The inhomogeneous Allen–Cahn equation and the existence of prescribed-mean-curvature hypersurfaces

Costante Bellettini∗ and Neshan Wickramasekera

Abstract

We prove that for any given compact Riemannian manifold $N$ of dimension $n+1 \geq 3$ and any non-negative function $g$ of class $C^{0,\alpha}$ ($0 < \alpha < 1$) on $N$, there exists a two-sided hypersurface $M \subset N$ of class $C^{2,\alpha}$ such that the mean curvature vector of $M$, with respect to an appropriate choice of continuous unit normal $\nu$, is given by $g\nu$. The hypersurface $M$ is quasi-embedded, i.e. $M$ is the image of an immersion and near every point where $M$ is not embedded, it is the union of two embedded disks intersecting tangentially with each disk lying on one side of the other; moreover its singular set $\Sigma = M \setminus M$ is empty if $2 \leq n \leq 6$, discrete if $n = 7$ and satisfies $\mathcal{H}^{n-7+\gamma}(\Sigma) = 0$ for every $\gamma > 0$ if $n \geq 8$. If $g > 0$ then $M$ is the boundary of a Caccioppoli set.

Our proof of this theorem is PDE theoretic, and considers first the case that $g$ is a positive function of class $C^{1,1}$. The theorem for non-negative $g$ of class $C^{0,\alpha}$ follows by approximation, based on the estimates we establish. In the case of positive $g$ of class $C^{1,1}$, the argument is based on: (i) a construction, using a simple mountain pass lemma, of uniformly energy bounded min-max solutions $u_\varepsilon$ to inhomogeneous Allen–Cahn equations $-\varepsilon \Delta u + \varepsilon^{-1} W'(u) = \sigma g$ for small $\varepsilon > 0$, where $\sigma$ is a fixed normalising constant and $W$ is a fixed double-well potential, and (ii) a proof of regularity of any limit-varifold $V$ that arises from a sequence $(u_{\varepsilon_j})$ of such solutions—in fact from any Morse-index bounded, energy bounded sequence of solutions—for $\varepsilon_j \to 0^+$. Parts of $V$ may be minimal (i.e. have zero mean curvature), but regularity of $V$ ensures that minimal portions, if there are any, can be smoothly excised. The remaining part of $V$ is a mean-curvature $g$ hypersurface $M$ as desired, unless $V$ is entirely supported on a minimal hypersurface, a possibility we do not rule out. If this possibility arises and $V$ is a minimal hypersurface $M_0$ (with multiplicity, necessarily, even integer valued and locally constant), we appeal to PDE techniques again. Working at the level of the Allen–Cahn approximation, we use a semi-linear gradient flow with carefully chosen initial data constructed from $M_0$ to produce, for each $j$, a stable solution $v_{\varepsilon_j}$ to the above Allen–Cahn equation satisfied by $u_{\varepsilon_j}$, in such a way that the sequence $(v_{\varepsilon_j})$ leads to a non-trivial limit-varifold that is not entirely supported on a minimal hypersurface. This new limit-varifold is regular by (ii), and after excision of its minimal portions we again obtain a mean-curvature $g$ hypersurface $M$ as desired.

Contents

1 Introduction 2

2 The case $g \equiv 0$, new difficulties when $g > 0$ and an outline of the proof 7

∗Partially supported by the EPSRC under the grant EP/S005641/1.
3 Preliminaries: Allen–Cahn solutions and limit \((g, 0)\)-varifolds

4 Regularity of limit \((g, 0)\)-varifolds

5 A min-max construction of Allen–Cahn solutions

6 Proof of the existence theorem

7 Extension to the case of non-negative \(g\) of class \(C^{0, \alpha}\)

1 Introduction

Our main purpose here is to establish the following theorem:

**Theorem 1.1.** Let \(N\) be a compact Riemannian manifold of dimension \(n + 1\) with \(n \geq 2\), and let \(g : N \to [0, \infty)\) be a (non-negative) function of class \(C^{0, \alpha}\) for some \(\alpha \in (0, 1)\). There exists a two-sided quasi-embedded immersed hypersurface \(M \subset N\) of class \(C^{2, \alpha}\) with \(\text{dim}_{\mathbb{H}} (\overline{M} \setminus M) \leq n - 7\) if \(n \geq 7\), \(\overline{M} \setminus M\) discrete if \(n = 7\) and \(\overline{M} \setminus M = \emptyset\) if \(2 \leq n \leq 6\), such that the mean curvature \(H_M\) of \(M\) is given by \(H_M = g\nu\) for some choice of continuous unit normal \(\nu\) to \(M\).

**Remark 1.1.** Here a quasi-embedded hypersurface is the image of a very special type of immersion that is natural in the context of prescribed-mean-curvature hypersurfaces. Specifically, in the case of \(M\) in Theorem 1.1 quasi-embedded means that for every \(y \in M\) there exists a neighbourhood of \(y\) (in \(N\)) in which \(M\) is either an embedded \(C^{2, \alpha}\) disk, or the union of two embedded \(C^{2, \alpha}\) disks with each disk lying on one side of the other.

**Remark 1.2.** Theorem 1.1 remains valid if the prescribing function for mean curvature is a non-positive \(C^{0, \alpha}\) function rather than a non-negative one. For if \(\tilde{g} : N \to (-\infty, 0)\) is of class \(C^{0, \alpha}\) with \(\alpha \in (0, 1)\), we may apply Theorem 1.1 with \(g = -\tilde{g}\). Reversing the choice of unit normal on the resulting quasi-embedded hypersurface provides the resolution of the existence question for \(\tilde{g}\).

**Remark 1.3.** Once \(C^2\) regularity is obtained for a two-sided immersion with mean curvature \(g\nu\), higher regularity follows by standard elliptic theory, depending on the regularity assumed on \(g\). In particular, if \(g \in C^{k, \alpha}\) then the immersion is \(C^{k+2, \alpha}\).

**Remark 1.4.** We will first prove the theorem assuming \(g \in C^{1,1}(N)\) and \(g > 0\). The theorem for non-negative \(g \in C^{0, \alpha}(N)\) will follow from an approximation argument using the estimates we establish. In the case \(g > 0\), we have the additional fact
that |M| = |∂*E| for a Caccioppoli set E ⊂ N (where |T| denotes the multiplicity 1 varifold associated with an n-rectifiable set T) and that at every non-embedded point y ∈ M, the two disks (as in Remark 1.1) whose union is M in a neighborhood of y intersect (only tangentially) along a closed set containing y and contained in an (n − 1)-dimensional embedded C¹ submanifolds (see [BelWic-1, Remark 2.6] and [BelWic-2] Remark 3.3).

In view of Remark 1.4 from now on until the end of Section 6.10 we shall assume that g ∈ C¹,¹(N) and g > 0 unless stated otherwise explicitly.

Our approach to the above theorem is by means of an inhomogeneous Allen–Cahn approximation scheme. Results from the general varifold regularity theory developed in our recent work [BelWic-1], [BelWic-2] play a key role in this approach.

In the work of Guaraco ([Gua18]), a homogeneous Allen–Cahn approximation process is employed to provide a strikingly simple PDE theoretic proof of the case g ≡ 0 of Theorem 1.1, i.e. the existence of minimal hypersurfaces. This approach employs a classical PDE min-max construction and is based on an earlier advance in the regularity theory for stable, area-stationary varifolds ([Wic14]) and its application to the limit varifolds arising from the Allen–Cahn approximation ([HutTon00], [TonWic12]).

For the case g > 0, our starting point is still a classical PDE min-max construction of solutions of the relevant inhomogeneous Allen–Cahn equation. In contrast to the case g ≡ 0 however, two new phenomena arise in the case g > 0. One is the possibility that two regular pieces of the corresponding limit varifold may touch without coinciding on a set as large as having codimension 1 (leading to the quasi-embeddedness conclusion in Theorem 1.1 as opposed to embeddedness). Hence a regularity theory with sufficiently general hypotheses that allow this possibility is needed. The second new aspect is that the inhomogeneous Allen–Cahn approximation process may lead to a loss of mean curvature information in the limit. The proof of the theorem has to account for the difficulties arising from this latter phenomenon, including the possibility that the min-max process may fail to produce the desired prescribed-mean-curvature hypersurface and may instead produce a minimal hypersurface. Nonetheless, the spirit of simplicity afforded by the PDE techniques, already seen in the minimal hypersurface case, continues in our proof of Theorem 1.1. Indeed, a general GMT framework ([BelWic-1], [BelWic-2], together with an extension of techniques from [BelWic-1]) enables us to prove regularity of a limit varifold, up to a lower dimensional singular set, allowing: (i) for sheets to touch without coinciding, and (ii) for the mean curvature to become zero somewhere (or everywhere) in the limit varifold; given this, the possible failure of the min-max process to produce the desired hypersurface can be handled by employing PDE principles in semilinear parabolic theory. More precisely, in the event that the min-max construction leads to a limit varifold that is a minimal hypersurface M₀, we use a semilinear gradient flow (namely, an Allen–Cahn flow) to obtain the desired prescribed mean curvature hypersurface M. The initial conditions for this flow are constructed out of M₀, capitalising on the lower dimensionality of its singular set.

In broad terms, our proof of Theorem 1.1 can be described as follows. We consider, for small positive ε, the functional (energy) \( \mathcal{F}_{\varepsilon, \sigma} : W^{1,2}(N) \to \mathbb{R} \) defined by

\[
\mathcal{F}_{\varepsilon, \sigma}(u) = \int_N \varepsilon \frac{|\nabla u|^2}{2} + \int_N \frac{W(u)}{\varepsilon} - \int_N \sigma g u,
\]

where g : N → \( \mathbb{R} \) is a positive C¹,¹ function and σ is a fixed positive normalising constant depending only on W. The first two terms of \( \mathcal{F}_{\varepsilon, \sigma} \) add up to give the
usual Allen–Cahn energy

\[ \mathcal{E}_\varepsilon(u) = \int_N \varepsilon \frac{|\nabla u|^2}{2} + \int_N \frac{W(u)}{\varepsilon}. \]

The “double well” potential \( W : \mathbb{R} \to [0, \infty) \) is a (non-negative) function of class \( C^2 \) having precisely two non-degenerate minima at \( \pm 1 \) with value 0 and appropriate (in fact quadratic) growth outside \([-2, 2]\).

The (relatively straightforward) starting point is to choose a sequence \( \varepsilon_j \to 0^+ \) and construct, in \( W^{1,2}(N) \), a sequence of 1-parameter min-max critical points \( u_{\varepsilon_j} \) of \( \mathcal{F}_{\varepsilon_j, \sigma g} \) such that the corresponding sequence of Allen–Cahn energies \( \mathcal{E}_{\varepsilon_j}(u_{\varepsilon_j}) \) is uniformly bounded from above and below by positive numbers. This is done by applying a mountain pass lemma based on the fact that \( \mathcal{F}_{\varepsilon_j, \sigma g} \) satisfies the Palais–Smale condition. The Morse index of \( u_{\varepsilon_j} \) will then automatically be \( \leq 1 \).

By general principles, the energy bounds imply the existence of a non-zero Radon measure \( \mu \) on \( N \) and a function \( u_{\infty} \in BV(N) \) with \( u_{\infty}(x) = \pm 1 \) for a.e. \( x \in N \) such that after passing to a subsequence without relabeling,

\[ (2\sigma)^{-1} \left( \varepsilon_j \frac{1}{2} \nabla u_{\varepsilon_j}^2 + \varepsilon_j^{-1} W(u_{\varepsilon_j}) \right) \, d\text{vol}_N \to \mu \]

weakly and \( u_{\varepsilon_j} \to u_{\infty} \) in \( L^1 \). Set \( E = \{ u_{\infty} = 1 \} \) and note that \( E \) is a Caccioppoli set in \( N \) with its reduced boundary \( \partial^* E \subset \text{spt} \mu \).

By the combined results and ideas in the work of Ilmanen ([Ilm93]), Schätzle ([Sch04]), Hutchinson–Tonegawa ([HutTon00]), Röger–Tonegawa ([RogTon08]) and Tonegawa ([Ton05-2]) (see Theorem 3.1 below), this limit measure \( \mu \) is the weight measure \( ||V|| \) of an integral \( n \)-varifold \( V \) on \( N \) with first variation \( \delta V = -H_V ||V|| \) and with \( H_V \), the generalized mean curvature of \( V \), locally bounded. Furthermore, by these works we know that if \( E = N \) then \( V = 0 \) (this uses \( g > 0 \)), and also the following:

(a) if \( V \) is of multiplicity 1, then \( E \) is non-empty, \( V = |\partial^* E| \) (the multiplicity 1 varifold associated with the reduced boundary \( \partial^* E \)), and \( H_V = g \nu \mathcal{H}^n \) a.e. where \( \nu \) is the unit normal to \( \partial^* E \) pointing into \( E \); in other words, in this case, \( V \) is a critical point of the functional \( A - \text{Vol}_g \) where \( A \) is the area of \( \partial^* E \) and \( \text{Vol}_g \) is the enclosed \( g \) volume \( \int_E g \, d\mathcal{H}^{n+1} \).

(b) if \( g \equiv 0 \) then regardless of whether \( V \) is of multiplicity 1, \( H_V = 0 \), i.e. \( V \) is a critical point of \( A \).

Putting aside for the moment the obvious regularity questions, when \( g \neq 0 \) and in the absence of the multiplicity 1 assumption on \( V \) as in (a), another key difficulty arises; that is, a combination of (a) and (b) may occur, in the sense that \( V \) may consist both of a higher multiplicity part \( V_0 \) on which \( H_V = 0 \) a.e. and the multiplicity 1 part \( V_\theta = |\partial^* E| \) on which \( H_V = g \nu \) a.e. The two parts \( V_0 \) and \( V_\theta \) may merge together in a potentially very complicated manner on not too small a set (see Figure 1 in Section 2). This would prevent their separation from each other as critical points, in \( N \), of \( A \) and \( A - \text{Vol}_g \), respectively.

In light of this, to prove Theorem 1.1 one approach would be to first rule out higher multiplicity in \( V \). If this succeeds we would be in case (a). As regards regularity in this case, the central difficulty would be that even though \( V \) may be of multiplicity 1, a priori \( V \) may still have a large (but \( \mathcal{H}^n \)-null) set of singularities with higher multiplicity planar tangent cones. However this possibility can be ruled out by a direct application of [BelWic-2, Theorem 5.1], [BelWic-2, Theorem 9.1] and [BelWic-2, Theorem 6.4]. From there, it is in fact a few easy steps to the full regularity conclusion claimed in Theorem 1.1.

In the generality in which we work here though, showing that no part of \( V \) can develop higher multiplicity appears to be difficult. In the absence of Morse
index control in the limiting process higher multiplicity can in fact occur (see the example in [HutTon98, Section 6.3]; see also [Sch01] for a related example, which seems to have originally appeared in [Gro98], where a sequence of CMC hypersurfaces with fixed mean curvature converges to a multiplicity 2 plane). In any case we here follow a strategy that avoids altogether the need to rule out higher multiplicity.

At the core of this strategy is establishing complete regularity of $V$ allowing multiplicity. For this we employ the regularity theory in our recent work ([BelWic-1], [BelWic-2]), as well as extensions of ideas therein to account for the non-variational nature of $V$, i.e. the presence of both a region with $H_V = 0$ and a region with $H_V = g
u$. This first step is perhaps of interest independently of Theorem 1.1 and produces the following general result concerning Allen–Cahn hypersurfaces with fixed mean curvature converges to a multiplicity 2 plane). In which seems to have originally appeared in [Gro98], where a sequence of CMC the example in [HutTon98, Section 6.3]; see also [Sch01] for a related example, index control in the limiting process higher multiplicity can in fact occur (see

**Theorem 1.2** (Theorem 4.1 below). Let $g \in C^{1,1}(N)$ with $g > 0$. Let $\mu$ be a non-zero Radon measure on $N$ such that

$$
\mu_j = (2\sigma)^{-1} \left( \frac{\varepsilon_j}{2} |\nabla u_{\varepsilon_j}|^2 + \varepsilon_j^{-1} W(u_{\varepsilon_j}) \right) \mathrm{dvol}_N \rightarrow \mu
$$

where $\sigma = \int_{-1}^1 \sqrt{\frac{W(0)}{2}} \mathrm{ds}$, $\varepsilon_j \rightarrow 0^+$ and $u_{\varepsilon_j}$ is a critical point of $F_{\varepsilon_j, \sigma}$ whose Morse index is bounded independently of $j$. Then

$$
\mu = \mathcal{H}^n \mathbf{L} M_0 + q\mathcal{H}^n \mathbf{L} M_0
$$

where:

(a) $M_0$ is a (possibly empty) immersed, quasi-embedded hypersurface bounding a Caccioppoli set $E \subset N$ such that its mean curvature $H_{M_0} = g
u$ where $\nu$ is the unit normal pointing into $E$, and such that its singular set $\overline{M_0} \setminus M_0$ is empty if $2 \leq n \leq 6$, discrete if $n = 7$ and has Hausdorff dimension $\leq n - 7$ if $n \geq 7$; in fact $E = \{u_\infty = 1\}$ where $u_\infty$ is the $L^1_{\text{loc}}$ limit of $u_{\varepsilon_j}$, with $u_\infty(x) = \pm 1$ for a.e. $x \in N$ and $E \neq N$ (since $\mu \neq 0$).

