilon |\nabla u_\varepsilon|^2 \geq c\, \mathrm{dist}_{\mathbb{RP}^n}(e_{2r}(X_0), e_r(X_0))^2 r^n.
\end{aligned}$$

For $k < k_0$, by Lemma 10.1 this gives

$$\mathrm{dist}_{\mathbb{RP}^n}(e_{k+1}(X_0), e_k(X_0)) \leq K_2 \delta_0 r_k^{\alpha/2} = K_2 \delta_0 \theta^{\alpha k/2}.$$

Summing in $i$ from $k$ to $k_0$, we see

$$\mathrm{dist}_{\mathbb{RP}^n}(e_{k_0}(X_0), e_k(X_0)) \leq \frac{K_2}{1 - \theta^{\alpha/2}} \delta_0 \theta^{\alpha k/2} = \frac{K_2}{1 - \theta^{\alpha/2}} \delta_0 r_k^{\alpha/2}, \quad \forall k < k_0. \tag{10.11}$$

For $k \in [k_0, k_1)$, we have

$$\mathrm{dist}_{\mathbb{RP}^n}(e_{k+1}(X_0), e_k(X_0)) \le K_2 \varepsilon r_k^{-1} = K_2 \varepsilon \theta^{-k}. \tag{10.12}$$

Let $k_2 \le k_1$ be the largest number satisfying

$$\frac{K_2}{\theta^{-1}-1}\varepsilon\theta^{-k_2-1} + K_2^2\varepsilon^2\theta^{-2k_2-2} \le \frac{\theta^n}{4}\sigma. \tag{10.13}$$

Note that there exists a constant $K_3(\sigma)$ such that $r_{k_2} = \theta^{k_2} \le K_3(\sigma)\varepsilon$. Summing (10.12) from $k$ to $k_2$, we get

$$\mathrm{dist}_{\mathbb{RP}^n}(e_{k_2}(X_0), e_k(X_0)) \le \varepsilon\frac{K_2}{\theta^{-1}-1}\theta^{-k_2-1} \le \frac{\theta^n}{4}\sigma, \qquad \forall k_0 \le k \le k_2. \tag{10.14}$$

In particular,

$$\mathrm{dist}_{\mathbb{RP}^n}(e_{k_2}(X_0), e_{k_0}(X_0)) \le \frac{\theta^n}{4}\sigma. \tag{10.15}$$

Let $e(X_0) = e_{k_2}(X_0)$. By (10.11)–(10.15) we obtain, for any $k \in (0, k_2)$,

$$\mathrm{dist}_{\mathbb{RP}^n}(e_k(X_0), e(X_0)) \le \frac{\theta^n}{4}\sigma + \frac{K_2}{1-\theta^{\alpha/2}}\delta_0 r_k^{\alpha/2}.$$

For any $k \ge 0$, similar to Remark 3.2, we have

$$1 - (\nu_\varepsilon \cdot e(X_0))^2 \le [1 - (\nu_\varepsilon \cdot e_k)^2] + 2\,\mathrm{dist}_{\mathbb{RP}^n}(e_k, e(X_0)).$$

Together with (10.1) and (10.13), this gives

$$r_k^{-n}\int_{\mathcal{B}_{r_k}} [1-(\nu_\varepsilon\cdot e(X_0))^2]\varepsilon|\nabla u_\varepsilon|^2 \le \frac{3\theta^n}{4}\sigma + \left(\frac{2K_2}{1-\theta^{\alpha/2}} + K_2^2\right)\delta_0 r_k^{\alpha/2}.$$

For any $r \in (K_3\varepsilon, \theta)$, by choosing a $k$ so that $r \in (r_k, r_{k+1}]$, we obtain

$$r^{-n}\int_{\mathcal{B}_r} [1-(\nu_\varepsilon\cdot e(X_0))^2]\varepsilon|\nabla u_\varepsilon|^2 \le \sigma + \theta^{-n}\left(\frac{2K_2}{1-\theta^{\alpha/2}} + K_2^2\right)\delta_0 r_k^{\alpha/2}. \tag{10.16}$$

By taking

$$K_4 := \theta^{-n}\left(\frac{2K_2}{1-\theta^{\alpha/2}} + K_2^2\right),$$

which is indeed a universal constant and does not depend on $\sigma$, we get (10.10). □

The following result will be used in the proof of Lipschitz regularity of $\{u_\varepsilon = 0\}$.

**Corollary 10.5.** *For any $X_0 \in \{u_\varepsilon = 0\} \cap \mathcal{B}_1$,*

$$|e(X_0) - e_{n+1}| \le C(\sigma^{1/2} + \delta_0^{1/2}).$$

*Proof.* By taking $r = \theta$ in (10.10), we have

$$\theta^{-n} \int_{\mathcal{B}_\theta(X_0)} [1 - (\nu_\varepsilon \cdot e(X_0))^2] \varepsilon |\nabla u_\varepsilon|^2 \leq \sigma + K_4 \delta_0.$$

On the other hand, by (9.1) and Lemma 4.6, we also have

$$\theta^{-n} \int_{\mathcal{B}_\theta(X_0)} [1 - (\nu_\varepsilon \cdot e_{n+1})^2] \varepsilon |\nabla u_\varepsilon|^2 \leq C \delta_0^2.$$

Similar to the proof of the previous lemma, combining these two and using Lemma B.5, we get

$$\begin{aligned} &\mathrm{dist}_{\mathbb{RP}^n}(e(X_0), e_{n+1})^2 \\ &\quad \leq \theta^{-n} \int_{\mathcal{B}_{1/4}(X_0)} [1 - (\nu_\varepsilon \cdot e(X_0))^2] \varepsilon |\nabla u_\varepsilon|^2 + \theta^{-n} \int_{\mathcal{B}_{1/4}(X_0)} [1 - (\nu_\varepsilon \cdot e_{n+1})^2] \varepsilon |\nabla u_\varepsilon|^2 \\ &\quad \leq C(\sigma + \delta_0). \end{aligned}$$

Finally, we can fix $e(X_0)$ so that it points upwards. Thus the estimate on the distance in $\mathbb{RP}^n$ can be lifted to an estimate in $\mathbb{S}^n$. □

What we have proved can be roughly stated as follows: level sets of $u_\varepsilon$ are Lipschitz graphs in the form of $x_{n+1} = h_\varepsilon(x)$ up to the scale $K_3\varepsilon$. However, this may break down for smaller scales, because in Lemma 10.4, $K_3$ depends on $\sigma$. To obtain further control on the scale smaller than $K_3\varepsilon$, we first give a direct proof of Theorem 1.1 and then use this to prove the full regularity of level sets of $u_\varepsilon$.

## 11. A direct proof of Theorem 1.1

This section is devoted to a direct proof of Theorem 1.1. In fact, we prove something more.