(b) $M_0$ is a (possibly empty) embedded minimal hypersurface of $N$ with its singular set $\overline{M_0} \setminus M_0$ empty if $2 \leq n \leq 6$, discrete if $n = 7$ and having Hausdorff dimension $\leq n - 7$ if $n \geq 7$; the multiplicity function $q$ is equal to a constant positive even integer on each connected component of $M_0$, and the multiplicity 1 varifold $|M_0|$ is stationary in $N$;

Additionally, we also have that $\overline{M_0} \cap \overline{M_0} = (M_0 \cap M_0) \cup ((\overline{M_0} \setminus M_0) \cap (\overline{M_0} \setminus M_0))$, i.e. $\overline{M_0}$ and that $\overline{M_0}$ intersect only at common regular (quasi-embedded) points or common singular points; moreover, $M_0$ and $M_0$ can only have tangential intersection, with $M_0 \cap M_0$ locally contained in the union of two embedded $(n - 1)$-dimensional $C^1$ submanifolds.

The crucial advantage Theorem 1.2 affords is the following: it says that in order to prove Theorem 1.1 one only has to address the possibility that the limit varifold arising from $(u_{\varepsilon_j})$ is entirely a minimal hypersurface with a small singular set and multiplicity $\geq 2$, i.e. the possibility that $\mu = q\mathcal{H}^n \mathbf{L} M_0$, where $q \geq 2$ and $M_0$ is a minimal hypersurface embedded away from a closed set of codimension $\geq 7$. For if this is not the case, then we will have produced the desired hypersurface $M$ in Theorem 1.1 by taking $M = M_0$ with $M_0$ as in Theorem 1.2.

In the final step of our proof of Theorem 1.1, we show not that the possibility $\mu = q\mathcal{H}^n \mathbf{L} M_0$ cannot arise, but that if it does then we can produce the desired hypersurface $M$ in Theorem 1.1 by a different method: rather than insisting on
obtaining it from a limit of saddle-type critical points of \( F_{\varepsilon_j, \sigma g} \), we will obtain it from a limit of stable critical points \( v_{\varepsilon_j} \) (as in Proposition 1.1 below). The construction of these stable critical points uses a negative gradient flow of \( F_{\varepsilon_j, \sigma g} \) with a well-chosen initial condition built from \( M_0 \). For this construction embeddedness of \( M_0 \) away from a small set is important, and it is carried out in such a way that there is a fixed non-empty open set \( \Omega \subset \{ v_{\varepsilon_j} > 3/4 \} \) for each \( j \), ensuring that \( \Omega \subset \{ v_{\infty} = 1 \} \) and hence \( \{ v_{\infty} = 1 \} \neq \emptyset \). We may then apply Theorem 1.2 again, this time to the sequence \( (v_{\varepsilon_j}) \). The fact that the original functions \( u_{\varepsilon_j} \) are min-max critical points guarantees that the limit measure corresponding to \( (v_{\varepsilon_j}) \) is non-zero. Thus \( \partial^* \{ v_{\infty} = 1 \} \neq \emptyset \), so by the regularity and separation property guaranteed by Theorem 1.2 for the limit varifold corresponding to \( (v_{\varepsilon_j}) \), we can take the regular (i.e. quasi-embedded) part of \( |\partial^* \{ v_{\infty} = 1 \}| \) to be the desired hypersurface \( M \) in Theorem 1.1.

In summary, our proof of Theorem 1.1 consists of the following three main steps:

**STEP (i)** a proof of Theorem 1.2 above (Theorem 4.1 below).

**STEP (ii)** a min-max construction of a critical point \( u_\varepsilon \) of \( F_{\varepsilon, \sigma g} \) for each sufficiently small \( \varepsilon \), with the property that \( 0 < L \leq E_\varepsilon(u_\varepsilon) \leq K < \infty \) for constants \( L, K \), independent of \( \varepsilon \) and \( g \). (Proposition 5.1, Lemma 5.2 and Lemma 5.3). This step is based on standard PDE tools: an application of a mountain pass lemma based on the fact that \( F_{\varepsilon, \sigma g} \) satisfies the Palais–Smale condition. It is then automatically true that the Morse index of \( u_\varepsilon \) is at most 1. Taking \( \varepsilon = \varepsilon_j \) in this construction for a sequence \( \varepsilon_j \to 0^+ \), the uniform upper bound on \( E_\varepsilon(u_{\varepsilon_j}) \) implies that along a subsequence \( u_{\varepsilon_j} \to \mu \) (notation as in Theorem 1.2) for some Radon measure \( \mu \) on \( N \), with \( \mu \neq 0 \) in view of the positive lower bound on \( E_\varepsilon(u_{\varepsilon_j}) \). If \( E = \lim \{ u_{\infty} = 1 \} \neq \emptyset \), Theorem 1.1 holds in view of Theorem 1.2 by just setting \( M \) to be the quasi-embedded part of \( \partial E \).

**STEP (iii)** a proof of the following result (Proposition 6.1 below) which leads to the conclusion of Theorem 1.1 if in step (ii) we have that \( \{ u_{\infty} = 1 \} = \emptyset \).

**Proposition 1.1.** Let \( (u_{\varepsilon_j}) \) be the sequence as step (ii), constructed as described in Section 5. Let \( \mu \) be the (non-zero) limit Radon measure corresponding to (a subsequence of) \( (u_{\varepsilon_j}) \), and suppose that \( \mu = g H^n \llcorner M_0 \) where \( M_0 \) and \( g \) are as in Theorem 1.2(b). Then there exist \( v_{\varepsilon_j} : N \to \mathbb{R} \) satisfying \( F_{\varepsilon_j, \sigma g}(v_{\varepsilon_j}) = 0 \) and \( F_{\varepsilon_j, \sigma g}^\prime(v_{\varepsilon_j}) \geq 0 \) (i.e. stable critical points of \( F_{\varepsilon_j, \sigma g} \)) with \( \liminf_{\varepsilon_j \to 0} v_{\varepsilon_j}(v_{\varepsilon_j}) > 0 \) and \( \limsup_{\varepsilon_j \to 0} E_{\varepsilon_j}(v_{\varepsilon_j}) < \infty \); moreover, there exists a (fixed) non-empty open set contained in \( \{ v_{\varepsilon_j} > \frac{1}{2} \} \) for all \( \varepsilon_j \).

In certain cases, for instance if \( g \) is constant and \( N \) has positive Ricci curvature, it follows from step (iii) of our proof that the possibility \( V_0 = \emptyset \) cannot happen for the sequence of min-max critical points \( (u_{\varepsilon_j}) \) constructed in step (ii) (see Remark 6.7 below). It is also conceivable that this possibility does not occur for more general metrics, for instance for \( n \leq 6 \) and for metrics on \( N \) for which all minimal hypersurfaces are non-generate (a dense subset of the set of smooth metrics by a theorem of White (White91)). We emphasize that here we bypass this question altogether by proceeding as in step (iii).

We end this introduction with remarks on some recent and old work related to the present work.

**Remark 1.5.** The existence of hypersurfaces with mean curvature prescribed by an ambient function \( g (\neq 0) \) on a compact manifold \( N^{n+1} \) has recently been addressed by techniques very different from those used here—specifically, by an Almgren–Pitts minmax construction—in the following cases: for \( 2 \leq n \leq 6 \) when \( g \) is any non-zero constant by Zhou–Zhu (ZhouZhu19); for \( 2 \leq n \leq 6 \) when \( g : N \to \mathbb{R} \).
is smooth and satisfies certain conditions on its nodal set \( \{ g = 0 \} \) (the resulting collection of functions being generic in a Baire sense) by Zhou–Zhu \([ZhoZhu20]\); for \( n \geq 7 \) and \( g \) any non-zero constant by Dey \([Dey]\).

For \( n \leq 6 \), the compactness theorem needed in these works to prove regularity of the minmax varifolds is a straightforward consequence of the a priori curvature estimates due to Schoen–Simon–Yau \([SSY75]\) (for \( n \leq 5 \)) and Schoen–Simon \([SchSim81]\) (for \( n = 6 \)). For general dimensions such estimates do not hold and the corresponding compactness theorems have to be obtained by different methods. The results of \([BelWic-1]\), \([BelWic-2]\) and \([SchSim81]\) provide a complete (regularity and) compactness theory in general dimensions needed for both the Almgren–Pitts approach as in \([Dey]\) as well as our approach in the present work.

We point out some methodological differences between the Almgren–Pitts min-max construction and the Allen–Cahn method pursued here. Just like in the original Almgren–Pitts method used to construct minimal hypersurfaces, the arguments in \([ZhoZhu19]\), \([ZhoZhu20]\), \([Dey]\) seem to require a “pull-tight” procedure (to make up for the lack of a Palais–Smale condition), the fulfilment of an “almost-minimizing” condition (as a substitute for uniform Morse index control), and the use of “stable replacements” (to obtain the regularity conclusions). None of these steps are required in the Allen–Cahn approach, which instead capitalises on elementary, general PDE principles and a sharp varifold regularity theory of independent interest.

Remark 1.6. Historically the use of the Allen–Cahn energy to approximate the area functional goes back to an idea of De Giorgi and to the work of Modica–Mortola \([ModMor77]\), in the case of minimizers. In particular minimizers of \( \mathcal{F}_{\varepsilon, \sigma g} \) converge in \( L^1 \) to a \( BV \) limit taking values \( \pm 1 \), and whose +1 phase \( E \) is a minimizer of the functional \( \text{Per}(E) - \int_E g \). Higher multiplicity issues (including touching singularities) addressed here do not arise in the case of minimizers, and regularity (i.e. embeddedness away from a codimension 7 singular set) of the limit follows from the work of Gonzalez–Massari–Tamanini \([GMT80], [GMT83]\).

2 The case \( g \equiv 0 \), new difficulties when \( g > 0 \) and an outline of the proof

In this section we provide a brief description of the Allen–Cahn approach to the case \( g \equiv 0 \) of Theorem 1.1 (i.e. the existence of minimal hypersurfaces), and a more detailed overview of step (i), step (ii) and step (iii) above in the case \( g > 0 \), highlighting the differences between that case and the case \( g \equiv 0 \).

In recent years, the Allen–Cahn approach has been shown to provide a simple PDE theoretic alternative to the classical Almgren–Pitts geometric measure theory machinery originally used to produce min-max minimal hypersurfaces. The analogue of step (ii) in this case is carried out in the work of Guaraco \([Gua18]\) which shows, via a classical mountain pass lemma, the existence of a critical point \( u_\varepsilon \) of \( \mathcal{E}_\varepsilon \) for small \( \varepsilon \in (0, 1) \), with Morse index at most 1 and with uniform upper and lower energy bounds. Choosing next \( \varepsilon_j \to 0^+ \) and applying a theorem of Hutchinson–Tonegawa \([HutTon00]\), one obtains a sequence of \( n \)-varifolds \( \{ V_j \} \) associated with the measures \( \mu_j = (2\sigma)^{-1} (\varepsilon_j^{-1} |\nabla u_{\varepsilon_j}|^2 + 2^{-1} W(u_{\varepsilon_j})) \), such that along a subsequence, \( V_j \) converges in the first instance to a non-trivial area-stationary integral \( n \)-varifold \( V \) on \( N \), with its weight measure \( \|V\| = \mu = \lim_{j \to \infty} \mu_j \). The proof is completed by deducing regularity of \( \mu \) as follows: the uniform Morse index bound on \( u_{\varepsilon_j} \) implies that in a neighborhood of each point in \( N \) a subsequence of \( (u_{\varepsilon_j}) \) is stable (i.e. has index zero). Hence by a theorem of Tonegawa \([Ton05-1]\), locally near every point of \( \text{spt} \mu \) the embedded part of \( \text{spt} \mu \) is stable with respect to the area functional for compactly supported normal variations (of the embedded
Regularity of $\mu$ (i.e. that $\mu = g\mathcal{H}^n \bigcap M_0$ where $M_0$ is a minimal hypersurface smoothly embedded away from a closed singular set (equal $M_0 \setminus M_0$) of Hausdorff dimension $\leq n - 7$, and $g$, the multiplicity function, is a positive integer valued and locally constant on $M_0$) follows by using local stability of the sequence $(u_{\varepsilon_j})$ to rule out a certain very specific type of singularities of $\mu$ (classical singularities, where at least one tangent cone is made up of three or more half-hyperplanes meeting along a common axis) ([Ton05-2]), and applying the regularity theory for stable codimension 1 stationary varifolds ([Wic13]). Note that whether the set $E = \{u_{\infty} = 1\}$ is empty or not is irrelevant in the case of minimal hypersurfaces, and hence an additional step like step (iii) above is not necessary in this case. (We remark that in dimension $n = 2$, more recent work of Chodosh–Mantoulidis ([ChoMan20]) provides an alternative proof of the regularity of the limit surface by establishing strong convergence of the level sets of the Allen–Cahn solutions.)

For the case $g > 0$, step (ii) is again, in essence, an application of a standard PDE mountain pass lemma. For each small $\varepsilon \in (0, 1)$, this step produces two functions $a_{\varepsilon}, b_{\varepsilon} \in W^{1,2}(N)$ close to the constant functions $-1$ and $1$ respectively, and a critical point $u_{\varepsilon}$ of $F_{\varepsilon, \sigma g}$ in $W^{1,2}(N)$ such that $F_{\varepsilon, \sigma g}(u_{\varepsilon})$ is the min-max width of $F_{\varepsilon, \sigma g}$ over all continuous paths in $W^{1,2}(N)$ joining $a_{\varepsilon}$ to $b_{\varepsilon}$, i.e.

$$F_{\varepsilon, \sigma g}(u_{\varepsilon}) = \min_{\gamma \in \Gamma_{\varepsilon}} \max_{t \in [-1, 1]} F_{\varepsilon, \sigma g}(\gamma(t))$$

where $\Gamma_{\varepsilon} = \{\text{continuous } \gamma : [-1, 1] \to W^{1,2}(N) \text{ with } \gamma(-1) = a_{\varepsilon}, \gamma(1) = b_{\varepsilon}\}$.

We shall now provide a discussion of the key new aspects of step (i) in the case $g > 0$, followed by an overview of step (iii).

The starting point of step (i) is the work of Röger–Tonegawa ([RogTon08]) and Tonegawa ([Ton05-2]) (Theorem 5.1 below). These works, which extend [HutTon00], imply that the limit varifold $V$ is an integral $n$-varifold having locally bounded generalized mean curvature $H_V$ in $N$, that the (possibly empty) set $E$ (equal $\{u_{\infty} = 1\}$) is a Caccioppoli set in $N$ with reduced boundary $\partial^* E \subset \mathrm{supp} \, \mu$, that $H_V(x) = 0$ for $\mu$-a.e. $x \in N \setminus E$ and $H_V(x) = g(x)\nu(x)$ for $\mu$ a.e. $x \in \partial^* E$ where $\nu$ is the unit normal to $\partial^* E$ pointing into $E$, and that the density $\Theta(\mu, x) = \alpha$ for $\mu$ a.e. $x \in \partial^* E$; moreover, in the case of positive $g$ as considered here, we have that $\mathrm{supp} \, \mu \subset N \setminus E$ and $\Theta(\mu, x)$ is an even integer for $\mu$ a.e. $x \in \mathrm{supp} \, \mu \setminus E$.

As regards regularity of $\mu$, the first important difference between the case $g \equiv 0$ and the case $g > 0$ is that as mentioned above, a limit varifold $V$ in the latter case may not in its entirety be a solution to a variational problem (whereas in the former case it is, namely a critical point of $n$-dimensional area). As the next best option we would like to say that the part of $V$ corresponding to the reduced boundary (i.e. the multiplicity 1 varifold $|\partial^* E|$, called the phase boundary) and the complementary part (i.e. $V \setminus N \setminus E$, called the hidden boundary) are separately critical points, in $N$, of the functionals $A - \text{Vol}_g$ and $A$ respectively, where $A$ is the $n$-dimensional area functional and $\text{Vol}_g$ is the enclosed $g$-volume. However even under a uniform Morse index bound on $u_{\varepsilon_j}$ this does not follow a priori from the results of [RogTon08], [Ton05-2]. In addition to the lack of regularity at this stage, a serious difficulty impeding such a decomposition is the topologically complicated ways in which the two parts $|\partial^* E|$ and $V \setminus (N \setminus E)$ may merge together on a set of points $y$ of positive $(n - 1)$-dimensional Hausdorff measure where $V$ has planar tangent cones of multiplicity $> 1$, e.g. as depicted schematically in Figure 1.

In an inductive argument to resolve this regularity question, we first apply a key varifold regularity result (Theorem 4.4 below) from our recent work ([BelWic-1], [BelWic-2]) to show (in Theorem 4.2 that subject to a uniform bound on the Morse index of $u_{\varepsilon_j}$, such topologically complicated behavior indeed does not occur; in fact an ordered $C^{1,\alpha}$ graph structure for $V$ must hold near such a point $y$. Once this is achieved, we combine (in Theorem 4.5) local stability with an adaptation of a PDE argument from [BelWic-1] to prove $C^2$ (and hence $C^{1,\alpha}$) regularity of
these graphs near \( y \) (Theorem 4.2 part (ii)). The difficulty in this higher regularity question is that the union of the graphs need not be embedded; indeed, two graphs can touch a priori on a large (but \( H^n \)-null) set, and the size of this coincidence set is only shown to be lower dimensional (in fact to be locally contained in an \((n-1)\)-dimensional \( C^1 \) submanifold) after \( C^2 \) regularity of the graphs is established.

Once \( C^2 \) regularity is established in a neighbourhood of a point \( y \) where \( V \) has a planar tangent cone, it follows that \( \text{spt} |\partial^* E| \) and \( \text{spt} \mu_{N \setminus E} \) must merge at \( y \), if at all, smoothly (i.e. as \( C^{3, \alpha} \) graphs each of which is entirely minimal or entirely PMC with mean curvature \( g \)) and tangentially; moreover, this can happen only in one of the two possible ways described in Definition 2 parts (iv), (v) and Figure 2. This result, together with the absence of classical singularities in \( V \) (Theorem 4.2 part (i)), leads to the global decomposition and regularity conclusion for \( \mu \) as in Theorem 1.2.

There is also a key difference between the case \( g \equiv 0 \) and the case \( g > 0 \) in the way in which local stability of \( u_{\varepsilon_j} \) (implied by the uniform Morse index bound on \( u_{\varepsilon_j} \)) is used in the analysis of the limit varifolds \( V \) described above. In either case, by [Ton05-1], stability of (a subsequence of) \( u_{\varepsilon_j} \) in a ball implies that the limit varifold \( V \) in that ball admits a generalized second fundamental form satisfying an “ambient” stability inequality, namely, inequality (7); this inequality has the form of the usual “intrinsic” stability inequality for embedded stable hypersurfaces, but with one important difference: it is valid only for ambient \( C^1 \) compactly supported test functions with the ambient gradient (of the test function) appearing where the intrinsic (hypersurface) gradient appears in the intrinsic stability inequality. The ambient and intrinsic stability inequalities are equivalent on the embedded part of a hypersurface (any \( C^1 \) test function \( \varphi \) supported on the embedded part has a \( C^1 \) compactly supported extension \( \tilde{\varphi} \) to the ambient space such that on the hypersurface the ambient gradient of \( \tilde{\varphi} \) agrees with the intrinsic gradient of \( \varphi \)). For this reason, and also since the regularity theory of [Wic14] only requires stability of the embedded part of the varifold, ambient stability inequality for limit varifolds is sufficient in the case \( g \equiv 0 \).

This is not so in the case \( g > 0 \). In this case, unlike in the case \( g \equiv 0 \), two \( C^2 \) pieces of the limit \( n \)-varifold may intersect tangentially along a set (the coincidence set) of finite, positive \((n-1)\)-dimensional measure (consider for instance two
touching unit cylinders in Euclidean space). Because of this possibility stability of the embedded part alone cannot be expected to yield any useful regularity estimates for the $C^2$ quasi-embedded part. (A coincidence set of positive $(n-1)$-dimensional measure, or even infinite $(n-2)$-dimensional measure for that matter, is too large to be removable for the (intrinsic) stability inequality on the embedded part.) Stability of the quasi-embedded part as an immersion, which would imply the intrinsic stability inequality on the quasi-embedded part, would suffice in the present context, but this does not follow from the ambient stability inequality (i.e. inequality (7) below, which is implied by stability of $(u_{\varepsilon_j})$) even though the latter holds on the quasi-embedded part. (A $C^1$ function on a quasi-embedded hypersurface is not necessarily the restriction to the hypersurface of an ambient $C^1$ function.) The question then is which alternative stability condition, valid across the coincidence set, would suffice.

As we identify in the present work, of crucial significance is the fact that it suffices to have an “ambient Schoen inequality” on the quasi-embedded part. This is of course fully consistent with the case $g \equiv 0$. In that case the quasi-embedded part is the embedded part, and the (stronger) intrinsic Schoen inequality ([SchSim81, Lemma 1]) on the embedded part is valid and is implied by the intrinsic stability inequality. This implication is the extent to which the intrinsic stability plays a role in the proof of the main estimate in [SchSim81, Theorem 1], and the application of [SchSim81, Theorem 1] is the way the stability condition enters the proof in [Wic13]. In the present case of $g > 0$ we do not have an intrinsic Schoen inequality on the quasi-embedded part, but instead we obtain its ambient analogue where any ambient test function and its ambient gradient take the place of the intrinsic test function and its intrinsic gradient. This ambient Schoen inequality is obtained directly from the stability of the Allen–Cahn critical points $u_{\varepsilon}$. For the Euclidean ambient space (and for $g \equiv 0$), the derivation of this inequality was carried out by Tonegawa in [Ton05-1]. We here (in Lemma 4.3 below) generalise this to Riemannian manifolds (and to $g \not\equiv 0$). (This generalisation turns out to be somewhat subtle and require a careful choice of coordinates; we provide a full account of it in Lemma 4.3.)

In [BelWic2] a “non-variational” version of our regularity theory (namely, [BelWic2, Theorem 9.1 and Theorem 5.1]) is established taking this Schoen–Tonegawa inequality as the stability assumption on the quasi-embedded region. With these results from [BelWic2] in hand, as well as extensions of some of the arguments in [Bel-Wic-1, Section 7] to account for the presence of both minimal and PMC parts, in Sections 4.4 and 4.5 below we complete the proof of Theorem 1.2.

Finally we briefly describe step (iii). This is the way our argument produces a hypersurface with mean curvature $g$ even in the event that the limit measure $\mu$ arising from the min-max critical points $u_{\varepsilon_j}$ constructed in step (ii) ends up being supported entirely on a minimal hypersurface $M_0$. This step is in part inspired by the ideas in the recent work of the first author ([Bel]), where it is shown that if the ambient manifold has positive Ricci curvature, then any minimal hypersurface arising from a sequence of 1-paramater min-max critical points of the homogeneous Allen–Cahn equation must have multiplicity 1. Our argument here requires the construction, for each sufficiently small $\varepsilon$, of a function $h = h^\varepsilon$. The construction takes $M_0$ (and its regularity) as starting point; $h$ is smooth for $n \leq 6$, while for $n \geq 7$ it is Lipschitz. The function $h$ (after an appropriate smoothing in case $n \geq 7$) is then used as initial data for a negative $F_{\varepsilon, \sigma g}$-gradient flow. This flow is then proved to converge to a stable solution $v_{\varepsilon}$ satisfying the properties given in Proposition 1.1. Special care has to be taken to ensure certain conditions on $h$, namely:

(I) a mean convexity condition, namely, $-F'_{\varepsilon, \sigma g}(h) > 0$ (where the first variation $-F'_{\varepsilon, \sigma g}(h) = \varepsilon \Delta h - W'(h) + g$ should be understood distributionally:

10
when \( n \geq 7, -F'_{\varepsilon, \sigma g}(h) \) is in fact a Radon measure, while for \( n \leq 6 \) it is a smooth function);

(II) the existence (for all sufficiently small \( \varepsilon \)) of a continuous path in \( W^{1,2}(N) \) that joins \( a_\varepsilon \) (the valley point close to \(-1\) identified in step (ii)) to \( h \), such that all along the path we have \( (2\sigma)^{-1}F_{\varepsilon, \sigma g} \leq 2\mathcal{H}^n(M_0) - \varsigma/2 \), where \( \varsigma > 0 \) depends only on the geometric data \( M_0 \subset N \).

The construction of \( h \) is based on a certain deformation of the minimal immersion \( \varepsilon \) of the oriented double cover of \( M_0 \) into \( N \), that covers \( M_0 \) twice. This deformation is schematically depicted in Figure 3 in (i), two disks \( D_1 \) and \( D_2 \) are removed from \( M_0 \), leaving a double cover of \( M_0 \setminus (D_1 \cup D_2) \) (an immersion with boundary); in (ii), one of the disks, say \( D_1 \), is pasted back in (with multiplicity 2); in (iii) the double disk \( 2|D_1| \) is deformed as an immersion: since \( M_0 \) has vanishing mean curvature, and is therefore not stationary for the functional \( \text{Area} - \text{Vol}_g \), this deformation decreases the value of \( \text{Area} - \text{Vol}_g \); in (iv) and (v) the same type of operations are performed on \( D_2 \). The deformation represented in Figure 3 is constructed in order to ensure that the value of \( \text{Area} - \text{Vol}_g \) as we go from (i) to (v) stays strictly below \( 2\mathcal{H}^n(M_0) \), by an amount \( \varsigma > 0 \) that only depends on the geometric properties of \( M_0 \) in \( N \). Note that, for \( \varepsilon \), \( \text{Vol}_g \) vanishes and therefore \( \text{Area} - \text{Vol}_g \) is just \( 2\mathcal{H}^n(M_0) \).

For any \( \varepsilon \) sufficiently small, we exhibit a continuous path in \( W^{1,2}(N) \) using functions that replicate, in a suitable sense, the geometric behaviour identified by the deformations represented in Figure 3. These functions are identically \(-1\) or identically \(+1\) on certain regions and the transitions between the values \( \pm 1 \) happen in a small neighborhood of the hypersurfaces depicted in Figure 3. Moreover, the energy \( (2\sigma)^{-1}F_{\varepsilon, \sigma g} \) of these functions is a very close approximation of the value attained by evaluating \( \text{Area} - \text{Vol}_g + \frac{1}{2} \int_N g \) on the hypersurfaces depicted in Figure 3 – with an error of size \( O(\varepsilon |\log \varepsilon|) \).

An extremely effective feature of the Allen–Cahn framework is that the transitions from (i) to (ii) and from (iv) to (v) can be replicated continuously, while the corresponding geometric deformations present discontinuities (with the sudden appearance of a portion). In similar spirit, we can find a continuous path that joins \( a_\varepsilon \) to the function corresponding to the immersion with boundary in (i). (The geometric counterpart is the sudden appearance of the hypersurface \( M_0 \setminus (D_1 \cup D_2) \) with multiplicity 2.) The path “from \( a_\varepsilon \) to (i)” is produced so that \( F_{\varepsilon, \sigma g} \) is almost increasing, i.e. it is equal to an increasing function plus a function bounded in modulus by \( O(\varepsilon |\log \varepsilon|) \). The function \( h \) mentioned above corresponds to the immersion represented in (v) and it is close to \(+1\) in the two domains obtained by deforming the double disks \( D_1 \) and \( D_2 \) and close to \(-1\) in their complement. Moreover, the construction ensures the mean convexity condition in (I). The mean convexity implies that the negative \( F_{\varepsilon, \sigma g} \)-gradient flow starting at \( h \) evolves towards a stable solution \( v_\varepsilon \) of \( \mathcal{F}_{\varepsilon, \sigma g} = 0 \), that has to be \( \geq h \). The assumption that the minmax solutions give rise to \( \mu \) supported entirely on \( M_0 \) with multiplicity \( \geq 2 \) implies that \( v_\varepsilon \) cannot be the second valley point \( b_\varepsilon \). Indeed, if \( v_\varepsilon = b_\varepsilon \), we would have found a continuous path in \( W^{1,2}(N) \) that connects \( a_\varepsilon \) to \( b_\varepsilon \) such that along this path, \( (2\sigma)^{-1}F_{\varepsilon, \sigma g} \) stays below the minmax value \( (2\sigma)^{-1}F_{\varepsilon, \sigma g}(u_\varepsilon) \) by a fixed amount \( \varsigma/2 \) (recall that we have assumed \( (2\sigma)^{-1}F_{\varepsilon, \sigma g}(u_\varepsilon) \to |\mu|(N) + \frac{1}{2} \int_N g \)); this contradicts the minmax characterisation. Very crucially, the fact that \( v_\varepsilon \neq b_\varepsilon \) leads to a uniform positive lower bound on \( \mathcal{E}_\varepsilon(v_\varepsilon) \) independent of \( \varepsilon \), while the energy decreasing property of the flow implies a uniform upper bound on \( \mathcal{E}_\varepsilon(v_\varepsilon) \). Moreover, the condition \( v_\varepsilon \geq h \) guarantees the presence of a fixed open set (corresponding to the two domains obtained by deforming the double disks \( D_1 \) and \( D_2 \)) on which \( v_\varepsilon \geq 3/4 \) for all \( \varepsilon \), as claimed in Proposition 1.1. Then we can consider the varifold associated to \( v_\varepsilon \) and apply Theorem 1.3 (with \( \varepsilon j \) in place of \( \epsilon_j \) for a sequence \( \varepsilon_j \to 0^+ \)) to obtain the desired hypersurface with mean curvature \( g \).
We remark that the arguments given in Section 6 for step (iii) become considerably shorter if \( n \leq 6 \), since in that case \( M_0 \) has no singularities. In general dimensions, while the basic geometric idea remains the same, a finer analysis of the distance function to \( M_0 \) is necessary to handle the presence of a small singular set in \( M_0 \) (as well as some extra technical asides).

3 Preliminaries: Allen–Cahn solutions and limit \((g,0)\)-varifolds

In this section and in Section 4 the ambient manifold \( N \) need not be complete. For \( \varepsilon \in (0,1) \) let \( u_\varepsilon : \mathbb{R}^n \to \mathbb{R} \) satisfy \( \mathcal{E}_\varepsilon(u_\varepsilon) \leq K \), for \( K > 0 \) independent of \( \varepsilon \). Let \( w_\varepsilon = \Phi(u_\varepsilon) \), for \( \Phi(s) = \int_0^s \sqrt{\frac{W(s)}{2}} ds \) and \( \sigma = \int_{-1}^1 \sqrt{\frac{W(s)}{2}} ds \). The BV-norm of \( w_\varepsilon \) is controlled by \( \mathcal{E}_\varepsilon(\sigma) \) by the computation in [HutTon00] (and [ModMor87]), hence uniformly bounded by our assumption. This allows to obtain a subsequential BV-limit \( w_\infty \) of \( w_\varepsilon \), as \( \varepsilon \to 0 \). We also denote \( u_\infty = \Phi^{-1}(w_\infty) \), and we have \( u_\varepsilon \to u_\infty \) in \( L^1 \) (see section 5). We define, as in [HutTon00], the \( n \)-varifold \( V^{u_\varepsilon} \) associated to \( u_\varepsilon \)

\[
V^{u_\varepsilon}(A) = \int_{-\infty}^{\infty} V\{w_\varepsilon = t\}(A)dt,
\]

where \( V\{w_\varepsilon = t\} \) is the multiplicity-1 varifold associated to the level set \( \{w_\varepsilon = t\} \). Since \( u_\varepsilon \) is smooth by elliptic theory, for a.e. \( t \) this level set is smooth.

We recall the main results of [RogTon08], and of [Ton05-2] (which rely on [HutTon00]), concerning \( V \) in the case \( g > 0 \).

**Theorem 3.1.** ([RogTon08, Theorem 3.2] and [Ton05-2, Theorem 2.2]) Let \( \alpha \in (0,1) \), \( g \in C^{1,1}(\mathbb{R} \setminus \{1\}) \), \( g_j \in C^{1,1}(\mathbb{R}) \) with \( g_j \to g \) locally in \( C^{1,1} \) and \( \varepsilon_j \to 0^+ \). Let \( u_{\varepsilon_j} \in W^{1,2}_{\text{loc}}(\mathbb{R} \setminus \{1\}) \) be a critical point of \( F_{\varepsilon_j,\sigma_j} \) for each \( j \), and suppose that \( \lim_{j \to \infty} V^{u_{\varepsilon_j}} = V \) for some \( n \)-varifold \( V \) on \( \mathbb{R} \setminus \{1\} \), where \( V^{u_{\varepsilon_j}} \) is the \( n \)-varifold on \( \mathbb{R} \setminus \{1\} \) associated with \( u_{\varepsilon_j} \) as described above. Let \( u_\infty \in BV_{\text{loc}}(\mathbb{R} \setminus \{1\}) \) be such that \( u_{\varepsilon_j} \to u_\infty \) locally in \( L^1 \), and note that such \( u_\infty \) exists possibly after passing to a subsequence of \( (\varepsilon_j) \). Let \( E = \{ x \in \mathbb{R} : u_\infty(x) = 1 \} \). We have the following:

(i) \( \sigma^{-1} V \) is an integral \( n \)-varifold, with \( V \) having locally bounded generalised mean curvature \( \mathcal{H}_V \) and first variation \( \delta V = -\mathcal{H}_V \lVert V \rVert \) in \( E \), and \( E \) is a Caccioppoli set with \( \mathcal{H}^* E \subset \text{spt} \lVert V \rVert \subset \mathbb{R} \setminus \text{int} E \).

If additionally \( g > 0 \), and if we denote the inward pointing unit normal to the reduced boundary \( \mathcal{H}^* E \) by \( \nu \) (i.e. if \( \nu = \sum_{x \in E^*} x \)), then:

(ii) \( E \neq N \).