**Theorem 11.1.** *Suppose that $u$ is a smooth solution of (2.1) on $\mathbb{R}^{n+1}$ satisfying*

$$\lim_{R\to\infty} R^{-n} \int_{\mathcal{B}_R} \left(\frac{1}{2}|\nabla u|^2 + W(u)\right) \leq (1 + \tau_A)\omega_n \sigma_0. \tag{11.1}$$

*Then there exists a unit vector $e$ and a constant $t \in \mathbb{R}$ such that $u(X) \equiv g(e \cdot X + t)$.*

In the following we will show that if $u$ is a minimizing solution of (2.1) on $\mathbb{R}^{n+1}$, where $n \leq 6$, then (11.1) is satisfied. Thus Theorem 1.1 is a corollary of this theorem.

Since $u$ is an entire solution, by the main result of [17], $u$ satisfies the Modica inequality and hence the monotonicity formula of Proposition 4.1 for any $X \in \mathbb{R}^{n+1}$ and $r > 0$. This monotonicity ensures the existence of the limit in (11.1). It also implies that, for any ball $\mathcal{B}_R(X) \subset \mathbb{R}^{n+1}$,

$$R^{-n} \int_{\mathcal{B}_R(X)} \left(\frac{1}{2}|\nabla u|^2 + W(u)\right) \leq (1 + \tau_A)\omega_n \sigma_0.$$

With this bound, we can study the asymptotic behavior of $u$ through the scaling

$$u_\varepsilon(X) := u(\varepsilon^{-1}X).$$

As before, by Hutchinson–Tonegawa theory, the varifolds $V_\varepsilon$ associated to $u_\varepsilon$ converge to a stationary varifold $V$ with integer multiplicity.

Furthermore, we claim:

**Proposition 11.2.** *$V$ is a cone with respect to the origin* 0.

*Proof.* This is because for any $R > 0$, by the convergence of $\|V_\varepsilon\|$ and (4.2),

$$\begin{aligned} R^{-n}\|V\|(\mathcal{B}_R) &= \lim_{\varepsilon\to 0} R^{-n}\|V_\varepsilon\|(\mathcal{B}_R) \\ &= \lim_{\varepsilon\to 0} R^{-n}\int_{\mathcal{B}_R}\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon}W(u_\varepsilon)\right) \qquad \text{(by the definition of } V_\varepsilon) \\ &= \lim_{\varepsilon\to 0}(\varepsilon^{-1}R)^{-n}\int_{\mathcal{B}_{\varepsilon^{-1}R}}\left(\frac{1}{2}|\nabla u|^2 + W(u)\right) \qquad \text{(by the definition of } u_\varepsilon). \end{aligned} \tag{11.2}$$

In the last line, the existence of the limit follows from the energy bound (11.1) and the monotonicity formula of Proposition 4.1. Note that this limit is independent of $R$. Then by the monotonicity formula for stationary varifolds [16, Theorem 6.3.2], we deduce that $V$ is a cone with respect to the origin. □

By (11.1) and (11.2),

$$\|V\|(\mathcal{B}_1) \le (1+\tau_A)\omega_n\sigma_0.$$

Hence we can apply Allard's $\varepsilon$-regularity theorem to deduce that spt $\|V\|$ is a smooth hypersurface in a neighborhood of the origin. Then by the previous proposition, spt $\|V\|$ must be a hyperplane and $V$ is the standard varifold associated to this plane with unit density.

Let

$$\Phi_\varepsilon := g_\varepsilon^{-1}\circ u_\varepsilon$$

be the distance type function (see Appendix A). Combining this blowing down analysis and Proposition A.2, we get

**Proposition 11.3.** *As $\varepsilon \to 0$, $\Phi_\varepsilon$ converges (up to a subsequence of $\varepsilon \to 0$) to a linear function of the form $e\cdot X$ in $C_{\mathrm{loc}}(\mathbb{R}^{n+1})$, where $e$ is a unit vector.*

However, this argument does not give the uniqueness of this limit. Different subsequences of $\varepsilon \to 0$ may lead to different limits. To obtain the uniqueness of the blowing down limit, we use the following lemma.

**Lemma 11.4.** *There exists a universal constant $C$ such that for any ball $\mathcal{B}_R(X)$ with $R \ge 1$, we can find a unit vector $e_R$ satisfying*

$$\int_{\mathcal{B}_R(X)}[1-(\nu\cdot e_R)^2]\,|\nabla u|^2 \le CR^{n-2}. \tag{11.3}$$

The proof is similar to the one of Lemma 10.1 (see also the proof of [26, Theorem 2.3]).

Note that $e_R$ in this theorem may not be unique. In the following we assume that for each $R > 1$, such a vector $e_R$ has been fixed.

If $n = 1$, then as $R \to \infty$, since $e_R$ are unit vectors, we can take a subsequence of $R_i \to \infty$ such that $e_{R_i} \to e_\infty \in \mathbb{S}^1$. Assume $e_\infty = e_2$. Then by taking the limit in (11.3), we get

$$\int_{\mathcal{B}_R(X)} \left(\frac{\partial u}{\partial x_1}\right)^2 = 0, \qquad \forall R > 0.$$

Thus $u(x_1, x_2) \equiv u(x_2)$.

Now consider the case $n \geq 2$. Similar to Lemma 10.4, we also have

**Lemma 11.5.** *There exists a unit vector $e_\infty$ and a universal constant $C$ such that*

$$\int_{\mathcal{B}_R(X)} [1 - (\nu \cdot e_\infty)^2] \, |\nabla u|^2 \leq C R^{n-2}, \qquad \forall R > 1. \tag{11.4}$$

For the blowing down sequence $u_\varepsilon$, (11.4) implies that

$$\int_{\mathcal{B}_1} [1 - (\nu_\varepsilon \cdot e_\infty)^2] \varepsilon |\nabla u_\varepsilon|^2 \leq C\varepsilon^2. \tag{11.5}$$

Note that this estimate just says that the assumption $\delta_\varepsilon \gg \varepsilon$ in Theorem 3.3 is not satisfied.

For any $\eta \in C_0^\infty(\mathbb{R}^{n+1})$, let $\Phi(X, S) = \eta(X)^2 \langle S e_\infty, e_\infty \rangle \in C_0^\infty(\mathbb{R}^{n+1} \times G(n))$. Passing to the limit in (11.5) gives

$$0 = \lim_{\varepsilon \to 0} \langle V_\varepsilon, \Phi \rangle = \langle V, \Phi \rangle.$$

Thus for $\|V\|$-a.a. $X$, the tangent plane of $V$ at $X$ is the hyperplane orthogonal to $e_\infty$. It can be directly checked that $V$ must be the standard varifold associated to this hyperplane. (This can also be seen by noting that we have proved that spt $\|V\|$ is a hyperplane.)

The uniqueness of $V$ also implies that the limit of $\Phi_\varepsilon$ in Proposition 11.3 is independent of the choice of subsequences of $\varepsilon \to 0$, i.e.,

$$\Phi_\varepsilon \to e_\infty \cdot X \quad \text{ in } C_{\mathrm{loc}}(\mathbb{R}^{n+1}).$$

Without loss of generality, assume $e_\infty = e_{n+1}$.

Then by Theorem A.4, for any $\delta > 0$,

$$\nabla \Phi_\varepsilon \to e_{n+1} \quad \text{ uniformly on } \mathcal{B}_1 \cap \{|x_{n+1}| > \delta\}.$$

By compactness, this still holds true if the base point is replaced by any $X_0 \in \{u = 0\}$. Thus we arrive at

**Lemma 11.6.** *For any $\delta > 0$, there exists an $L(\delta)$ such that, for any $X \in \{|\Phi| \geq L(\delta)\}$,*

$$|\nabla \Phi(X) - e_{n+1}| \leq \delta.$$

In particular, in $\{|\Phi| > L(\delta)\}$, $u$ is increasing along directions in the cone $\{e : e \cdot e_{n+1} \geq \delta\}$. Then we can proceed as in [12] to deduce that $u$ is increasing along directions in this cone everywhere in $\mathbb{R}^{n+1}$. After letting $\delta \to 0$, we deduce that for any unit vector $e$ orthogonal to $e_{n+1}$,

$$e \cdot \nabla u \geq 0, \quad -e \cdot \nabla u \geq 0, \quad \text{in } \mathbb{R}^{n+1}.$$

Thus $\frac{\partial u}{\partial x_i} \equiv 0$ in $\mathbb{R}^{n+1}$ for all $1 \leq i \leq n$. This then implies that $u$ depends only on $x_{n+1}$.

Finally, by using (11.1), it can be checked directly that $u(X) \equiv g(x_{n+1} + t)$ for some $t \in \mathbb{R}$ (see again the proof of Lemma B.2).

Next we prove Theorem 1.1. Let $u$ be a minimizing solution of (2.1) on $\mathbb{R}^{n+1}$, where $n \leq 6$. First we can use standard comparison functions to deduce an energy bound.

**Lemma 11.7.** *There exists a universal constant $C$ such that*

$$\int_{\mathcal{B}_R(X)} \left(\frac{1}{2}|\nabla u|^2 + W(u)\right) \leq CR^n \tag{11.6}$$

*for any ball $\mathcal{B}_R(X)$.*

As before, consider the blowing down sequence $u_\varepsilon$ and the associated varifold $V_\varepsilon$. By [15, Theorem 2], its limit varifold $V$ has unit density. In fact, in this case spt $\|V\| = \partial\Omega$, where $\Omega$ has minimizing perimeter [18].

Moreover, by Proposition 11.2, $\partial\Omega$ is a cone. Because $n \leq 6$, $\partial\Omega$ must be a hyperplane [23]. Then (11.2) gives

$$\lim_{R \to \infty} R^{-n} \int_{\mathcal{B}_R} \left(\frac{1}{2}|\nabla u|^2 + W(u)\right) = \|V\|(\mathcal{B}_1) = \omega_n \sigma_0.$$

Hence $u$ satisfies all of the assumptions in Theorem 11.1. By applying Theorem 11.1 we get Theorem 1.1.

## 12. The Lipschitz regularity of intermediate layers

Now we continue the proof of Theorem 9.1. In this section we first prove that $\{u_\varepsilon = t\}$ can be represented by a Lipschitz graph in the $x_{n+1}$ direction. This is a consequence of Corollary 10.5 and Lemma 12.2 below.

Before coming to Lemma 12.2, we need the following lemma, which is an easy consequence of Theorem 11.1.

**Lemma 12.1.** *Let $v$ be a solution of (2.1) in $\mathbb{R}^{n+1}$. Assume there exists a constant $\sigma$ small such that for all $r$ large,*

$$\int_{\mathcal{B}_r} [1 - (\nu \cdot e_{n+1})^2]\, |\nabla v|^2 \leq \sigma^2 r^n, \tag{12.1}$$

*and*

$$\lim_{r \to \infty} r^{-n} \int_{\mathcal{B}_r} \left(\frac{1}{2}|\nabla v|^2 + W(v)\right) \leq (1 + \tau_A)\sigma_0 \omega_n. \tag{12.2}$$

*Then there exists a constant $t \in \mathbb{R}$ and a unit vector $e$ satisfying*

$$|e - e_{n+1}| \leq C\sigma, \tag{12.3}$$

*such that $v(X) \equiv g(e \cdot X + t)$.*

*Proof.* The only thing we need to check is that (12.1) implies (12.3). This can be directly verified by substituting $u(X) \equiv g(e \cdot X + t)$ into (12.1). □

**Lemma 12.2.** *For any $b \in (0, 1)$, $R > 1$ and $\sigma > 0$ small, there exists $\bar{R} > R$ such that the following holds. Let $v$ be a solution of (2.1) in $\mathcal{B}_{\bar{R}}$, satisfying $|v(0)| \le 1 - b$, the Modica inequality (2.4) and*

$$\bar{R}^{-n} \int_{\mathcal{B}_{\bar{R}}} \left( \frac{1}{2} |\nabla v|^2 + W(v) \right) \le (1 + \tau_A) \sigma_0 \omega_n.$$

*Suppose that for any $r \in (R, \bar{R})$,*

$$\int_{\mathcal{B}_r} [1 - (\nu \cdot e_{n+1})^2] \, |\nabla v|^2 \le \sigma^2 r^n.$$

*Assume that $v > 0$ when $x_{n+1} \gg 0$. Then for $\Phi := g^{-1} \circ v$,*

$$\sup_{\mathcal{B}_R} |\nabla \Phi - e_{n+1}| \le 1/4.$$

*Proof.* Assume that, on the contrary, there exists an $R > 0$, a sequence of $R_i \to \infty$ and a sequence of solutions $v_i$ to (2.1) defined on $\mathcal{B}_{R_i}$, satisfying $|v_i(0)| \le 1 - b$, the Modica inequality (2.4),

$$R_i^{-n} \int_{\mathcal{B}_{R_i}} \left( \frac{1}{2} |\nabla v_i|^2 + W(v_i) \right) \le (1 + \tau_A) \sigma_0 \omega_n, \tag{12.4}$$

and

$$\int_{\mathcal{B}_r} [1 - (\nu_i \cdot e_{n+1})^2] \, |\nabla v_i|^2 \le \sigma^2 r^n, \qquad \forall r \in (R, R_i). \tag{12.5}$$

But

$$\sup_{\mathcal{B}_R} |\nabla \Phi_i - e_{n+1}| > 1/4. \tag{12.6}$$

Then we can assume $v_i$ converges to a smooth solution $v_\infty$ on any compact set of $\mathbb{R}^{n+1}$. By the monotonicity formula and (12.4), for any $r > 0$,

$$r^{-n} \int_{\mathcal{B}_r} \left( \frac{1}{2} |\nabla v_\infty|^2 + W(v_\infty) \right) \le (1 + \tau_A) \sigma_0 \omega_n.$$

Passing to the limit in (12.5) we also have

$$\int_{\mathcal{B}_r} [1 - (\nu_\infty \cdot e_{n+1})^2] \, |\nabla v_\infty|^2 \le \sigma^2 r^n, \qquad \forall r > R.$$

Then by the previous lemma (noting that $v_\infty(0) = \lim_{i \to \infty} v_i(0)$ and $v_\infty > \gamma$ in the part far above $\mathbb{R}^n$), $v_\infty(X) \equiv g(e \cdot X + g^{-1}(v_\infty(0)))$ for some unit vector $e$ satisfying

$$|e - e_{n+1}| \le 1/8.$$

Consequently,

$$\Phi_i(X) := g^{-1} \circ v_i(X) \to e \cdot X \quad \text{in } C^1(\mathcal{B}_R).$$