(iii) \( \sigma^{-1}(\mathcal{H}^* \lVert V \rVert, x) = 1 \) and \( \mathcal{H}_V(x) \cdot \nu(x) = g(x) \) for \( \mathcal{H}^n \)-a.e. \( x \in \mathcal{H}^* E \).

(iv) \( \mathcal{H}_V(x) = 0 \) for \( \mathcal{H}^n \)-a.e. \( x \in \text{spt} \lVert V \rVert \setminus \mathcal{H}^* E \).

(v) \( \sigma^{-1}(\mathcal{H}^* \lVert V \rVert, x) \) is an even integer \( \geq 2 \) for \( \mathcal{H}^n \)-a.e. \( x \in \text{spt} \lVert V \rVert \setminus \mathcal{H}^* E \).

**Remark 3.1.** In [RogTon08] and [Ton05-2], these conclusions are established in the case of Euclidean ambient space. Adaptation of the arguments to the case of Riemannian ambient space is routine.

Under the hypotheses of Theorem 3.1, no regularity result for the limit varifold \( V \) is known beyond the fact that the regular set is dense in \( \text{spt} \lVert V \rVert \) (which follows from Allard’s regularity theorem). Heuristically, the theorem says that the minimal portions, if there are any, always appear with even multiplicity and lie in the \( \{ u_\infty = 1 \} \)-phase. In principle, minimal and non-minimal portions may come together in irregular fashion (e.g. as depicted in Figure 1). More threatening for the success of the minmax approach is the possibility that the limit interface
could be completely minimal, i.e. the possibility that \( u_\infty \equiv -1 \) a.e. on \( N \) and \( \text{spt} \| V \| \) all consists of "hidden boundary". This latter possibility in fact arises under the hypotheses of Theorem 3.1 even for \( g \equiv 1 \) and \( N = \mathbb{R}^n \); see [HutTon00] Section 6.3.

Throughout the rest of the article, it will be convenient to use the terminology defined as follows:

**Definition 1.** Let \( g \in C^{1,1}(N) \). We say that an \( n \)-varifold \( V \) on \( N \) is a limit \((g_j, 0)\)-varifold if there is a sequence of numbers \( \varepsilon_j \to 0^+ \) and for each \( j \), a function \( g_j \in C^{1,1}(N) \) and a critical point \( u_{\varepsilon_j} \in W^{1,2}(N) \) of \( F_{\varepsilon_j, \sigma g_j} \) such that \( g_j \to g \) locally in \( C^{1,1} \) and \( V = \lim_{j \to \infty} V_{\varepsilon_j} \), where \( V_{\varepsilon_j} \) is the \( n \)-varifold on \( N \) associated with \( u_{\varepsilon_j} \) as described above. We say that \( V \) is a stable limit \((g, 0)\)-varifold if \( V \) is a limit \((g, 0)\)-varifold and the associated critical points \( u_{\varepsilon_j} \) of \( F_{\varepsilon_j, \sigma g_j} \) are stable, i.e. satisfy \( \frac{\partial^2}{\partial s^2} F_{\varepsilon_j, \sigma g_j}(u_{\varepsilon_j} + s \varphi) \geq 0 \) for each \( \varphi \in C^1(N) \).

This terminology is motivated by Theorem 3.1 according to which a varifold \( V \) which is the limit of a sequence of varifolds \( V_{\varepsilon_j} \) as described above admits generalised mean curvature, and consists of an oriented portion where the generalised scalar mean curvature is equal to \( g \), and a complementary portion where the generalised mean curvature is 0.

### 4 Regularity of limit \((g, 0)\)-varifolds

In this section we use the varifold regularity theory developed in our earlier work ([BelWic1], [BelWic2]) to completely analyse the limit \((g, 0)\)-varifolds corresponding to Morse index bounded Allen–Cahn solutions. Results in this section may be of interest in contexts other than that of Theorem 1.1.

#### 4.1 Main theorems: smooth excision of hidden boundary

Our main regularity theorem (Theorem 4.1 below) says that if \( g \in C^{1,1}(N) \) is positive, then a limit \((g, 0)\)-varifold \( V \) for which the associated sequence \( (u_{\varepsilon_j}) \) has uniformly bounded Morse index with respect to \( F_{\varepsilon_j, \sigma g_j} \) is regular (i.e. has quasi-embedded \( \text{PMC}(g, 0) \) structure as in Definition 2 below) away from a closed set of codimension \( \geq 7 \), and moreover, that the "hidden boundary" \( V \cap (N \setminus \text{E}) \) (on which \( H_V = 0 \)) and the "phase boundary" \( \partial \text{E} \) (where \( H_V = g \nu \)) can be separated smoothly globally, i.e. neither the hidden boundary nor the phase boundary has singular first variation in \( N \) (and each is separately regular away from a closed set of codimension \( \geq 7 \)).

**Definition 2.** Let \( g : N \to \mathbb{R} \) be a positive continuous function and let \( V \) be an integral \( n \)-varifold on \( N \). We say that \( V \) has quasi-embedded \( \text{PMC}(g, 0) \) structure near a point \( Y \in \text{spt} \| V \| \) if one of the following (depicted in figure 2) holds:

(i) \( V \) near \( Y \) is equal to the multiplicity 1 varifold \( |D| \) for some \( C^2 \) embedded disk \( D \) with a choice of continuous unit normal with respect to which the scalar mean curvature of \( D \) is equal to \( g \) everywhere;

(ii) \( V \) near \( Y \) is equal to \( q|D'| \) for some even integer \( q \) and a \( C^2 \) embedded minimal disk \( D' \);

(iii) \( \Theta(|V|, Y) = 2 \) and \( V \) near \( Y \) is equal to \( |D_1| + |D_2| \) for two distinct \( C^2 \) embedded disks \( D_1, D_2 \) having only tangential intersection, with each disk lying on one side of the other, having mean curvature vector pointing away from the other and having scalar mean curvature (with respect to the unit normal in the direction of the mean curvature vector) equal to \( g \) everywhere;
(iv) $\Theta([V], Y) = q$ and $V$ near $Y$ is equal to $|D| + (q-1)|D'|$ for some odd integer $q \geq 3$ and two distinct $C^2$ embedded disks $D$, $D'$ having only tangential intersection, with each disk lying on one side of the other; $D$ having mean curvature vector pointing away from $D'$ and scalar mean curvature (with respect to the unit normal in the direction of the mean curvature vector) equal to $g$ everywhere, and with $D'$ being minimal;

(v) $\Theta([V], Y) = q$ and $V$ near $Y$ is equal to $|D_1| + |D_2| + (q-2)|D'|$ for some even integer $q \geq 4$ and three distinct $C^2$ embedded disks $D_1$, $D_2$, $D'$ with each pair of disks having only tangential intersection, and where $D_1$, $D_2$ are precisely as in (iii) and $D'$ is minimal and lies between $D_1$ and $D_2$.

Figure 2: Quasi-embedded PMC$(g, 0)$ structure of a varifold $V$. In each case, $V$ corresponds to a certain number of disks, depicted by curves. The disks depicted by thin curves have multiplicity 1 and mean curvature $g\nu$ for a choice of unit normal $\nu$, and those depicted by thick curves are minimal and have even multiplicity. The numbers $+1$ and $-1$ indicate the phase values, i.e. the values of $u_\infty$, when $V$ is the limit varifold arising from a sequence of Allen–Cahn critical points $u_{\varepsilon_j}$ with $g > 0$. In this case $\nu$ always points into the $+1$ phase.

**Definition 3.** For a limit $(g, 0)$-varifold $V$ on $N$ (Definition 1), we define the singular set $\text{sing} V$ to be the set of points $Y \in \text{spt} [V]$ such that there is no neighborhood of $Y$ in which $V$ has quasi-embedded PMC$(g, 0)$ structure.

**Remark 4.1.** Let $W$ be an integral $n$-varifold on $N$ and let $g$ be a positive $C^{1,1}$ function on $N$. In Theorem 4.1 and subsequently, we denote by gen-reg $W$ the set of points $Y \in \text{spt} [W]$ near which $W$ has the structure as in (i) or (iii) of Definition 2. Thus near a point in gen-reg $W$, either $\text{spt} [W]$ is a $C^2$ embedded disk, or is equal to precisely two $C^2$ embedded disks intersecting tangentially with each disk on one side of the other, and in either case with the scalar mean curvature given by $g$. Note that this terminology is consistent with that of [BelWic-2, Definition 1.7].
Theorem 4.1. Let \( g \in C^{1,1}(N) \) be positive, and let \( V \) be a limit \( g \)-varifold on \( N \) with associated sequences \( \varepsilon_j \to 0^+ \), \( g_j \in C^{1,1}(N) \) and \( u_{\varepsilon_j} \in W^{1,2}_{\text{loc}}(N) \) where \( g_j \to g \) locally in \( C^{1,1} \) and \( u_{\varepsilon_j} \) is a critical point of \( F_{\varepsilon_j,g_j} \) for each \( j \). Suppose further that the sequence \( (u_{\varepsilon_j}) \) converges to \( u_\infty \in BV_{\text{loc}}(N;\{-1,1\}) \) locally in \( L^1 \), noting that such \( u_\infty \) exists possibly after passing to a subsequence of \( (u_{\varepsilon_j}) \). Let \( E = \{ x \in N : u_\infty(x) = 1 \} \) and note, by Theorem 3.7, that \( E \) is a Caccioppoli set with \( E \neq N \). Finally suppose that the Morse index of \( u_{\varepsilon_j} \) with respect to \( F_{\varepsilon_j,g_j} \) is bounded independently of \( j \). Then we have \( \sigma^{-1}V = V_0 + V_g \), where:

(i) \( V_0 \) is a (possibly zero) stationary integral \( n \)-varifold on \( N \) with \( \text{spt } V_0 = \emptyset \)

if \( 2 \leq n \leq 6 \), \( \cup \text{sing } V_0 \) discrete if \( n = 7 \) and \( \dim_H (\text{sing } V_0) \leq n - 7 \) if \( n \geq 8 \); moreover, \( V_0 \) has locally constant even multiplicity on \( \text{reg } V_0 \) and \( \sup \| V_0 \| \subset N \setminus \text{int } (E) \).

(ii) if \( E \neq \emptyset \) then \( V_g = |\partial^* E| \), and we have that \( \text{spt } V_g \setminus \text{gen-reg } V_g = \emptyset \) if

\( 2 \leq n \leq 6 \), \( \cup \text{spt } V_g \setminus \text{gen-reg } V_g \) is discrete if \( n = 7 \) and \( \dim_H (\text{spt } V_g \setminus \text{gen-reg } V_g) \leq n - 7 \) if \( n \geq 8 \); moreover, the (classical) mean-curvature of the immersed hypersurface \( \text{gen-reg } V_g \) is given by \( H_{V_g} = gu \) where \( u \) is the unit normal vector to \( \text{gen-reg } V_g \) pointing inward (i.e. towards \( E \)), and \( H_{V_g} \) is the generalized mean curvature of \( V_g \) in \( N \).

Furthermore, \( \sigma^{-1}V \) has quasi-embedded \( \text{PMC}(g,0) \) structure near each point \( y \in \cup \text{spt } V_g \cup (\text{spt } V_g \setminus \text{gen-reg } V_g) \). (See Figure 3 below).

Theorem 4.1 is a direct consequence of Theorem 3.1 above and the following:

Theorem 4.2. Let \( g \in C^{1,1}(N) \) be a positive function, and let \( V \) be a stable limit \( (g,0) \)-varifold on \( N \). Noting by Theorem 3.2 that \( \sigma^{-1}V \) is integral and that \( V \) admits locally bounded generalized mean curvature, we have that:

(i) no tangent cone to \( V \) is supported on three or more half-hyperplanes meeting along an \((n-1)\)-dimensional subspace.

(ii) if \( Y \in \text{spt } V \) and if there is an \( n \)-dimensional plane \( P \) and a positive integer \( q \) such that \( qP \in \text{VarTan}_V (\sigma^{-1}V) \) then \( \sigma^{-1}V \) has quasi-embedded \( \text{PMC}(g,0) \) structure near \( Y \).

(iii) \( \dim_H (\text{sing } V) \leq n - 7 \).

In fact we have quantitative versions of parts (i) and (ii) of the preceding theorem, given by the estimates in Theorem 4.3 below. These estimates will only be needed in Section 7 where we use an approximation argument to remove the current assumption that \( g \) is a positive, \( C^{1,1} \) function, thereby establishing Theorem 4.1 for non-negative \( g \in C^{0,(\alpha)}(N) \).

Theorem 4.3. Let \( q \) be a positive integer and let \( \Gamma > 0 \). Let \( C \) be a stationary integral cone in \( \mathbb{R}^{n+1} \) supported on a union of three or more \( n \)-dimensional half-hyperplanes meeting along a common \((n-1)\)-dimensional subspace. There are constants \( \eta_0 = \eta_0(n,C,\Gamma) \in (0,1) \), \( \epsilon_0 = \epsilon_0(n,q,\Gamma) \in (0,1) \) and \( \mu = \mu(N,\Gamma) \) such that if \( g \in C^{1,1}(N) \) is a positive function with \( \sup_N |g| \leq \Gamma \), \( V \) is a stable limit \((g,0)\)-varifold on \( N \), \( \rho \in (0,\epsilon_0^4 N) \) and \( \tilde{V} \) is the varifold on \( B^+_{\rho+1}(0) \subset T_y N \approx \mathbb{R}^{n+1} \) obtained by pulling back \( V \cup N_{\rho}(y) \) by the exponential map \( \exp_y \), then the following hold:

(i) if \( \frac{\|\tilde{V}\|_{B^+_{\rho+1}(0)}^{(n+1)}}{\mu \rho \rho^+} \leq \Theta (\|C\|_q, 0) + 1/4 \) then

\[ \mu \rho + \rho^+ \text{dist}_H (\text{spt } C \cap B^+_{\rho+1}(0), \text{spt } \tilde{V} \cap B^+_{\rho+1}(0)) \geq \eta_0; \]

(ii) if \( \frac{\|\tilde{V}\|_{B^+_{\rho+1}(0)}^{(n+1)}}{\mu \rho \rho^+} \leq q + 1/2 \), \( q - 1/2 \leq \frac{\|\tilde{V}\|_{B^+_{\rho+1}(0)}^{(n+1)}}{\mu \rho \rho^+} \leq q + 1/2 \) and

\[ E_{\rho} \equiv \mu \rho + \rho^{-n-2} \int_{B^+_{\rho+1}(0) \times R} \left| x^{n+1} \right|^2 |\tilde{V}| < \epsilon_0 \]
then
\[ \widetilde{V} \cap (B_{\rho/4}(0) \times \mathbb{R}) = \sum_{j=1}^{q} \left| \text{graph } u_j \right| \]

where \( u_j \in C^{2,\alpha}(B_{\rho/4}^0(0)) \) for any \( \alpha \in (0,1) \), \( u_1 \leq u_2 \leq \ldots \leq u_q \) and
\[
\sup_{B_{\rho/4}^0(0)} \left( \rho^{-1}|u_j| + |Du_j| + \rho|D^2u_j| \right) + \rho^{1+\alpha} \sup_{x,y \in B_{\rho/4}^0(0), x \neq y} \frac{|D^2u_j(x) - D^2u_j(y)|}{|x-y|^\alpha} \leq C \sqrt{\rho}
\]

for some constant \( C = C(n,q,\alpha,\Gamma) \) and each \( j \in \{1,2,\ldots,q\} \).

The proofs of Theorem 4.2 and Theorem 4.3 will rely heavily on the general varifold regularity theory developed in our earlier work [BelWic-1], [BelWic-2]. We shall give their proofs in Section 4.5 below after discussing the necessary preliminary results, but point out next how Theorem 4.1 follows from Theorem 4.2.

**Proof of Theorem 4.1 assuming Theorem 4.2.** Let \( y \in \text{spt } |V| \) be arbitrary. Since Morse index of \( u_j \) is bounded independently of \( j \), it follows from standard arguments that there is \( \rho_j > 0 \) and a subsequence \((\varepsilon_j)\) of \((\varepsilon_j)\) such that \( u_{\varepsilon_j} \) is stable in \( B_{\rho_j}(y) \) for each \( k = 1,2,3,\ldots \), and consequently \( \tilde{V} = V \cap B_{\rho_j}(y) \) is a stable limit \((g,0)\)-varifold on \( B_{\rho_j}(y) \). Hence by Theorem 4.2 we have that
\[
\dim_H \left( \text{spt } |V| \setminus \text{genreg } V \right) \cap B_{\rho_j}(y) \leq n - 7,
\]
and consequently \( \dim_H \left( \text{spt } |V| \setminus \text{genreg } V \right) \leq n - 7 \). By Theorem 4.2, part (ii) and Theorem 3.1 we have that for each \( z \in \text{genreg } V \), there is a ball \( B_{\delta_k}(z) \subset \mathbb{N} \) such that \( \sigma^{-1}V \cap B_{\delta_k}(z) = V_0(z) + V_0^z \) where \( V_0(z) \) is a possibly zero stationary (i.e. zero-mean curvature) \( n \)-varifold on \( B_{\delta_k}(z) \) with no singularities and constant even integer multiplicity and \( V_0^z = |\partial^* (E \cap B_{\delta_k}(z))| \) with \( \text{spt } |V_0^z| \setminus \text{genreg } V_0^z \) satisfies \( H = \nu \) where \( \nu \) is the inward pointing unit normal to \( \partial^* E \).

Now define varifolds \( V_0, V_g \) on  \( N \setminus \text{spt } |V| \setminus \text{genreg } V \) as follows: pick any \( \varphi \in C^2_c((N \setminus \text{spt } |V| \setminus \text{genreg } V) \times G(n)) \). If \( \text{spt } \varphi \subset B_{\delta_k}(z) \times G(n) \) for some \( z \in \text{genreg } V \), set \( V_0(\varphi) = V_0(z)(\varphi) \) and \( V_g(\varphi) = V_g(z)(\varphi) \), noting that these definitions are independent of the choice of \( z \). For an arbitrary function \( \varphi \in C^0_c((N \setminus \text{spt } |V| \setminus \text{genreg } V) \times G(n)) \), set \( V_0(\varphi) = \sum_{\alpha \in I} V_0(\psi_\alpha \varphi) \) and \( V_g(\varphi) = \sum_{\alpha \in I} V_g(\psi_\alpha \varphi) \) where \( \psi_\alpha \) is a smooth, locally finite partition of unity subordinate to the collection of open sets
\[
\left\{ B_{\delta_k}(z) : z \in \text{genreg } V \right\} \cup \{ I \subset \text{genreg } V, B_{\delta_k}(z) \},
\]
where \( I \) is some indexing set. Since \( \text{spt } |V_0| \subset \text{spt } |V| \), \( \text{spt } |V_g| \subset \text{spt } |V| \) and \( V \) has locally bounded generalized mean curvature in \( N \), it follows that \( V_0, V_g \) have Euclidean volume growth everywhere. Since \( \dim_H \left( \text{spt } |V| \setminus \text{genreg } V \right) \leq n - 7 \), then \( V_0 \) is stationary in \( N \), and that \( V_g \) has generalized mean curvature \( H_{V_g} \) in \( N \), given on \( \text{genreg } V \) by \( H_{V_g} = \nu \). We thus have that \( V = V_0 + V_g \) on \( N \) with \( V_0, V_g \) satisfying all of the requirements of the conclusion of the theorem. \( \square \)

**Remark 4.2.** It follows from [BelWic-2, Remark 1.22] that \( \text{gen-reg } V_g \) (where \( V_g \) is as in Theorem 4.1) is the image of a \( C^{3,\alpha} \) immersion.
4.2 A general varifold regularity theorem

The central ingredient in the proofs of Theorem 4.2 and Theorem 4.3 is the following general varifold regularity theorem established in [BelWic-2].

**Theorem 4.4.** ([BelWic-2, Theorem 9.1]) Let $N$ be an $(n + 1)$-dimensional Riemannian manifold (not assumed complete), $q$ be a positive integer, $\beta \in (0, 1)$ and let $p > n$. Let $\mathcal{V}$ be a class of of integral $n$-varifolds $V$ on $N$ satisfying the following properties (a)-(c):

(a) each $V \in \mathcal{V}$ has generalised mean curvature $H_V \in L^p_{loc}(\|V\|)$ and first variation $\delta V = -H_V\|V\|$, i.e. $V$ satisfies

$$\delta V(\psi) = -\int_N <H_V, \psi> d\|V\|$$

for some $H_V \in L^p_{loc}(\|V\|)$ and any compactly supported $C^1$ vector field $\psi$ on $N$;

(b) if $V \in \mathcal{V}$ then no point $Y \in \text{spt}\|V\|$ with $\Theta(\|V\|, Y) = q$ is a classical singularity of $V$;

(c) if $V \in \mathcal{V}$, $X \in N$, $\tilde{V} = (\Gamma \circ \exp_X^{-1})_# V \cap N_{\text{inj}_X}(X)$ for some orthogonal rotation $\Gamma : T_X N \approx \mathbb{R}^{n+1} \to T_X N$, $\sigma \in (0, \min\{1, \text{inj}_X N\}]$, then

$$(\omega_n2^n)^{-1}\|\eta_{0,\sigma} \# \tilde{V}\|(B_{1/2}^{n+1}(0)) \leq q + 1/2,$$

$q - 1/2 \leq \omega_n^{-1}\|\eta_{0,\sigma} \# \tilde{V}\|(B_{1/2}^{n+1}(0) \times \mathbb{R}) \cap B_{1/2}^{n+1}(0)) \leq q + 1/2,$

$\Theta(\|\eta_{0,\sigma} \# \tilde{V}\|, Y) < q$ for each $Y \in B_{1/2}^{n+1}(0)$ and if

$$E \equiv \int_{(B_{1/2}^{n+1}(0) \times \mathbb{R}) \cap B_{1/2}^{n+1}(0)} |x|^{n+1} d\|\eta_{0,\sigma} \# \tilde{V}\|$$

$$+ \sigma \left(\int_{(B_{1/2}^{n+1}(0) \times \mathbb{R}) \cap B_{1/2}^{n+1}(0)} |H_V(\exp_X(\Gamma^{-1}(\sigma Y)))|^{p} \ d\|\eta_{0,\sigma} \# \tilde{V}\|(Y)\right)^{1/p} + \sigma < \beta,$$

then $\eta_{0,\sigma} \# \tilde{V} \cap (B_{1/2}^{n+1}(0) \times \mathbb{R}) \cap B_{1/2}^{n+1}(0(0)) = \sum_{j=1}^{q} |\text{graph}\ u_j|$ for some $u_j \in C^2(B_{1/2}^{n+1}(0))$, $j = 1, 2, \ldots, q$, satisfying $u_1 \leq u_2 \leq \ldots \leq u_q$ and

$$\|u_j\|_{C^{1,\alpha}(B_{1/2}^{n+1}(0))} \leq C\sqrt{E}$$

for any $\alpha \in (0, 1)$ and some constant $C = (N, p, q, \alpha)$.

**Conclusion:** there exists $\varepsilon = \varepsilon(n, p, q, N, \beta, \mathcal{V}) \in (0, 1)$ such that if $V \in \mathcal{V}$, $X \in N$, $\tilde{V} = (\Gamma \circ \exp_X^{-1})_# V \cap N_{\text{inj}_X}(X)$ for some orthogonal rotation $\Gamma : T_X N \approx \mathbb{R}^{n+1} \to T_X N$, $\sigma \in (0, \min\{1, \text{inj}_X N\}]$, then

$$(\omega_n2^n)^{-1}\|\eta_{0,\sigma} \# \tilde{V}\|(B_{1/2}^{n+1}(0)) \leq q + 1/2,$$

$q - 1/2 \leq \omega_n^{-1}\|\eta_{0,\sigma} \# \tilde{V}\|(B_{1/2}^{n+1}(0) \times \mathbb{R}) \cap B_{1/2}^{n+1}(0(0)) \leq q + 1/2,$

and if

$$E \equiv \int_{(B_{1/2}^{n+1}(0) \times \mathbb{R}) \cap B_{1/2}^{n+1}(0)} |x|^{n+1} d\|\eta_{0,\sigma} \# \tilde{V}\|$$

$$+ \sigma \left(\int_{(B_{1/2}^{n+1}(0) \times \mathbb{R}) \cap B_{1/2}^{n+1}(0)} |H_V(\exp_X(\Gamma^{-1}(\sigma Y)))|^{p} \ d\|\eta_{0,\sigma} \# \tilde{V}\|(Y)\right)^{1/p} + \sigma < \varepsilon,$$

17
then
\[ \eta_{0, \sigma} \# \tilde{V}(\{(B_{1/2}^n(0) \times \mathbb{R}) \cap B_{1}^{n+1}(0)\}) = \sum_{j=1}^{q} |\text{graph } u_j| \]
for some \( u_j \in C^{1,\alpha}(B_{1/2}^n(0)) \), \( j = 1, 2, \ldots, q \), with \( u_1 \leq u_2 \leq \ldots \leq u_q \), where \( \alpha = \alpha(n, p) \in (0, 1) \). Furthermore, we have that
\[ \|u_j\|_{C^{1,\alpha}(B_{1/2}^n(0))} \leq C \sqrt{E} \]
for each \( j \in \{1, 2, \ldots, q\} \), where \( C = C(N, p, q) \).

The content of this theorem (for a given varifold) can roughly be described as follows: fix a positive integer \( q \). In the absence of classical singularities (as in hypothesis (b)), if an integral \( n \)-varifold with generalized mean curvature locally in \( L^p \) for some \( p > n \) has the property that its flat regions where density is \( \leq q - 1 \) are “well-behaved” (as in hypothesis (c)), then the varifold is well-behaved also near any point where there is a tangent plane of multiplicity \( q \). Note the important difference in the meaning of “well-behaved” for the hypotheses and for the conclusion: in the hypotheses (i.e. in hypothesis (c)) it means separation into \( C^2 \) graphs, whereas in the conclusion it means separation into \( C^{1,\alpha} \) graphs. Easy examples show that in general, \( C^2 \) regularity need not hold in the conclusion.

In the present application (i.e. Theorem 4.2) however, where \( V \) is a stable limit \((g, 0)\)-varifold, \( C^2 \) regularity does hold (as we show in Section 4.4 below) because of the additional constraints imposed by Theorem 3.1, including the fact that the set \( E = \{u_\infty = 1\} \) (notation as in Theorem 3.1) is a Caccioppoli set.

Note also that in the abstract form above Theorem 4.4 requires no stability hypothesis, but in the present application stability plays a key role. Indeed, in the proof of Theorem 4.2, we employ an inductive argument in which we apply Theorem 4.4 to a stable limit \((g, 0)\)-varifold \( V \) near a point where one of its tangent cones is a multiplicity \( q \) hyperplane. Validity of hypotheses (b) and (c) of Theorem 4.4 in this application is a consequence of stability of the critical points \( u_{\varepsilon_j} \) associated with \( V \).

Of fundamental importance in verifying hypothesis (c) in this setting is the existence of generalized second fundamental form of \( V \) satisfying a stability inequality (inequality [6] below) as well as the Schoen–Tonegawa inequality (Lemma 4.3 below). These inequalities are respectively the analogues of the usual stability inequality and the Schoen inequality ([SchSim81, Lemma 1]) for embedded stable hypersurfaces, with the important difference that for stable limit \( g \)-varifolds both inequalities hold for ambient test functions, with the ambient gradient (i.e. gradient on \( N \)) of the test function appearing where the intrinsic gradient (i.e. gradient on \( V \)) appears in the classical counterparts. These inequalities, which are derived independently of each other unlike in the classical setting (in which the Schoen inequality is derived from the stability inequality), were first established by Tonegawa ([Ton05-1]) for the case of Euclidean ambient space and \( g = 0 \).

We next discuss adaptation of the arguments of [Ton05-1] to the Riemannian ambient space and for general \( g \). While for the stability inequality this adaptation is standard and has appeared in several places in the literature, the derivation of the Schoen–Tonegawa inequality in the Riemannian setting is more subtle and requires care, and the right choice of coordinates; we provide a complete account of the latter (Lemma 4.3) below.

### 4.3 Stability inequality and the Schoen–Tonegawa inequality

Let \( g \in C^{1,1}(N) \) with \( \sup_N |\nabla g| \leq \Gamma \), and let \( V \) be a stable limit \((g, 0)\)-varifold on \( N \), with associated sequence \( \varepsilon_j \to 0^+ \), functions \( g_{\varepsilon_j} \in C^{1,1}(N) \) converging
locally in $C^{1,1}$ to $g$, and associated sequence $(u_{\varepsilon j})$ of critical points of $F_{\varepsilon j,\sigma g_{\varepsilon j}}$. For each $j$, the function $u_j$ satisfies the Euler–Lagrange equation
\[ -\varepsilon_j \Delta u_{\varepsilon j} + \varepsilon_j^2 W'(u_{\varepsilon j}) = g_{\varepsilon j} \]
weakly on $N$. By elliptic regularity, $u_{\varepsilon j} \in C^{3,\alpha}(N)$ for every $\alpha \in (0,1)$, and the equation (1) holds classically. Since $u_{\varepsilon j}$ is stable, $u_{\varepsilon j}$ additionally satisfies the inequality
\[ \int_N \varepsilon_j |\nabla \varphi|^2 + \varepsilon_j^{-1} W''(u_{\varepsilon j}) \varphi^2 \geq 0 \]
for all $\varphi \in C_0^1(N)$.

We first outline the well-known derivation of the stability inequality on $V$ following [Ton05-1]. (See e.g. [Hie18] for details in the case of Riemannian ambient space). Replacing $\varphi$ in (2) with $|\nabla u_{\varepsilon j}| \varphi$ (a valid choice by an approximation argument), and using in the resulting inequality the differentiated equation (1) (i.e. $-\varepsilon_j^2 \Delta u_{\varepsilon j} + W''(u_{\varepsilon j}) \nabla u_{\varepsilon j} = \varepsilon_j \nabla g$) and the Bochner identity
\[ \frac{1}{2} |\nabla u_{\varepsilon j}|^2 = |\nabla \varphi|^2 + <\nabla \Delta u_{\varepsilon j}, \nabla u_{\varepsilon j}> + \int \text{Ric}(\nabla u_{\varepsilon j}, \nabla u_{\varepsilon j}), \]
we obtain
\[ \int_N \left( \text{Ric}(\nabla u_{\varepsilon j}, \nabla u_{\varepsilon j}) + |\nabla u_{\varepsilon j}|^2 - |\nabla |\nabla u_{\varepsilon j}|^2| \right) \varepsilon_j \varphi^2 \]
\[ \leq \int_N |\nabla \varphi|^2 \varepsilon_j |\nabla u_{\varepsilon j}|^2 + \int <\nabla g_{\varepsilon j}, \nabla u_{\varepsilon j}> \varphi^2. \]
Integrating by parts in the last term on the right, this leads to
\[ \int_N \left( |\nabla u_{\varepsilon j}|^2 - |\nabla |\nabla u_{\varepsilon j}|^2| \right) \varepsilon_j \varphi^2 \leq c \int N \varphi^2 \varepsilon_j |\nabla u_{\varepsilon j}|^2 + \int N |\nabla \varphi|^2 \varepsilon_j |\nabla u_{\varepsilon j}|^2 \]
\[ - \int N \text{div} \left( \varphi^2 \nabla g_{\varepsilon j} \right) u_{\varepsilon j}, \]
where $c = c(N)$. This implies in the first instance that
\[ \int_{B_{\varepsilon_0}(0)} B^2_{\omega_{\varepsilon j}} \varphi^2 \varepsilon_j |\nabla u_{\varepsilon j}|^2 \leq C \sup \left( |\varphi|^2 + |\nabla \varphi|^2 \right) \]
for each $\varphi \in C_0^1(N)$, where $C = C(N, g, ||V|| (\text{spt } \varphi))$. Here, for a $C^3$ function $u$, the non-negative function $B_u$ is defined by
\[ B^2_u = |\nabla u|^2 - |\nabla |\nabla u|^2| \]
on $\{|\nabla u| > 0\}$ and $B_u = 0$ on $\{|\nabla u| = 0\}$. It follows from (5) (see also [Hie18]) that the limit varifold $V$ admits a (unique) generalized second fundamental form $B_V$, and passing to the limit in (4) we obtain that $V$ satisfies the following stability inequality:
\[ \int_N |B_V|^2 \varphi^2 d||V|| \leq \int_N |\nabla \varphi|^2 d||V|| + c \int_N \varphi^2 d||V|| \]
for all $\varphi \in C_0^1(N)$, where $c = c(N, \Gamma)$.

Remark 4.3. Note that in the preceding argument we have used the hypotheses $g_{\varepsilon j} \in C^{1,1}(N)$ and $g_{\varepsilon j} \to g$ in $C^{1,1}$ to integrate by parts in the last term in inequality (3) and to pass to the limit in inequality (4). These hypotheses will also be used again in a similar way in deriving inequality (15) below. Note also that in passing to the limit in the last term in (4), we have used the fact that $u_{\varepsilon j} \to u_{\infty}$ locally in $L^1$, $u_{\infty} = \pm 1$ a.e. in $N$ and that $\{u_{\infty} = 1\}$ is a Caccioppoli set.
Remark 4.4. Although we do not need it here, we note that by integrating by parts in the last term of (5) and passing to the limit as $\varepsilon_j \to 0^+$, we obtain the following “ambient” stability inequality:

$$
\int_N (\text{Ric}(\nu, \nu) + |B_V|^2) \varphi^2 d|V| \leq \int_N |\nabla \varphi|^2 d|V| + 2 \int_{N \cap \partial^* \{ u_{\infty} = 1 \}} \varphi^2 D_e g_d \mathcal{H}^n
$$

(7)

for all $\varphi \in C_c^1 (N \setminus \text{sing } V)$, where $\nabla$ is the gradient on $N$, and $\nu$ is the unit normal to $\partial^* \{ u_{\infty} = 1 \}$ “outward pointing” i.e. points away from $\{ u_{\infty} = 1 \}$.