In particular, for all $i$ large,

$$\sup_{\mathcal{B}_R} |\nabla \Phi_i - e| \le 1/4.$$

This contradicts (12.6) and finishes the proof. □

We can apply this lemma to $v(X) := u_\varepsilon(X_0 + \varepsilon X)$, where $u_\varepsilon$ is as in Theorem 9.1 and $X_0 \in \{u_\varepsilon = u_\varepsilon(0)\} \cap \mathcal{B}_1$. Combined with Corollary 10.5 (provided $\sigma$, and then $\varepsilon$, are sufficiently small), this results in

**Lemma 12.3.** *For any $X_0 \in \{u_\varepsilon = u_\varepsilon(0)\} \cap \mathcal{B}_1$, $\nabla u_\varepsilon \ne 0$ in $\mathcal{B}_{K_3\varepsilon}(X_0)$ and*

$$|\nu_\varepsilon - e_{n+1}| \le 1/2 \quad \text{in } \mathcal{B}_{K_3\varepsilon}(X_0).$$

Here we only need to note that at the beginning we have assumed that $u_\varepsilon > u_\varepsilon(0)$ in $\{x_{n+1} > 1/2\} \cap \mathcal{B}_1$. Then by Lemma 10.2, $u_\varepsilon < u_\varepsilon(0)$ in $\{x_{n+1} < -1/2\} \cap \mathcal{B}_1$.

Next by combining Lemmas 10.2 and 10.4, for any $r \ge K_3\varepsilon$,

$$\{(X - X_0) \cdot e(X_0) \ge r/2\} \cap \mathcal{B}_r(X_0) \subset \{u_\varepsilon > u_\varepsilon(0)\}, \tag{12.7}$$

thanks to the continuous dependence on $r$.

By (12.7), for any $x \in B_1$, there exists a unique $x_{n+1} \in (-1, 1)$ such that $(x, x_{n+1}) \in \{u_\varepsilon = u_\varepsilon(0)\}$. Combined with the previous lemma, this then implies that

$$\{u_\varepsilon = u_\varepsilon(0)\} \cap \mathcal{B}_1 = \{x_{n+1} = h_\varepsilon(x)\}, \qquad x \in B_1.$$

Here $h_\varepsilon$ is a function with Lipschitz constant bounded by 4. (This constant can be made as small as we wish by decreasing $\varepsilon$, $\tau_A$ and $\sigma$.)

To complete the proof of Theorem 9.1, we directly apply the main result in [6]. Note that instead of the minimizing condition assumed in that paper, with our assumption (9.1) the argument still goes through.

## Appendix A. A distance type function

In this appendix, $u_\varepsilon$ always denotes a solution of (2.3) satisfying the Modica inequality (2.4). Here we introduce a distance type function associated to $u_\varepsilon$ and study its convergence as $\varepsilon \to 0$. This is perhaps well known (see for example [11] for the parabolic Allen–Cahn case). However, we have not found an exact reference, so we include some details.

Recall that $g_\varepsilon(t) := g(\varepsilon^{-1}t)$ is a one-dimensional solution of (2.2). Define

$$\Phi_\varepsilon(X) = g_\varepsilon^{-1}(u_\varepsilon(X)).$$

Then

$$-\varepsilon\Delta\Phi_\varepsilon = f(\varepsilon^{-1}\Phi_\varepsilon)(1 - |\nabla\Phi_\varepsilon|^2), \tag{A.1}$$

where $f(t) := -W'(g(t))/\sqrt{2W(g(t))} \in C^2(\mathbb{R})$. Note that

$$\lim_{t\to\pm\infty} f(t) = \pm\sqrt{W''(\pm 1)},$$

where the convergence rate is exponential.

The following result is a consequence of the Modica inequality.

**Proposition A.1.** $|\nabla\Phi_\varepsilon| \le 1$.

*Proof.* Since $u_\varepsilon = g_\varepsilon(\Phi_\varepsilon)$, we have

$$|\nabla u_\varepsilon| = g_\varepsilon'(\Phi_\varepsilon)|\nabla\Phi_\varepsilon|.$$

Thus $|\nabla\Phi_\varepsilon| \le 1$ is equivalent to

$$|\nabla u_\varepsilon|^2 \le g_\varepsilon'(\Phi_\varepsilon)^2.$$

The first integral for $g_\varepsilon$ is

$$\frac{\varepsilon}{2}(g_\varepsilon')^2 = \frac{1}{\varepsilon}W(g_\varepsilon).$$

Then the final equivalent statement is exactly the Modica inequality for $u_\varepsilon$. □

By using (A.1), the limit of $\Phi_\varepsilon$, say $\Phi_0$, can be characterized as a viscosity solution of the eikonal equation: In $\{\Phi_0 > 0\}$, $\Phi_0$ is a viscosity solution of

$$|\nabla\Phi_0|^2 - 1 = 0.$$

In $\{\Phi_0 < 0\}$, $\Phi_0$ is a viscosity solution of

$$1 - |\nabla\Phi_0|^2 = 0.$$

This is similar to the *vanishing viscosity* method (see for example Fleming–Souganidis [13]). However, here we would like to give a direct proof in our special setting.

**Proposition A.2.** *For any $\delta > 0$, there exist constants $\varepsilon_\sharp$, $\tau_\sharp$, $R_\sharp > 0$ such that the following holds. Let $u_\varepsilon$ satisfy all of the assumptions in Theorem 9.1 (with $R_A$, $\varepsilon_A$ and $\tau_A$ replaced by $R_\sharp$, $\varepsilon_\sharp$ and $\tau_\sharp$ respectively). Then there exists a set $\Omega \subset \mathcal{B}_1$ with $0 \in \partial\Omega$ and $\partial\Omega$ being a smooth minimal hypersurface such that*

$$\sup_{\mathcal{B}_1} |\Phi_\varepsilon - d_{\partial\Omega}| \le \delta. \tag{A.2}$$

*Here $d_{\partial\Omega}$ is the signed distance function to $\partial\Omega$, which is positive in $\Omega$.*

*Proof.* By Hutchinson–Tonegawa [15], the varifolds $V_\varepsilon$ converge to a stationary rectifiable varifold $V$ with integer multiplicity. Moreover, (9.1) and the monotonicity formula imply that

$$2^{-n}\|V\|(\mathcal{B}_2(X)) \le (1+2\tau_\sharp)\sigma_0\omega_n, \quad \forall X \in \mathcal{B}_2,$$

provided $R_\sharp$ has been chosen large enough.

If $\tau_\sharp$ is sufficiently small, Allard's $\varepsilon$-regularity theorem implies that $\operatorname{spt}\|V\| \cap \mathcal{B}_2$ is a smooth hypersurface and $V \cap \mathcal{B}_2$ is the standard varifold associated to this hypersurface with unit density. This hypersurface divides $\mathcal{B}_1$ into two parts (see Remark 4.5), say $\Omega$ and $\mathcal{B}_1 \setminus \Omega$. As in Remark 4.5, $u_\varepsilon$ converges to 1 uniformly in any compact subset of $\Omega$, and to $-1$ uniformly in any compact subset of $\Omega^c$.

Thus if $\varepsilon_\sharp$ is small enough, we can assume that there exists a set $\Omega$ with $0 \in \partial\Omega$ and $\partial\Omega$ being a smooth minimal hypersurface, such that

$$\operatorname{dist}_H(\{u_\varepsilon > 0\} \cap \mathcal{B}_1, \Omega \cap \mathcal{B}_1) \le \delta/8.$$

In particular,

$$\sup_{\mathcal{B}_1} |\operatorname{dist}_{\{u_\varepsilon=0\}} - d_{\partial\Omega}| \le \delta/8.$$

By Proposition A.1, in $\{u_\varepsilon > 0\} \cap \mathcal{B}_1$,

$$\Phi_\varepsilon(X) \le \operatorname{dist}_{\{u_\varepsilon=0\}}(X) \le \operatorname{dist}_{\partial\Omega}(X) + \delta.$$

Similarly, in $\{X : |\operatorname{dist}_{\partial\Omega}(X)| \le \delta/4\}$, we have $|\Phi_\varepsilon| \le \delta/2$. Thus in this part,

$$|\Phi_\varepsilon - d_{\partial\Omega}| \le |\Phi_\varepsilon| + |d_{\partial\Omega}| \le \delta.$$

In order to prove (A.2), it remains to show that if $X \in \Omega \cap \{X : \operatorname{dist}_{\partial\Omega}(X) \ge \delta/4\}$, where $\operatorname{dist}_{\{u_\varepsilon=0\}} \ge \delta/8$, then

$$\Phi_\varepsilon(X) \ge \operatorname{dist}_{\{u_\varepsilon=0\}}(X) - \delta/16.$$

However, if we have chosen $\varepsilon_\sharp$ sufficiently small (compared to $\delta$), this can be proved directly by constructing a comparison function in the ball $\mathcal{B}_{\operatorname{dist}_{\{u_\varepsilon=0\}}(X)-\delta/16}(X)$, by noting that $u_\varepsilon$ is close to 1 in this ball. □

In the following we assume that as $\varepsilon \to 0$, $\Phi_\varepsilon$ converges to a distance function $\Phi_0$ uniformly. Now we present a fact about the $C^1$ convergence of $\Phi_\varepsilon$ near a $C^1$ point of $\Phi_0$.

First we establish the uniform semiconcavity of $\Phi_\varepsilon$.

**Lemma A.3.** *Let $\Phi_\varepsilon$ satisfy (A.1) in $\mathcal{B}_1$. Assume $\Phi_\varepsilon > 1/2$ and $|\nabla\Phi_\varepsilon| \le 1$ in $\mathcal{B}_1$. Then*

$$\nabla^2\Phi_\varepsilon(0) \le C,$$

*where $C$ is a constant depending only on the dimension $n$.*

The constant $1/2$ is not essential here. It can be replaced by any positive constant.

*Proof.* We shall work in the setting where $\varepsilon = 1$ and the ball is $\mathcal{B}_R$, where $R = \varepsilon^{-1}$. For simplicity, all subscripts will be dropped.

Take a unit vector $\xi$. By directly differentiating (A.1) in the direction $\xi$, we get

$$
\begin{aligned}
-\Delta\Phi_\xi &= f'(\Phi)(1-|\nabla\Phi|^2)\Phi_\xi - 2f(\Phi)\sum_{k=1}^{n+1}\Phi_k\Phi_{\xi k},\\
-\Delta\Phi_{\xi\xi} &= f'(\Phi)(1-|\nabla\Phi|^2)\Phi_{\xi\xi} + f''(\Phi)(1-|\nabla\Phi|^2)\Phi_\xi^2\\
&\quad - 4f'(\Phi)\sum_{k=1}^{n+1}\Phi_k\Phi_{\xi k}\Phi_\xi - 2f(\Phi)\sum_{k=1}^{n+1}\Phi_k\Phi_{\xi\xi k} - 2f(\Phi)\sum_{k=1}^{n+1}\Phi_{\xi k}^2.
\end{aligned}
$$

Take an $\eta \in C_0^\infty(\mathcal{B}_{R/2})$ such that $\eta \equiv 1$ in $\mathcal{B}_{R/4}$, $0 \le \eta \le 1$, $|\nabla\eta| \le 8R^{-1}$ and $\eta^{-1}|\nabla\eta|^2 + |\Delta\eta| \le 100R^{-2}$. Denote $w := \eta\Phi_{\xi\xi}$. Since $w$ is 0 on $\partial\mathcal{B}_{R/2}$, it attains its maximum at an interior point $X_0$, where

$$
\nabla w = \eta\nabla\Phi_{\xi\xi} + \Phi_{\xi\xi}\nabla\eta = 0, \tag{A.3}
$$

and

$$
\begin{aligned}
0 \ge \Delta w &= \Delta\Phi_{\xi\xi}\eta + 2\nabla\Phi_{\xi\xi}\nabla\eta + \Phi_{\xi\xi}\Delta\eta\\
&\ge -f'(\Phi)(1-|\nabla\Phi|^2)w - f''(\Phi)(1-|\nabla\Phi|^2)\Phi_\xi^2\eta\\
&\quad + 4f'(\Phi)\sum_{k=1}^{n+1}\Phi_k\Phi_{\xi k}\Phi_\xi\eta + 2f(\Phi)\sum_{k=1}^{n+1}\Phi_k\Phi_{\xi\xi k}\eta + 2f(\Phi)\eta\sum_{k=1}^{n+1}\Phi_{\xi k}^2\\
&\quad + 2\nabla\Phi_{\xi\xi}\nabla\eta + w\eta^{-1}\Delta\eta.
\end{aligned}
$$

Substituting (A.3) into this, and applying the Cauchy inequality to the third term, we obtain

$$
\begin{aligned}
&4\frac{f'(\Phi)^2}{f(\Phi)}|\nabla\Phi|^2\Phi_\xi^2\eta + f''(\Phi)(1-|\nabla\Phi|^2)\Phi_\xi^2\eta\\
&\ge -f'(\Phi)(1-|\nabla\Phi|^2)w - 2f(\Phi)\nabla\Phi\nabla\eta\eta^{-1}w + f(\Phi)\eta^{-1}w^2 + w\eta^{-1}\Delta\eta - 2w\eta^{-2}|\nabla\eta|^2.
\end{aligned}
$$

This can be written as

$$
Aw(X_0)^2 + Bw(X_0) \le D,
$$

where $A > 0$, $B$ and $D$ are constants. From this we deduce that

$$
w(X_0) \le |B|/A + \sqrt{|D|/A}.
$$

More precisely,

$$
\begin{aligned}
w(X_0) \le{}& \frac{|f'(\Phi)|}{f(\Phi)}(1-|\nabla\Phi|^2)\eta + 2|\nabla\Phi|\,|\nabla\eta| + \frac{1}{f(\Phi)}|\Delta\eta| + \frac{2}{f(\Phi)}\eta^{-1}|\nabla\eta|^2\\
&+ \frac{1}{\sqrt{f(\Phi)}}\left(4\frac{f'(\Phi)^2}{f(\Phi)}|\nabla\Phi|^2\Phi_\xi^2\eta + f''(\Phi)(1-|\nabla\Phi|^2)\Phi_\xi^2\eta\right)^{1/2}.
\end{aligned} \tag{A.4}
$$