We next obtain the Schoen–Tonegawa inequality, i.e. the Riemannian analogue of [Ton05-1, Theorem 4]. For an arbitrary point $x \in N$ we consider in the first instance normal coordinates centred at $x$, pulling back the Riemannian metric to a ball $B^{n+1}_0(x) \subset T_x N$ via the exponential map. Fix any hyperplane $P \subset T_x N$.

Now consider a Fermi system of coordinates centred at the disk $P \cap B^{n+1}_{\frac{3}{4}\text{inj}(N)}(x)$.

We will denote by $\{ x_1, \ldots, x_{n+1} \}$ this coordinate system. These coordinates are chosen so that $\{ x_1, \ldots, x_n, 0 \}$ agree with the normal coordinates restricted to $P \cap B^{n+1}_{\frac{3}{4}\text{inj}(N)}(x)$; on the other hand, the curve $s \in (-\sigma, \sigma) \to (x_1, \ldots, x_n, s)$ represents the geodesic orthogonal to $P \cap B^{n+1}_{\frac{3}{4}\text{inj}(N)}(x)$ at the point $(x_1, \ldots, x_n, 0)$, and $\sigma > 0$ is the semi-width of a tubular neighbourhood of $P \cap B^{n+1}_{\frac{3}{4}\text{inj}(N)}(x)$. This coordinate system covers therefore an open subset (a cylinder) in $B^{n+1}_{\frac{3}{4}\text{inj}(N)}(x) \subset T_x N$. Note that $\partial x_{n+1}$ is a unit vector field (for the Riemannian metric induced from $N$), that is determined (up to sign) at $x$ by the initial choice of $P$. Moreover, the coordinate chart has differential equal to the identity at the point $(0, \ldots, 0)$ (which is mapped to $x$). In view of this latter property, the area functional $A$, and more generally the functional $A - \text{Vol}_n$, can be written in the chosen Fermi coordinates by means of a non-parametric integrand $F$ as done in [BelWic-2, Section 3.1], maintaining the validity of [SchSim81] (1.2)-(1.5).

The following lemma is the Riemannian analogue of [Ton05-1, Proposition 4].

Lemma 4.1. Let $x \in N$ and $e \in T_x N$ be an arbitrarily chosen unit vector and fix the coordinate chart $B^n \times (-\sigma, \sigma)$ described in the previous discussion, with $(0, \ldots, 0)$ mapped to $x$ and with $\partial x_{n+1}$ identified with $e$ at 0. Let $g_e \in C^{1,1}(N)$ and $u_e : B^n \times (-\sigma, \sigma) \to \mathbb{R}$ satisfy $F^{e, g_e}(u_e) = 0$ and $F^{e, g_e, \nu}(u_e) \geq 0$. Then we have for any $\phi \in C_c^\infty (B^n \times (-\sigma, \sigma))$

$$
\int \varepsilon \left( |\nabla^2 u_e|^2 - |\nabla (\nabla u_e \cdot \partial_{n+1})|^2 - |\nabla |\nabla u_e - (\nabla u_e \cdot \partial_{n+1})\partial_{n+1}|^2 \right) \phi^2 \leq
$$

$$
\leq \int \varepsilon (|\nabla u_e|^2 - (\nabla u_e \cdot \partial_{n+1})^2) |\nabla \phi|^2 + \int \phi^2 \nabla g_e \cdot (\nabla u_e - (\nabla u_e \cdot \partial_{n+1})\partial_{n+1}) + (8)
$$

$$
+ \int \varepsilon \phi^2 \text{Ric}_N (\nabla u_e, \nabla u_e) + \int \phi^2 E,
$$

where $|E| \leq C_N \varepsilon |\nabla u_e|^2$. Here $\nabla^2 u_e$ denotes the Hessian (with respect to the Levi–Civita connection), $\nabla u_e$ and $\nabla g_e$ are metric gradients and the norms and scalar products are taken with respect to the Riemannian metric.

Remark 4.5. Recall that $|\nabla^2 u_e|^2 = \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \sum_{k=1}^{n+1} \sum_{l=1}^{n+1} g_{ik} g_{lj} u_{ik} u_{lj}$, where $g$ stands for the Riemannian metric tensor, $u_{ac} = \partial_e^a u - \Gamma^a_{ac} \partial_e u$ and $\Gamma^a_{ac}$ are the Christoffel symbols for the Levi–Civita connection.

---

\footnote{In [BelWic-2] we chose to work in an exponential chart, however the only property that was needed was the fact that $D\exp = \text{Id}$ at the centre of coordinates.}
Proof. For notational ease we will write, within this proof, \( u, g \) instead of \( u_c, g_c \). Given \( \phi \in C^\infty_c (B^n \times (-\sigma, \sigma)) \), we use the stability assumption \( \int \varepsilon |\nabla \varphi|^2 + W''(u)\varphi^2 \geq 0 \) with the choice of test function \( \varphi = \phi |\nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}| \). We remark that the full justification of the argument would require the use of the test function \( \varphi = \phi \partial^2 + |\nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}|^2 \) for \( \delta > 0 \), and then taking the limit \( \delta \to 0 \) (as in the proof of [Ton05-1, Proposition 4]). With this understood, we will set straight away \( \delta = 0 \), to make notation and ideas more transparent. We recall that \( \partial_{n+1} \) is a coordinate vector and has everywhere unit length. We get

\[
\int \varepsilon \phi^2 |\nabla \nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}|^2 + \int \varepsilon |\nabla \varphi|^2 (|\nabla u|^2 - (\nabla u \cdot \partial_{n+1})^2) + (9)
\]

\[
\int 2 \varepsilon \phi |\nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}| \nabla \varphi \cdot \nabla |\nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}| +
\]

\[
\int \frac{W''(u)}{\varepsilon} \varphi^2 (|\nabla u|^2 - (\nabla u \cdot \partial_{n+1})^2) \geq 0.
\]

Differentiating the PDE \( \varepsilon \partial u - W'(u) = -g \), we have \( \varepsilon \nabla \nabla u - \frac{W''(u)}{\varepsilon} \nabla u = -\nabla g \); taking the scalar product on both sides with \( \nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1} \) we obtain

\[
\varepsilon \nabla \nabla u (\nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}) + \nabla g (\nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}) =
\]

\[
= \frac{W''(u)}{\varepsilon} (|\nabla u|^2 - (\nabla u \cdot \partial_{n+1})^2) \quad (10)
\]

We substitute (10) in the last term of (9). Moreover we manipulate the third term of (11) by using the identity \( f \nabla f = \frac{1}{2} \nabla f^2 \) twice, and then performing a partial integration: this term then becomes \( - \int \varepsilon \phi^2 \Delta (|\nabla u|^2 - (\nabla u \cdot \partial_{n+1})^2) \). This leads to

\[
\int \varepsilon \phi^2 |\nabla \nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}|^2 + \int \varepsilon |\nabla \varphi|^2 (|\nabla u|^2 - (\nabla u \cdot \partial_{n+1})^2) + (11)
\]

\[
- \int \varepsilon \phi^2 \Delta |\nabla u|^2 + \int \varepsilon \phi^2 \Delta (\nabla u \cdot \partial_{n+1})^2 + \int \varepsilon \phi^2 \nabla \Delta u \nabla u - \int \varepsilon \phi^2 (\nabla u \cdot \partial_{n+1}) \nabla \Delta u \nabla u \partial_{n+1}
\]

\[
+ \int \phi^2 \nabla g (\nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}) \geq 0.
\]

Using Bochner’s identity (for the third and fifth terms of (11)) we get

\[
\int \varepsilon \phi^2 |\nabla \nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}|^2 + \int \varepsilon |\nabla \varphi|^2 (|\nabla u|^2 - (\nabla u \cdot \partial_{n+1})^2) + (12)
\]

\[
- \int \varepsilon \phi^2 |\nabla u|^2 - \int \varepsilon \phi^2 \text{Ric}_N (\nabla u, \nabla u) + \int \varepsilon \phi^2 \Delta (\nabla u \cdot \partial_{n+1})^2 -
\]

\[
- \int \varepsilon \phi^2 (\nabla u \cdot \partial_{n+1}) \nabla \Delta u \nabla u \partial_{n+1} + \int \phi^2 \nabla g (\nabla u - (\nabla u \cdot \partial_{n+1})\partial_{n+1}) \geq 0.
\]

We will now focus on the sixth term in (12). Firstly, we compute the commutator \( \nabla \Delta u \cdot \partial_{n+1} - \Delta (\nabla u \cdot \partial_{n+1}) = \nabla_{n+1} (\Delta u) - \Delta (\nabla_{n+1} u) \), where \( \nabla_j \) stands for the covariant derivative along the (coordinate) vector field \( \partial_j \). Writing \( \Delta = g^{ij} \nabla_i \nabla_j \) (this is the expression of the rough Laplacian, which agrees with the Laplace-Beltrami when applied to functions), and writing \( R \) for the curvature tensor, we compute

\[
\Delta (\nabla_{n+1} u) - \nabla_{n+1} (\Delta u) = g^{ij} \nabla_i \nabla_j (\nabla_{n+1} u) - \nabla_{n+1} (g^{ij} \nabla_i \nabla_j u) = \frac{\partial g}{\partial n} \text{ is parallel}
\]
We write \( g^{ij} \nabla_i \nabla_j u = g^{ij} \nabla_i \nabla_{n+1} \nabla_j u = g^{ij} \nabla_i \nabla_{n+1} \nabla_j u + g^{ij} R^a_{(n+1)j} \nabla_a u \).

We can thus replace \( \nabla \Delta u \cdot \partial_{n+1} \) (in the sixth term of (12)) with \( \Delta (\nabla u \cdot \partial_{n+1}) + \tilde{E} \), where \( |\tilde{E}| \leq C_N|\nabla u| \). Working then on the fifth and sixth terms in (12), noting that 

\[
\int \varepsilon \phi^2 (\nabla u \cdot \partial_{n+1}) \Delta (\nabla u \cdot \partial_{n+1}) = - \int \varepsilon \phi^2 \Delta (\nabla u \cdot \partial_{n+1})^2 + \int \varepsilon \phi^2 |\nabla (\nabla u \cdot \partial_{n+1})|^2,
\]

we obtain

\[
\int \varepsilon \phi^2 |\nabla \nabla u - (\nabla u \cdot \partial_{n+1}) \partial_{n+1}||^2 + \int \varepsilon |\nabla \phi|^2 (|\nabla u|^2 - (\nabla u \cdot \partial_{n+1})^2) + (13)
\]

\[
- \int \varepsilon \phi^2 |\nabla^2 u|^2 - \int \varepsilon \phi^2 \text{Ric}_N(\nabla u, \nabla u) + \int \varepsilon \phi^2 |\nabla (\nabla u \cdot \partial_{n+1})|^2 -
\]

\[
- \int \varepsilon \phi^2 \tilde{E}(\nabla u \cdot \partial_{n+1}) + \int \varepsilon \phi^2 \varepsilon \tilde{g} (\nabla u - (\nabla u \cdot \partial_{n+1}) \partial_{n+1}) \geq 0,
\]

from which (8) follows immediately. \( \square \)

The next lemma provides the geometric significance of the integrand on the left-hand-side of (8). Let \( p \in B^n \times (-\sigma, \sigma) \) be such that \( \nabla u \neq 0 \). Choose a normal system of coordinates \((y_1, \ldots, y_{n+1})\) centred at \( p \), with \( \partial_{y_{n+1}} = \partial_{y_{n+1}} \) at \( p \) and \( \nabla u \in \text{span}\{\partial_{y_n}, \partial_{y_{n+1}}\} \) at \( p \). In analogy with the notation in [Ton05-1] we set

\[
|\tilde{B}^e|^2(p) = \sum_{i=1}^{n-1} \sum_{j=1}^{n+1} |\partial^2_j u_{e_i}|^2(p),
\]

where the partial derivatives are computed with respect to the coordinates \((y_1, \ldots, y_{n+1})\). Then \( |\tilde{B}^e|^2(p) \) is a well-defined function on \((B^n \times (-\sigma, \sigma)) \setminus \{\nabla u = 0\}\) (it does not depend on the choice at \( p \) of \( \partial_{y_j} \) for \( j \in \{1, \ldots, n-1\} \)).

Remark 4.6. In the sequel, \( |\tilde{B}^e|^2 \) will always appear multiplied by \( |\nabla u_{e_i}|^2 \) and integrated with respect to \( dH^{n+1} \cup (B^n \times (-\sigma, \sigma)) \). Therefore, we can view \( |\tilde{B}^e|^2 \) as a summable function with respect to the measure \( |\nabla u_{e_i}|^2 dH^{n+1} \cup (B^n \times (-\sigma, \sigma)) \) (regardless of the fact that it is undefined where \( \nabla u_{e_i} = 0 \)). In fact, later on, when taking the \( \varepsilon \to 0 \) limit in (8), we will treat \( |\tilde{B}^e|^2 \) and \( \varepsilon |\nabla u_{e_i}|^2 \) as a measure-function pair (as in [Ton05-1]).

Lemma 4.2. We have, on \((B^n \times (-\sigma, \sigma)) \setminus \{\nabla u_{e_i} = 0\}, \)

\[
|\nabla^2 u_{e_i}|^2 - |\nabla (\nabla u_{e_i} \cdot \partial_{n+1})||^2 - |\nabla |\nabla u_{e_i} - (\nabla u_{e_i} \cdot \partial_{n+1}) \partial_{n+1}||^2 = (14)
\]

\[
= |\nabla u_{e_i}|^2 |\tilde{B}^e|^2 + \tilde{E},
\]

where \( |\tilde{E}| \leq C_N|\nabla u_{e_i}|^2 \).

Proof. We write \( \nu = \nabla u_{e_i} / \|\nabla u_{e_i}\| \) in a neighbourhood of \( p \) and let \( \{e_1, \ldots, e_n, e_{n+1}\} \) be a smooth orthonormal frame in a neighbourhood of \( p \), chosen so that \( e_{n+1} \) agrees with \( \partial_{n+1} \) in the neighbourhood and \( e_n \) is determined at \( p \) by the condition \( \nu \in \text{span}\{e_n(p), e_{n+1}(p)\} \). We set \( \nu_j = \nu \cdot e_j \), so that \( \nu = \sum_{j=1}^{n+1} \nu_j e_j \). Then

\[
|\nabla |\nabla u_{e_i} - (\nabla u_{e_i} \cdot \partial_{n+1}) \partial_{n+1}||^2 = \left| \nabla \left( \sum_{j=1}^{n+1} \nu_j^2 |\nabla u_{e_i}| \right) \right|^2 =
\]

\[
\left| \sum_{j=1}^{n+1} \nu_j \nabla \nu_j \right| |\nabla u_{e_i}| + \left| \sum_{j=1}^{n+1} \nu_j^2 \nabla |\nabla u_{e_i}| \right|^2.
\]}
Evaluating at $p$, at which $\nu_j = 0$ for $j \in \{1, \ldots, n-1\}$, this expression becomes

$$\frac{|\nu_n \nabla u|}{|\nu_n|} |\nabla u| + |\nu_n| \nabla |\nabla u| = |(\nabla \nu_n) \nabla u| + \nu_n \nabla |\nabla u|.$$ 

This says (by writing $\nabla (\nabla u \cdot e_n) = \nabla (|\nabla u| \nu_n)$) that

$$|\nabla|\nabla u| - (\nabla u \cdot \partial_{n+1}) \partial_{n+1}||² (p) = |\nabla (\nabla u \cdot e_n)||² (p),$$

with the given choice of $\{e_1, \ldots, e_n, e_{n+1}\}$. Note that $|\nabla (\nabla u \cdot e_n)||² (p)$ only depends on the choice of the (unit) vector field $e_n$ (rather than the whole frame). We have thus obtained that, when evaluated at $p$, each of the two terms that are subtracted from $|\nabla² u||²$ in (4) is of the form $|\nabla (\nabla u e)|² (p)$, for a certain unit vector field $e$.

We write such a unit vector field $e$ in coordinates (around $p$) as $e = \eta^i \partial_i$. Then $\nabla (\nabla u e) = \nabla (\eta^i \partial_i u) = g^{ab} \partial_b (\eta^i \partial_i u) \partial_a$. If the coordinate system is orthonormal at $p$, taking the norm we get (with all the terms evaluated at $p$)

$$|\nabla (\nabla u e)|² = \sum_{a=1}^{n+1} (\partial_a \eta^i \partial_i u + \eta^i \partial_a^2 u)^2$$

$$= \sum_{a=1}^{n+1} [(\partial_a \eta^i)^2 (\partial_i u)^2 + \partial_u \eta^i \partial_a u \partial_a^2 u + (\eta^i)^2 (\partial_a^2 u)^2].$$

If the coordinate system is moreover chosen so that $\partial_a (p) = e(p)$ we obtain at $p$

$$|\nabla (\nabla u e)|² = \sum_{a=1}^{n+1} (\partial_a \eta^i)^2 (\partial_i u)^2 + 2 \partial_u \eta^i \partial_a u \partial_a^2 u + \sum_{a=1}^{n+1} (\partial_a^2 u)^2.$$ 

Finally, we note that if the coordinate system is chosen to be normal at $p$ (with coordinate frame orthonormal at $p$ and with $\partial_a = e$ at $p$) we have $(\partial_a \eta^a)(p) = 0$. This follows by expanding $0 = \partial_a |e|² = \partial_a (g_{ij} \eta^i \eta^j)$ and using the vanishing of the partial derivatives of $g_{ij}$ at $p$, which gives $2g_{ij} \partial_u \eta^i \eta^j = 0$. Evaluating at $p$, where $\eta^i = 0$ for $i \neq n$ and $\eta^n = 1$, we find $(\partial_a \eta^a)(p) = 0$. In conclusion, at $p$ we have the following equality:

$$|\nabla (\nabla u e)|² = \sum_{a=1}^{n+1} (\partial_a \eta^i)^2 (\partial_i u)^2 + \sum_{a=1}^{n+1} (\partial_a^2 u)^2$$

in normal coordinates at $p$ chosen so that $e(p) = \partial_n (p)$. With the same argument we get

$$|\nabla (\nabla u_{n+1} e)|² = \sum_{a=1}^{n+1} (\partial_a \tilde{\eta}^i)^2 (\partial_i u)^2 + \sum_{a=1}^{n+1} (\partial_a^2 u)^2$$

in normal coordinates at $p$ chosen so that $e_{n+1}(p) = \partial_{n+1}(p)$ (here $\tilde{\eta}^i$ denote the coordinates of $e_{n+1}$). Recalling (see Remark [4,5]) that, in normal coordinates at $p$, we have $|\nabla² u||² = \sum_{i,j=1}^{n+1} (\partial_i^2 u)^2$, the claim is proved.

We will now analyse the limiting behavior of (8) as $\varepsilon = \varepsilon_j \to 0^+$, under the assumption that $g_{\varepsilon} \to g$ in $C^{1,1}$. We begin with the second term on the right-hand-side of (8). Recall that $u_{\varepsilon} \to u_{\infty}$ in $L^1(N)$ (up to a subsequence that we
keep implicit) and that $u_\infty \in BV(N)$ and takes values in $\{-1, +1\}$.

\[
\int \phi^2 \nabla g \cdot \nabla u_\varepsilon - \int \phi^2 (\nabla g \cdot \partial_{n+1}) (\nabla u_\varepsilon \cdot \partial_{n+1}) = \int \phi^2 \nabla g \cdot \nabla u_\varepsilon - \int \phi^2 (\nabla g \cdot \partial_{n+1}) \text{div}(u_\varepsilon \cdot \partial_{n+1}) + \int \phi^2 (\nabla g \cdot \partial_{n+1}) u_\varepsilon \text{div}\partial_{n+1}
\]

\[
= - \int \text{div}(\phi^2 \nabla g) u_\varepsilon + \int u_\varepsilon \nabla (\phi^2 (\nabla g \cdot \partial_{n+1})) \cdot \partial_{n+1} + \int \phi^2 (\nabla g \cdot \partial_{n+1}) u_\varepsilon \text{div}\partial_{n+1}.
\]

Since $g_\varepsilon \to g$ in $C^{1,1}(N)$, in all the above terms the function $u_\varepsilon$ multiplies an $L^\infty$-bounded function and therefore we can pass to the limit as $\varepsilon \to 0$ to obtain (we denote by $\partial^*$ the reduced boundary and by $\hat{n}$ the measure theoretic normal)

\[
- \int \text{div}(\phi^2 g) u_\infty + \int u_\infty \nabla (\phi^2 (\nabla g \cdot \partial_{n+1})) \cdot \partial_{n+1} + \int \phi^2 (\nabla g \cdot \partial_{n+1}) u_\infty \text{div}\partial_{n+1}
\]

\[
= \int_{\partial^\ast (u_\infty = +1)} \phi^2 g \cdot \hat{n} + \int u_\infty \text{div} (\phi^2 (\nabla g \cdot \partial_{n+1}) \partial_{n+1})
\]

\[
= \int_{\partial^\ast (u_\infty = +1)} \phi^2 g \cdot \hat{n} + \int_{\partial^\ast (u_\infty = +1)} \phi^2 (\nabla g \cdot \partial_{n+1}) \partial_{n+1} \cdot \hat{n}.
\]

We have obtained

\[
\limsup_{\varepsilon \to 0} \left| \int \phi^2 \nabla g \cdot (\nabla u_\varepsilon - (\nabla u_\varepsilon \cdot \partial_{n+1}) \partial_{n+1}) \right| \leq C_g \int \phi^2 d||V|| \quad (15)
\]

where the constant $C_g$ depends only on an upper bound on $\sup_N |\nabla g|$.

**Remark 4.7.** For the last two terms on the right-hand-side of (8), on the other hand, it suffices to recall that $\varepsilon |\nabla u_\varepsilon|^2 dH^{n+1} \to ||V||$, so that, in the $\varepsilon \to 0$ limit, these two terms will be controlled by $C_N \int \phi^2 d||V||$. Similarly, recalling Lemma 4.2, leaving $\int \varepsilon |\nabla u_\varepsilon|^2 + \phi^2 |\tilde{B}^\varepsilon|^2$ on the left-hand-side of (8) and moving the term $\int \phi^2 \tilde{E}$ to the right-hand-side, we find that this latter term is also controlled by $C_N \int \phi^2 d||V||$ in the limit.

In view of (15) and of Remark 4.7, we can rewrite (8) as

\[
\int \varepsilon |\nabla u_\varepsilon|^2 \phi^2 |\tilde{B}^\varepsilon|^2 \leq \int \varepsilon (|\nabla u_\varepsilon|^2 - (\nabla u_\varepsilon \cdot \partial_{n+1})^2) |\nabla \phi|^2 + \text{other terms} \quad (16)
\]

and $\limsup_{\varepsilon \to 0} |\text{other terms}| \leq C_N \phi d||V||$. The only terms that are left to deal with, in the $\varepsilon \to 0$ limit, are therefore $\int \varepsilon |\nabla u_\varepsilon|^2 \phi^2 |\tilde{B}^\varepsilon|^2$ and the first term on the right-hand-side. The argument for these follows [Ton05-1] closely.

Using Lemma 4.2, we have, in the first instance (arguing similarly to [Ton05-1, Proposition 5]) that the measure-function pairs $(|\tilde{B}^\varepsilon|, V)$ satisfy a uniform $L^2$ bound and therefore there exists a well-defined (subsequential) limit $(|\tilde{B}|, V)$, where $V$ is the varifold limit of $V^\varepsilon$. Then arguing (pointwise) as in [Ton05-1, Lemma 1] and using Lemma 4.2 again, it follows that there exists $c = c(n) > 0$ such that $|B_V|^2 \leq c(|\tilde{B}|^2 + g^2)$, where $B_V$ is the generalized second fundamental form of $V$ as in (6). For the first term on the right-hand-side of (16), we also argue as in the (final line of the) proof of [Ton05-1, Proposition 5]. These arguments, together with (15) and Remark 4.7, give the following:
Lemma 4.3 (Schoen–Tonegawa inequality). Let $g \in C^{1,1}(N)$ with $\sup_N |\nabla g| \leq \Gamma$, and let $V$ be a stable limit $(g,0)$-varifold on $N$. Let $y \in N$ and let $B^n \times (-\sigma,\sigma) \subset T_y N$ be a Fermi coordinate neighborhood as described in the paragraph preceding Lemma 4.7. Let $\tilde{V}$ be the pullback of $V$ under this coordinate map. Then for any $\phi \in C^\infty_c(B^n \times (-\sigma,\sigma))$,

$$
\int \phi^2 |B_{\tilde{V}}| d\|\tilde{V}\| \leq c \int |\nabla \phi|^2 (1 - (\tilde{\nu} \cdot \partial_{n+1})^2) d\|\tilde{V}\| + C \int \phi^2 d\|\tilde{V}\| \tag{17}
$$

where $B_{\tilde{V}}$ is the generalized second fundamental form of $\tilde{V}$ (with respect to the pullback metric from $N$ under the coordinate map), $\tilde{\nu}(x)$ is a choice of unit normal to $T_x \tilde{V}$ (which exists for $\|\tilde{V}\|$-a.e. $x \in \text{spt} \|\tilde{V}\|$), $c = c(N)$ and $C = C(N,\Gamma)$.

Remark 4.8. All quantities in (17) are Riemannian. Recalling the choice of Fermi coordinates, we see that $\nu \cdot \partial_{n+1}$ is the same as the Euclidean scalar product of the vector $\nu$ with $\partial_{n+1}$ (because $\partial_{n+1}$ is orthogonal to $\partial_j$ for $j \neq n+1$ for the Riemannian scalar product). Therefore, upon, say, doubling the constants on the right-hand-side of (17), we can write the same inequality with gradients, scalar products, norms and second fundamental forms replaced with the corresponding quantities computed with respect to the Euclidean metric on $B^n \times (-\sigma,\sigma)$.

4.4 Regularity of stable limit $(g,0)$-varifolds, Part I: $C^{1,\alpha} \implies C^2$

In this section we show that if $g \in C^{1,1}(N)$ with $g > 0$, and if a stable limit $(g,0)$-varifold $V$ in some neighborhood of a point $y$ is given by a union of ordered $C^{1,\alpha}$ graphs for some $\alpha \in (0,1)$, then each of the graphs is of class $C^2$ (and hence, by elliptic regularity, are of class $C^{3,\alpha}$ for any $\alpha \in (0,1)$) and in fact $V$ has quasi-embedded PMC($g,0$) structure near $y$. This is the content of Theorem 4.5 below. The arguments are adapted from those in [BelWic-1] and [BelWic-2], the difference being that in the present case we have a quasi-embedded PMC($g,0$) structure, rather than mean curvature prescribed by $g$ (as a consequence, 3-fold touching singularities are allowed, as in the last picture in Figure 2 while only 2-fold touching singularities could arise in [BelWic-1], [BelWic-2]). We also point out that, by using local stability for the Allen–Cahn solutions $u_c$, we can simplify some of the steps carried out in [BelWic-1], [BelWic-2]; specifically, we will be able to obtain $W^{2,2,\nu}$-estimates directly from (6), while in [BelWic-1] Section 7.4 these estimates require additional work; this is due to the fact that the stability hypothesis in [BelWic-1], [BelWic-2] is only on gen-reg $V$ (where $C^2$ regularity is assumed) and hence there is no information on the second fundamental form on the second fundamental form on the coincidence set away from gen-reg $V$.

We recall, to begin with, that smooth prescribed-mean-curvature hypersurfaces are locally described by graphs of functions that solve a PDE of mean curvature type. Assume that $u : B^n_p(x) \to \mathbb{R}$ is $C^2$ and there exists a local system of normal coordinates centred at $p \in N$ and such that $B_p^n(x) \times I \subset \exp_p^{-1}(U)$ for some open interval $I$ that contains $[\min u, \max u]$, and that the mean curvature $H$ of $(\exp_p)_\# \text{graph}(u)$ satisfies $|H| = g$ everywhere or $H = 0$ everywhere. Then $u$ satisfies one of the following three PDEs: for any $\zeta \in C^\infty_c(B^n_p(x))$

$$
- \int F_j(x,u,Du)D_j \zeta + \int F_{n+1}(x,u,Du) \zeta = \begin{cases} 
\int g(x,u) \zeta 
& \text{if } H > 0 

\int -g(x,u) \zeta 
& \text{if } H < 0 

0 & \text{if } H = 0 
\end{cases} \tag{18}
$$

where $F_j(x,u,Du) = \left( \frac{\partial F}{\partial y_j} \right)((x,u),(Du,-1))$ and $F : \exp_p^{-1}(U) \times \mathbb{R}^{n+1} \to \mathbb{R}$ is as in [BelWic-2] Section 3 and [SchSim81], and satisfies conditions [SchSim81].
(1.2)-(1.5)]. The notation $D_j$ stands for $\frac{\partial}{\partial x_j}$ for $j \in \{1, \ldots, n\}$. Conditions 
\[\text{SchSim81} (1.2)-(1.5)\] guarantee in particular that $D_iF_j$ forms a positive definite matrix, so that the above are (uniformly) elliptic quasilinear PDEs. The function $g$ appearing on the right-hand-side of the first two PDEs in [18] is a $C^{1,1}$ function in $\exp_p^{-1}(U)$ obtained by composing the original $g|_U$ with $\exp_p$, and multiplying by the Riemannian volume element in normal coordinates (see [BelWic-2 Section 3.1]); by abuse of notation, we denote the resulting function by $g$. The PDEs in [18] are written in their weak formulation (with an implicit summation over repeated indices), however they are satisfied in the strong sense as well, since $u$ is $C^2$. The three PDEs in [18] correspond respectively to the cases where $\exp_p(\text{graph}(u))$ is a hypersurface with mean curvature $g\nu$, $-g\nu$ and 0, where $\nu$ is the unit normal that has positive scalar product with the (pushforward via $\exp_p$) of the upwards direction in the cylinder $B^n_p(x) \times I$. For this reason, when $u$ satisfies the first, middle or the last of the PDEs in [18], we will say that $\text{graph}(u)$ is PMC with mean curvature $g$, or PMC with mean curvature $-g$ or minimal respectively.

The main result of this section is the following:

**Theorem 4.5.** Let $g \in C^{1,1}(N)$ be positive. Let $y \in N$ be an arbitrary point, $\mathcal{O} = B^n_1 \times \mathbb{R} \subset T_y N \approx \mathbb{R}^{n+1}$ and let $V$ be the varifold on $\mathcal{O}$ obtained by applying an appropriate rescaling to the pull back of a (part of a) stable limit $(g, 0)$-varifold on $N$ by the exponential map $\exp_y$. Suppose that $\mathcal{VLO} = \sum_{j=1}^{q} |\text{graph}(u_j)|$, where $q \in \mathbb{N}$, $u_j : B^n_1 \to \mathbb{R}$ are of class $C^{1,\alpha}$ for some $\alpha \in (0, 1)$ and $u_1 \leq u_2 \leq \ldots \leq u_q$. Suppose further that there exists a point $y \in \text{spt} \|V\| \cap \mathcal{O}$ such that density $\Theta(||V||, y) = q$. Then, for each $j \in \{1, \ldots, q\}$, $u_j$ is of class $C^{2,\alpha}$ (and the scalar mean curvature of $\text{graph}(u_j)$ with respect to the upward pointing unit normal is either zero everywhere, or is equal to $g(x)$ for every $x \in \text{graph}(u_j)$, or is equal to $-g(x)$ for every $x \in \text{graph}(u_j)$). Moreover, $\text{spt} \|V\|$ is the union of the graphs of at most three $u_j$'s. More precisely:

(i) if $q$ is even, then either all of the $u_j$'s coincide and $\text{graph}(u_j)$ are minimal; or $\text{graph}(u_1)$ is PMC with mean curvature $-g$, $\text{graph}(u_q)$ is PMC with mean curvature $g$ and (if $q \geq 2$) $u_2 = \ldots = u_{q-1}$ with $\text{graph}(u_j)$ minimal for $2 \leq j < q - 1$;

(ii) if $q$ is odd, then either $\text{graph}(u_1)$ is PMC with mean curvature $-g$ and (if $q \geq 3$) $u_2 = \ldots = u_q$ with $\text{graph}(u_j)$ minimal for $2 \leq j < q$; or $\text{graph}(u_q)$ is PMC with mean curvature $g$ and (if $q \geq 3$) $u_1 = \ldots = u_{q-1}$ with $\text{graph}(u_j)$ minimal for $1 \leq j < q - 1$.

The strategy of the proof of Theorem 4.5 is to use the stability hypothesis and Theorem 3.1 to show that each $u_j$, which satisfies one of the above PDEs where it is $C^2$, must in fact be a weak solution (to one of the above PDEs) on the entire domain. Once this is done, standard elliptic theory implies the conclusion. As a preliminary step, we will need the following lemma.

**Lemma 4.4** (Absence of $\ell$-fold touching singularities for $\ell \geq 4$). Let $y \in N$ be an arbitrary point, $\mathcal{O} = B^n_1 \times \mathbb{R} \subset T_y N \approx \mathbb{R}^{n+1}$ and let $V$ be the varifold on $\mathcal{O}$ obtained by applying an appropriate rescaling to the pull back of a (part of a) limit $(g, 0)$-varifold on $N$ by the exponential map $\exp_y$. Suppose that $\mathcal{VLO} = \sum_{j=1}^{q} |\text{graph}(u_j)|$, where $q \in \mathbb{N}$, $u_j : B^n_1 \to \mathbb{R}$ are of class $C^{1,\alpha}$ for some $\alpha \in (0, 1)$ and $u_1 \leq u_2 \leq \ldots \leq u_q$. Suppose further that there exists a point $y \in \text{spt} \|V\| \cap \mathcal{O}$ such that density $\Theta(||V||, y) = q$ and that if $\Theta(||V||, x) \leq q - 1$ then there exists a neighbourhood of $x$ in $\mathcal{O}$ in which $V$ has quasi-embedded PMC$(g, 0)$ structure.