Since $\Phi \geq R/2$ in $\mathcal{B}_{R/2}$, by the definition of $f$ and some standard estimates on $g(t)$,

$$f(\Phi) > c \quad \text{in } \mathcal{B}_{R/2}, \qquad |f'(\Phi)| + |f''(\Phi)| \leq Ce^{-cR} \quad \text{in } \mathcal{B}_{R/2}.$$

Substituting these into (A.4), by using the condition $|\nabla\Phi| \leq 1$ and our assumptions on $\eta$, we obtain

$$\sup_{\mathcal{B}_{R/4}} \Phi_{\xi\xi} \leq w(x_0) \leq CR^{-1}.$$

Rescaling back we get the claimed estimate. □

**Theorem A.4.** *Assume that $\Phi_\varepsilon$ converges to $\Phi_0$ in $C^0(\Omega)$, where $\Omega \subset \mathbb{R}^{n+1}$ is an open set and $\Phi_0 > 0$ in $\Omega$. If $\Phi_0 \in C^1(\Omega)$, then $\Phi_\varepsilon$ converges to $\Phi_0$ in $C^1_{\mathrm{loc}}(\Omega)$.*

*Proof.* Fix an open set $\Omega_0 \subset\subset \Omega$. Take an arbitrary sequence $X_\varepsilon \in \Omega_0$ such that $X_\varepsilon \to X_0 \in \Omega_0$ as $\varepsilon \to 0$. By the uniform semiconcavity of $\Phi_\varepsilon$ in $\Omega_0$, there exists a constant $C(\Omega_0)$ such that, for all $\varepsilon > 0$,

$$\tilde{\Phi}_\varepsilon(X) := \Phi_\varepsilon(X) - C(\Omega_0)|X - X_\varepsilon|^2$$

is concave in $\Omega_0$. In particular, for any unit vector $e$ and $h < \mathrm{dist}(\Omega_0, \partial\Omega)$,

$$\tilde{\Phi}_\varepsilon(X_\varepsilon + he) \leq \tilde{\Phi}_\varepsilon(X_\varepsilon) + h\nabla\Phi_\varepsilon(X_\varepsilon)e. \tag{A.5}$$

Because $|\nabla\Phi_\varepsilon(X_\varepsilon)| \leq 1$, assume $\nabla\Phi_\varepsilon(X_\varepsilon)$ converges to a vector $\xi$. By the uniform convergence of $\Phi_\varepsilon$ in $\Omega$, passing to the limit in (A.5) leads to

$$\Phi_0(X_0 + he) \leq \Phi_0(X_0) + h(\xi \cdot e) + C(\Omega_0)h^2/2, \qquad \forall h > 0.$$

Since $\Phi_0$ is differentiable at $X_0$, letting $h \to 0$ gives $\xi = \nabla\Phi_0(X_0)$. From this argument we get the uniform convergence of $\nabla\Phi_\varepsilon$ in $\Omega_0$. □

## Appendix B. Several technical results

Here we collect some technical results used in this paper.

The first one is an exponential decay estimate. It has been used in many places and can be proved by various methods (see for example [6, Section 2]), so here we only state the result.

**Lemma B.1.** *If in the ball $\mathcal{B}_{2R}(0)$, $u \in C^2$ satisfies*

$$\Delta u \geq Mu, \qquad 0 \leq u \leq 1, \tag{B.1}$$

*then*

$$\sup_{\mathcal{B}_R(0)} u \leq Ce^{-cRM^{1/2}}.$$

The next one gives a control of the discrepancy using the excess.

**Lemma B.2.** *Given $M$, $L$ and $\tau > 0$, there exist constants $\delta > 0$ and $R(M, L, \tau) > 2L$ such that the following holds. Suppose that $u$ is a solution of (2.1) in $\mathcal{B}_R$ with $R \geq R(M, L, \tau)$, satisfying*

$$R^{-n}\int_{\mathcal{B}_R}\left(\frac{1}{2}|\nabla u|^2 + W(u)\right) \leq M, \qquad (2L)^{-n}\int_{\mathcal{B}_{2L}}[1-(\nu\cdot e_{n+1})^2]\,|\nabla u|^2 \leq \delta.$$

*Then*

$$L^{-n}\int_{\mathcal{B}_L}\left|W(u) - \frac{1}{2}|\nabla u|^2\right| \leq \tau.$$

*Proof.* Assume that, on the contrary, there exist constants $M$ and $\tau$, and a sequence of solutions $u_i$, defined in $\mathcal{B}_{R_i}$ with $R_i \to \infty$, satisfying

$$R_i^{-n}\int_{\mathcal{B}_{R_i}}\left(\frac{1}{2}|\nabla u_i|^2 + W(u_i)\right) \leq M, \tag{B.2}$$

$$(2L)^{-n}\int_{\mathcal{B}_{2L}}[1-(\nu_i\cdot e_{n+1})^2]\,|\nabla u_i|^2 \to 0, \tag{B.3}$$

but

$$L^{-n}\int_{\mathcal{B}_L}\left|W(u_i) - \frac{1}{2}|\nabla u_i|^2\right| \geq \tau. \tag{B.4}$$

Denote the limit of $u_i$ by $u_\infty$. By passing to the limit in (B.3) and the unique continuation principle, $u_\infty$ depends only on the $x_{n+1}$ variable. By the monotonicity formula, for any $R \in (0, R_i)$,

$$R^{-n}\int_{\mathcal{B}_R}\left(\frac{1}{2}|\nabla u_i|^2 + W(u_i)\right) \leq M.$$

This also holds for $u_\infty$ by passing to the limit. Because $u_\infty$ is one-dimensional, this implies

$$\int_{-\infty}^{\infty}\left(\frac{1}{2}\left|\frac{du_\infty}{dx_{n+1}}\right|^2 + W(u_\infty)\right) \leq M.$$

Note that except the heteroclinic solution $g$, all the other solutions of (2.1) in $\mathbb{R}^1$ are periodic, and hence their energy on $\mathbb{R}$ is infinite. From this fact we deduce that $u_\infty \equiv g(x_{n+1} + t)$ for some constant $t \in \mathbb{R}$. Hence

$$W(u_\infty) \equiv \frac{1}{2}\left|\frac{du_\infty}{dx_{n+1}}\right|^2.$$

Consequently,

$$\lim_{i\to\infty}\int_{\mathcal{B}_L}\left|W(u_i) - \frac{1}{2}|\nabla u_i|^2\right| = 0.$$

However, this contradicts (B.4). □

The following result says the energy $\varepsilon|\nabla u_\varepsilon|^2$ is mostly concentrated on the transition part $\{|u_\varepsilon| \leq 1 - b\}$.

**Lemma B.3.** *Let $u_\varepsilon$ be a solution of (2.2) defined in $\mathcal{B}_2$, satisfying the Modica inequality and*

$$\int_{\mathcal{B}_2}\left(\frac{\varepsilon}{2}|\nabla u_\varepsilon|^2+\frac{1}{\varepsilon}W(u_\varepsilon)\right)\le M.$$

*For any $\delta>0$, there exists a constant $b\in(0,1)$ such that*

$$\int_{\{|u_\varepsilon|>1-b\}\cap\mathcal{B}_1}\varepsilon|\nabla u_\varepsilon|^2\le\delta.$$

This is essentially [15, Proposition 5.1]. We just need to note that by the Modica inequality, we can bound $\varepsilon|\nabla u_\varepsilon|^2$ by $\varepsilon^{-1}W(u_\varepsilon)$.

Here we give a different proof. More precisely, we prove

**Lemma B.4.** *Let $u_\varepsilon$ be as in the previous lemma. For any $1\le i\le n+1$ and $\delta>0$, there exists a constant $b\in(0,1)$ such that*

$$\int_{\{|u_\varepsilon|>1-b\}\cap\mathcal{B}_1}\varepsilon\left(\frac{\partial u_\varepsilon}{\partial x_i}\right)^2\le\delta\int_{\{|u_\varepsilon|<1-b\}\cap\mathcal{B}_2}\varepsilon\left(\frac{\partial u_\varepsilon}{\partial x_i}\right)^2.$$

*Proof.* For simplicity, denote $\xi:=\varepsilon\big(\frac{\partial u_\varepsilon}{\partial x_i}\big)^2$, which satisfies

$$\Delta\xi\ge\frac{c}{\varepsilon^2}\xi\quad\text{ in }\{|u_\varepsilon|>\gamma\}.\tag{B.5}$$

By the gradient bound on the distance type function $\Phi_\varepsilon$, we know that for any $M>0$ there exists $0<b<1$ such that

$$|X_1-X_2|\ge M\varepsilon,\quad\forall X_1\in\{|u_\varepsilon|<1-2b\},\ X_2\in\{|u_\varepsilon|>1-b\}.$$

In other words, for any $X\in\{|u_\varepsilon|>1-b\}$, $B_{M\varepsilon}(X)\subset\{|u_\varepsilon|>1-2b\}$.

By (B.5), if $1-2b>\gamma$, then for any $X\in\{|u_\varepsilon|>1-b\}$,

- because $\xi$ is subharmonic,

$$\sup_{\mathcal{B}_{M\varepsilon/2}(X)}\xi\le CM^{-n-1}\varepsilon^{-n-1}\int_{\mathcal{B}_{M\varepsilon}(X)}\xi(Y)\,dY;$$

- by Lemma B.1,

$$\xi(X)\le Ce^{-cM}\sup_{\mathcal{B}_{M\varepsilon/2}(X)}\xi.$$

Thus

$$\xi(X)\le Ce^{-cM}M^{-n-1}\varepsilon^{-n-1}\int_{\mathcal{B}_{M\varepsilon}(X)}\xi(Y)\,dY.$$

Integrating this on $\{|u_\varepsilon| > 1-b\}$ and then using the Fubini theorem, we obtain

$$\begin{aligned}\int_{\{|u_\varepsilon|>1-b\}\cap\mathcal{B}_1} \xi(X)\,dX &\le Ce^{-cM}M^{-n-1}\varepsilon^{-n-1}\int_{\{|u_\varepsilon|>1-b\}\cap\mathcal{B}_1}\int_{B_{M\varepsilon}(0)} \xi(X+Y)\,dY\,dX\\ &= Ce^{-cM}M^{-n-1}\varepsilon^{-n-1}\int_{B_{M\varepsilon}(0)}\int_{\{|u_\varepsilon|>1-b\}\cap\mathcal{B}_1} \xi(X+Y)\,dX\,dY\\ &\le Ce^{-cM}M^{-n-1}\varepsilon^{-n-1}\int_{B_{M\varepsilon}(0)}\int_{\{|u_\varepsilon|>1-2b\}\cap\mathcal{B}_2} \xi(X)\,dX\,dY\\ &\le Ce^{-cM}\int_{\{|u_\varepsilon|>1-2b\}\cap\mathcal{B}_1} \xi(X)\,dX.\end{aligned}$$

Hence by choosing $b$ small enough, which implies that $M$ is sufficiently large, we get the claimed estimate. □

Finally, we give a lower bound of the energy in balls.

**Lemma B.5.** *For any $b \in (0,1)$ and $M > 0$, there exist constants $c(M,b)$ and $R(M,b)$ such that the following holds. Assume $u$ is a solution of (2.1) in $\mathcal{B}_R$ where $R \ge 2R(M,b)$, satisfying $|u(0)| \le 1-b$, the Modica inequality and the energy bound*

$$\int_{\mathcal{B}_R}\left(\frac{1}{2}|\nabla u|^2 + W(u)\right) \le MR^n.$$

*Then for any $r \in [1, R/2]$,*

$$\int_{\mathcal{B}_r} |\nabla u|^2 \ge c(M,b)r^n. \tag{B.6}$$

By the monotonicity formula, it is easy to get a constant $c(b)$ such that

$$\int_{\mathcal{B}_r}\left(\frac{1}{2}|\nabla u|^2 + W(u)\right) \ge c(b)r^n, \quad \forall 1 < r < R. \tag{B.7}$$

However, this is weaker than our statement.

*Proof of Lemma B.5.* We first prove that, under the assumptions of this lemma (with a different constant $R_1(M,b)$),

$$\int_{\mathcal{B}_{R/2}} |\nabla u|^2 \ge c(M,b)R^n \tag{B.8}$$

if $R \ge R_1(M,b)$.

Assume this is not true, that is, the claimed $R_1(M,b)$ does not exist. Then there exists an $M > 0$ and a sequence of $u_i$, which are solutions of (2.2) in $\mathcal{B}_{R_i}$ where $R_i \to \infty$, satisfying $|u_i(0)| \le 1-b$, the Modica inequality and the energy bound

$$\int_{\mathcal{B}_{R_i}}\left(\frac{1}{2}|\nabla u_i|^2 + W(u_i)\right) \le MR_i^n,$$

but

$$R_i^{-n}\int_{\mathcal{B}_{R_i/2}}|\nabla u|^2\to 0. \tag{B.9}$$

Let $\varepsilon_i=R_i^{-1}$ and $u_{\varepsilon_i}(X):=u_i(R_iX)$. Then

$$\int_{\mathcal{B}_1}\left(\frac{\varepsilon_i}{2}|\nabla u_{\varepsilon_i}|^2+\frac{1}{\varepsilon_i}W(u_{\varepsilon_i})\right)\le M,$$

and

$$\int_{\mathcal{B}_{1/2}}\varepsilon_i|\nabla u_{\varepsilon_i}|^2\to 0. \tag{B.10}$$

By the main result in [15], $\varepsilon_i|\nabla u_i|^2dX\rightharpoonup\mu$ as measures, where $\mu$ is a positive Radon measure (in fact, the weight measure associated to the limit varifold, as in Section 4). Moreover, $u_{\varepsilon_i}\to\pm1$ locally uniformly outside $\operatorname{spt}\mu$. However, (B.10) obviously implies that $\mu(\mathcal{B}_{1/2})=0$. Thus for all $i$ large, $|u_{\varepsilon_i}|>1-b$ in $\mathcal{B}_{1/2}$. This contradicts our assumption that $|u_{\varepsilon_i}(0)|\le 1-b$, and proves (B.8).

By the monotonicity formula (recall that we have assumed the Modica inequality), for any $r\in(1,R)$,

$$\int_{\mathcal{B}_r}\left(\frac{1}{2}|\nabla u|^2+W(u)\right)\le Mr^n. \tag{B.11}$$

Thus the above discussion covers the case $r\in[R_1(M,b),R]$ in (B.6), that is, (B.8) holds for every $r\in[R_1(M,b),R]$.

For the remaining case, we only need to note that it is impossible to have $u\equiv u(0)$, because otherwise

$$\int_{\mathcal{B}_r}\left(\frac{1}{2}|\nabla u|^2+W(u)\right)=\omega_{n+1}W(u(0))r^{n+1}\ge Mr^n,$$

provided $r\ge M/\omega_{n+1}(\inf_{s\in[-1+b,1-b]}W(s))$. Then it can be directly verified that

$$\int_{\mathcal{B}_1}|\nabla u|^2\ge c(M),$$

by using (B.11) with $r=M/\omega_{n+1}(\inf_{s\in[-1+b,1-b]}W(s))$.

Choosing $R(M,b)=\max\{R_1(M,b),M/\omega_{n+1}(\inf_{s\in[-1+b,1-b]}W(s))\}$ we finish the proof. □

*Acknowledgments.* The author is grateful to the referees for their careful reading and useful suggestions. This work was supported by NSFC No. 11301522.

## References


[1] Alberti, G., Ambrosio, A., Cabré, X.: On a long standing conjecture of E. De Giorgi: symmetry in 3D for general nonlinearities and a local minimality property. Acta Appl. Math. **65**, 9–33 (2001) Zbl 1121.35312 MR 1843784

[2] Allard, W.: On the first variation of a varifold. Ann. of Math. **95**, 417–491 (1972) Zbl 0252.49028 MR 0307015

[3] Ambrosio, A., Cabré, X.: Entire solutions of semilinear elliptic equations in $\mathbb{R}^3$ and a conjecture of De Giorgi. J. Amer. Math. Soc. **13**, 725–739 (2000) Zbl 0968.35041 MR 1775735

[4] Barlow, M., Bass, R., Gui, C.: The Liouville property and a conjecture of De Giorgi. Comm. Pure Appl. Math. **53**, 1007–1038 (2000) Zbl 1072.35526 MR 1755949

[5] Caffarelli, L., Córdoba, A.: An elementary regularity theory of minimal surfaces. Differential Integral Equations **6**, 1–13 (1995) Zbl 0783.35008 MR 1190161

[6] Caffarelli, L., Córdoba, A.: Phase transitions: Uniform regularity of the intermediate layers. J. Reine Angew. Math. **593**, 209–235 (2006) Zbl 1090.49019 MR 2227143

[7] De Giorgi, E.: Frontiere orientate di misura minima. Editrice tecnico scientifica, Pisa (1961)

[8] De Giorgi, E.: Convergence problems for functionals and operators. In: Proc. Int. Meeting on Recent Methods in Non-Linear Analysis (Rome, 1978), Pitagora, Bologna, 131–188 (1979) Zbl 0405.49001 MR 0533166

[9] del Pino, M., Kowalczyk, M., Wei, J., Yang, J.: Interface foliation near minimal submanifolds in Riemannian manifolds with positive Ricci curvature. Geom. Funct. Anal. **20**, 918–957 (2010) Zbl 1213.35219 MR 2729281

[10] del Pino, M., Kowalczyk, M., Wei, J.: On De Giorgi's conjecture in dimension $N \geq 9$. Ann. of Math. **174**, 1485–1569 (2011) Zbl 1238.35019 MR 2846486

[11] Evans, L. C., Soner, H., Souganidis, P.: Phase transitions and generalized motion by mean curvature. Comm. Pure Appl. Math. **45**, 1097–1123 (1992) Zbl 0801.35045 MR 1177477

[12] Farina, A.: Symmetry for solutions of semilinear elliptic equations in $\mathbb{R}^N$ and related conjectures. Ricerche Mat. **48** (suppl.), 129–154 (1999) Zbl 0940.35084 MR 1765681

[13] Fleming, W. H., Souganidis, P.: Asymptotic series and the method of vanishing viscosity. Indiana Univ. Math. J. **35**, 425–448 (1986) Zbl 0573.35034 MR 0833404

[14] Ghoussoub, N., Gui, C.: On a conjecture of De Giorgi and some related problems. Math. Ann. **311**, 481–491 (1998) Zbl 0918.35046 MR 1637919

[15] Hutchinson, J., Tonegawa, Y.: Convergence of phase interfaces in the van der Waals–Cahn–Hilliard theory. Calc. Var. Partial Differential Equations **10**, 49–84 (2000) Zbl 1070.49026 MR 1803974

[16] Lin, F., Yang, X.: Geometric Measure Theory: an Introduction. Adv. Math. (Beijing/Boston), Science Press, Beijing, and Int. Press, Boston (2002) Zbl 1074.49011 MR 2030862

[17] Modica, L.: A gradient bound and a Liouville theorem for nonlinear Poisson equations. Comm. Pure Appl. Math. **38**, 679–684 (1985) Zbl 0612.35051 MR 0803255

[18] Modica, L.: The gradient theory of phase transitions and the minimal interphase criterium. Arch. Ration. Mech. Anal. **98**, 125–142 (1987) Zbl 0616.76004 MR 0866718

[19] Modica, L.: Monotonicity of the energy for entire solutions of semilinear elliptic equations. In: Partial Differential Equations and the Calculus of Variations, Vol. II, Progr. Nonlinear Differential Equations Appl. 2, Birkhäuser Boston, Boston, MA, 843–850 (1989) Zbl 0699.35082 MR 1034031

[20] Savin, O.: Regularity of flat level sets in phase transitions. Ann. of Math. **169**, 41–78 (2009) Zbl 1180.35499 MR 2480601

[21] Savin, O.: Small perturbation solutions for elliptic equations, Comm. Partial Differential Equations **32**, 557–578 (2007) Zbl 1221.35154 MR 2334822

[22] Savin, O.: Entire solutions to a class of fully nonlinear elliptic equations, Ann. Scuola Norm. Sup. Pisa Sci. (4) **7**, 369–405 (2008) Zbl 1181.35111 MR 2466434

[23] Simons, J.: Minimal varieties in Riemannian manifolds. Ann. of Math. **88**, 62–105 (1968) Zbl 0181.49702 MR 0233295

[24] Tonegawa, Y.: Note on the small-scale structure of the phase boundaries. RIMS Kôkyûroku **1323**, 33–36 (2003)

[25] Valdinoci, E., Sciunzi, B., Savin, O.: Flat level set regularity of $p$-Laplace phase transitions. Mem. Amer. Math. Soc. **182**, no. 858, vi + 144 pp. (2006) Zbl 1138.35029 MR 2228294

[26] Wang, K.: Harmonic approximation and improvement of flatness in a singular perturbation problem. Manuscripta Math. **146**, 281–298 (2015) Zbl 1310.35023 MR 3294427