Then $\text{spt} \|V\| \cap \mathcal{O} = \cup_{j=1}^{q} \text{graph}(\tilde{u}_j)$, where $\tilde{q} \in \{1, 2, 3\}$, $\tilde{q} \leq q$, and $\tilde{u}_j \in C^{1,\alpha}(B^n_1)$. Moreover, if $\tilde{q} \geq 2$ then $\tilde{u}_j \leq \tilde{u}_{j+1}$ on $B^n_1$ for $j \in \{1, \tilde{q} - 1\}$, and there exists $x \in B^n_1$ such that $\tilde{u}_j(x) < \tilde{u}_{j+1}(x)$ for $j \in \{1, \tilde{q} - 1\}$. (In other
words, \( spt \|V\| \cap \mathcal{O} \) is the union of the graphs of at most three distinct ordered \( C^{1,\alpha} \) functions and is embedded in some non-empty open cylinder \( \Omega \times \mathbb{R} \subset \mathcal{O} \).

**Proof.** We note that the graph structure implies that \( \Theta(\|V\|, \cdot) \) is integer-valued everywhere on \( spt \|V\| \) (not just almost everywhere); the proof is obtained as in BelWic-1 Lemma A.2. Let \( \pi : \mathcal{O} \to B = B^n_{\rho} \) be the projection onto the first factor and let \( C = \pi(\{x \in spt \|V\| : \Theta(x,\|V\|) = \bar{q}\}) \). The set \( C \) is closed. If \( B \setminus C = \emptyset \), then \( u_j = u_1 \) for all \( j \) and \( spt \|V\| \) is a single graph, so the conclusion of the lemma holds with \( \bar{q} = 1 \). From now we therefore assume that \( B \setminus C \neq \emptyset \).

We denote by \( U \) the subset of \( B \setminus C \) made up of points \( p \) such that \( spt \|V\| \) is an embedded hypersurface at all the points in \( \pi^{-1}(p) \cap spt \|V\| = \bigcup_{j=1}^{k} \{(p, u_j(p))\} \) (there may be repeated points in the last expression). The set \( U \) is open and, by the regularity assumption on points with density \( \leq q - 1 \), \( U \) is dense in \( B \setminus C \) (the set of non-embedded points with multiplicity \( \leq q - 1 \) projects locally to a finite union of submanifolds with dimension at most \( n - 1 \)). We define on \( U \) the function \( \tilde{Q} \) that assigns to \( p \in U \) the number of distinct points in the set \( \pi^{-1}(p) \cap spt \|V\| \). By the embeddedness requirement that characterizes \( U \), the function \( \tilde{Q} \) is locally constant on \( U \). Clearly \( \tilde{Q} \leq q \) on \( U \).

We claim, in a first instance, that \( \tilde{Q} \) extends to a locally constant function on \( B \setminus C \). To see this, let \( a \in B \setminus C \). Then there exists at least a point (and at most \( q/2 \) points) \( x \in \pi^{-1}(a) \cap spt \|V\| \) at which \( spt \|V\| \) is not embedded; at such a point \( \Theta(x,\|V\|) \leq q - 1 \). By hypothesis, there exists a neighbourhood \( \mathcal{O}_x \) of \( x \) contained in \( \mathcal{O} \) such that \( V \cap \mathcal{O}_x \cap \mathcal{O}_x \) has one of the structures (iii), (iv) or (v) in Definition 2. Denote by \( \{x_1, \ldots, x_K\} \) the distinct points in \( \pi^{-1}(a) \cap spt \|V\| \) at which \( spt \|V\| \) is not embedded and by \( \{x_{K+1}, \ldots, x_M\} \) the points in \( \pi^{-1}(a) \cap spt \|V\| \) at which \( spt \|V\| \) is embedded. Let \( B_a \) be an open ball contained in \( B \) such that \( spt \|V\| \cap (B_a \times I) \) is the union of exactly \( M \) connected sets. Note that \( M - K \) of these are embedded disks. From the characterization of the structures (iii), (iv) or (v) we can check the following fact: \( Q(p) \) is an integer independent of \( p \in B_a \cap U \) (this follows by counting the distinct points in \( \pi^{-1}(p) \) for each of the five possible structures). Therefore, recalling that \( U \) is dense in \( B \setminus C \), \( \tilde{Q} \) extends (from \( U \)) to a locally constant integer-valued function on \( B \setminus C \), proving our first claim.

We next define on each connected component \( U_c \) of \( U \) the functions \( u_1, \ldots, u_{\tilde{Q}} \) such that \( u_j < u_{j+1} \) for all \( j \in \{1, \ldots, \tilde{Q} - 1\} \) and such that \( spt \|V\| \cap (U_c \times I) = \bigcup_{j=1}^{\tilde{Q}} \text{graph}(u_j) \). (Note that \( u_1 \) necessarily agrees with \( \bar{u}_1 \).) Let \( B_c \) be a connected component of \( B \setminus C \). By our first claim, \( \tilde{Q} \) does not vary among the connected components of \( U \) that lie in \( B_c \). This implies that the functions \( u_1, \ldots, u_{\tilde{Q}} \), initially defined only on \( U \cap B_c \) for a common \( \tilde{Q} \), can be extended continuously from \( U \cap B_c \) to \( B_c \), giving rise to \( \tilde{Q} \) ordered functions that we still denote by \( u_1 \leq \ldots \leq u_{\tilde{Q}} \). Moreover, \( u_1, \ldots, u_{\tilde{Q}} \) are \( C^2 \) on \( B_c \) by the characterization of the structures (iii), (iv) or (v) of Definition 2. A priori, \( \tilde{Q} \) may depend on \( B_c \), and we will focus on a single connected component \( B_c \). By the characterization of the structures (iii), (iv) or (v), each graph(\( u_k \)) is either PMC with mean curvature \( q \) on \( B_c \), or PMC with mean curvature \( -q \) on \( B_c \), or is minimal on \( B_c \). (These three options correspond respectively to the fulfillment, by \( u_k \), of the three PDEs in (18).)

Our next claim is that \( \tilde{Q} \leq 3 \) on \( B_c \). We let \( b \in U \cap B_c \) and consider an open ball \( D \) contained in \( B_c \), centred at \( b \) and such that \( \partial D \cap C \) is not empty. Let \( p \in \partial D \cap C \), then all the \( u_j \) extend continuously to \( p \) with the same value (= \( u_1(p) = \ldots = u_{\tilde{Q}}(p) \)). If \( \tilde{Q} \geq 4 \) we find a contradiction to Hopf boundary point lemma, as follows. If \( \tilde{Q} \geq 4 \) then there exist two indices \( j_1 \neq j_2 \) for which \( u_{j_1} \) and \( u_{j_2} \) both solve the same of the three PDEs in (18). We let \( v = u_{j_1} - u_{j_2} \) and
compute, following a standard argument, the PDE satisfied by $v$ and obtain, for any $\zeta \in C_c^\infty(B_\epsilon)$
\[- \int (a_{ij} D_i v + b_i v) D_j \zeta + \int (d_i D_i v + c v) \zeta = \begin{cases} \int f v \zeta, \\ 0 \end{cases},\]
where
\[a_{ij} = \int_0^1 \frac{\partial F}{\partial p}^j(x, (1 - s)u_1 + su_2, (1 - s)Du_1 + sDu_2)ds,\]
and $b_i, c, d_i$ are similarly defined by integration of the composition of a smooth function (depending on $F$ or $DF$) with $(x, u_1, D_u u_1, u_{j_1}, Du_{j_1})$. Such as, $a_{ij}, b_i, c, d_i$ are in $C^{0,1}(B_\epsilon)$. The left-hand-side of (19) is the difference of the left-hand-sides of (18) for $u_{j_1}$ and $u_{j_2}$. The right-hand-side of (19) is similarly obtained from the right-hand-sides of (18), an $f$ is the integration of the composition of a $C^{0,1}$ function (depending on derivatives of $g$) with $(x, u_{j_1}, u_{j_2})$ and therefore $f \in C^{0,1}(B_\epsilon)$. The function $a_{ij}$ is moreover symmetric in $i, j$ (by the definition of $F_j$ in terms of $\frac{\partial F}{\partial p}$, see (18) above) and $a_{ij}$ form a positive definite matrix by the conditions on $F$ [SchSim81, (1.2)-(1.5)], so the PDEs in (19) are uniformly elliptic. The first case in (19) arises when $u_{j_1}$ and $u_{j_2}$ either both solve the first or both solve the second of the PDEs in (18), while the second case in (19) arises when $u_{j_1}$ and $u_{j_2}$ both solve the third PDE in (18). Since $v \in C^2(B_\epsilon)$, the PDE is satisfied in the strong sense. Recall that all the graph($u_j$) intersect tangentially at the point $(p, u_1(p))$, therefore $v$ and $Dv$ extends continuously to $p$ with $v(p) = 0$, $Dv(p) = 0$. Hopf boundary point lemma (see Remark 4.9 below), applied to either of the two PDEs in (19), implies that $v \equiv 0$ on $D$, so $u_{j_1} = u_{j_2}$ on $D$, contradicting the initial definition of $u_j$ on $U$. We thus have $\tilde{Q} \leq 3$ on any connected component $B_\epsilon \setminus C$.

We then set $\tilde{q} = \max_{p \in U} \tilde{Q}$. Since we are working under the assumption $B \setminus C \neq \emptyset$, we have $\tilde{q} \geq 2$. The functions $u_1$ and $u_q$ (defined above on $B \setminus C$) agree respectively with (the restrictions to $B \setminus C$ of) $u_1$ and $u_q$. We set $\tilde{u}_1 = u_1$ and $\tilde{u}_q = u_q$ on $B$. If $\tilde{q} = 2$, then $\arg\max \|\nabla\| = \text{graph}($\tilde{u}_1$) \cup \text{graph}($\tilde{u}_q$) and all the conclusions of the lemma are satisfied. The last case to consider is $\tilde{q} = 3$. In this case, we let $\tilde{u}_2 = u_2$ on the (non-empty) open set $A$ given by the union of the connected components of $B \setminus C$ on which $\tilde{Q} = 3$. The function $\tilde{u}_2$ is $C^2(A)$ and we extend it to $B$ by setting it equal to $\tilde{u}_1$ on $B \setminus A$. The resulting function, still denoted $\tilde{u}_2$, is in $C^{1,\alpha}(\tilde{B})$: this only needs to be checked at points in $C$, since $\tilde{u}_2$ is $C^2$ by construction on each connected components of $B \setminus C$. At point in $C$ the conclusion follows by recalling that all the $u_j$’s and their differentials agree on $C$, and that each $u_j$ is in $C^{1,\alpha}(\tilde{B})$. By construction, there exists at least one connected component of $U$ on which $u_1 < u_2 < u_3$, hence on the same non-empty open set $\tilde{u}_1 < \tilde{u}_2 < \tilde{u}_3$.

Remark 4.9. We have used Hopf boundary point lemma in its version that is valid regardless of the sign of the 0-th order term, since we know that $v(p) = 0$ (see the discussion that follows [GilTru Theorem 3.5]). A more careful use of Hopf boundary point lemma gives additional information on $\tilde{u}_{j_1}$ and $\tilde{u}_{j_2}$, as we will point out now. While not necessary for the conclusion of Lemma 4.4, this will be used within the proof of Theorem 1.5 below. We assume that $\tilde{u}_{j_1} \leq \tilde{u}_{j_2}$ are $C^2$ on a ball $D$, $\tilde{u}_{j_1} \not\equiv \tilde{u}_{j_2}$ and that $\tilde{u}_j$ and $D\tilde{u}_j$, for $j \in \{j_1, j_2\}$, extend continuously to $p \in \partial D$ with $\tilde{u}_{j_1}(p) = \tilde{u}_{j_2}(p)$ and $D\tilde{u}_{j_1}(p) = D\tilde{u}_{j_2}(p)$. We have ruled out, in Lemma 4.4, the possibility that these two functions solve the same PDE among the three in (18). More precisely, we will now show that the only possibilities for the right-hand-sides $\int h_j(x, u_j, \zeta)$ in (18) are those for which $h_{j_1}(x, \tilde{u}_{j_1}) < h_{j_2}(x, \tilde{u}_{j_2})$. Here $h_j = g$, $h_j = -g$, or $h_j = 0$ are the three options for $h_j$. This claim follows,
arguing by contradiction, i.e. assuming that \( h_{\tilde{j}_1}(x, \tilde{u}_{\tilde{j}_1}) > h_{\tilde{j}_2}(x, \tilde{u}_{\tilde{j}_2}) \), by taking the difference of the two PDEs, which gives for \( v = \tilde{u}_{\tilde{j}_1} - \tilde{u}_{\tilde{j}_2} \leq 0 \) (recall that \( g > 0 \))

\[
- \int (a_{ij} D_i v + b_i v) D_j \zeta + \int (d_{ij} D_i v + c v) \zeta \geq 0,
\]

(20)

whenever \( \zeta \in C^\infty_0(D) \) and \( \zeta \geq 0 \). Here the coefficients \( a_{ij}, b_i, d_i, c \) are as in the proof of Lemma 4.4. We have from (20) a contradiction to Hopf boundary point lemma, because \( v(p) = 0, Dv(p) = 0 \) and \( v \neq 0 \) in \( D \).

**Remark 4.10.** We are not (yet) ruling out the possibility that \( \tilde{Q} \) in the proof of Lemma 4.4 (i.e. the number of distinct graphs needed to describe \( \text{spt} \|V\| \)) may change from one connected component of \( B \setminus C \) to another. This will be accomplished in the proof of Theorem 4.5 based on the fact that \( V \) is a limit \( g \)-varifold (using Theorem 5.1).

For the proof of Theorem 4.5 it is convenient to analyse the possible configuration that can arise depending on the maximum value of distinct graphs needed to describe \( \text{spt} \|V\| \), i.e. the value of \( \tilde{q} \) in Lemma 4.4. In the proof of Theorem 4.5 we directly make use of the assumption that \( V \) is a limit \( g \)-varifold. This assumption was absent in Lemma 4.4, although it should be kept in mind that this lemma is used within an induction. To avoid confusion: the induction is carried out on \( q \), however the cases below are differentiated depending on \( \tilde{q} \). For a fixed \( \tilde{q} \), only some of the cases analysed below can actually happen. After the case by case analysis (on \( \tilde{q} \)) is complete, we will show that in any of the possible configurations identified, one can extend the validity of the PDE of mean curvature type (i.e. one of the equations in (18)) from the \( C^2 \) portion of each graph, to the full \( C^{1,\alpha} \) graph in the weak form.

**Proof of Theorem 4.5.** Let \( \tilde{q} \) be as in Lemma 4.4. Then \( \tilde{q} \leq 3 \). Let \( C = \pi(\{ x \in \text{spt} \|V\| : \Theta(x, \|V\|) = q \}) \) where \( \pi : B^1_0 \times \mathbb{R} \to B^1_0 \) is the orthogonal projection. We consider the cases \( \tilde{q} = 1, \tilde{q} = 2 \) and \( \tilde{q} = 3 \) separately.

\( \tilde{q} = 1 \). In this case \( \text{spt} \|V\| \) is a single \( C^{1,\alpha} \) graph, \( \text{spt} \|V\| = \text{graph}(\tilde{u}_1) = \text{graph}(u_1) \). Then we will show by means of [RogTon08], see Theorem 5.1 above, that it is either completely minimal, or completely PMC (the latter means that \( \text{graph}(u_1) \) has either mean curvature everywhere equal to \( g(x, u_1) \), or has mean curvature everywhere equal to \( -g(x, u_1) \)). Indeed, the phases of \( u_\infty \) can only be:

- +1 on the supergraph of \( u_1 \) and −1 on the subgraph;
- −1 on the supergraph of \( u_1 \) and +1 on the subgraph;
- +1 on the supergraph of \( u_1 \) and +1 on the subgraph;
- −1 on the supergraph of \( u_1 \) and −1 on the subgraph.

By Theorem 5.1 the third option is not possible for \( g > 0 \). In the first or second case, we have again by Theorem 5.1 the following two facts. The multiplicity of \( V \) is a.e. 1 on the graph and the generalized mean curvature takes almost everywhere the value \( g \), or almost everywhere the value \( -g \), respectively in the first and second case. The first fact implies that the multiplicity is everywhere 1, by the condition that the generalized mean curvature is in \( L^p(\|V\|) \) (so this case can only happen for \( q = 1 \)). Using this fact and elliptic theory we get that the graph is \( C^{2,\alpha} \), with mean curvature \( g \) or \( -g \) respectively. In the fourth case we get by Theorem 5.1 that the mean curvature is almost everywhere 0 and the multiplicity is an even integer almost everywhere. Again, these fact imply that the graph is minimal (and smooth) and carries a constant even multiplicity (therefore this case can only arise for \( q \) even). If \( \tilde{q} = 1 \), the proof of Theorem 4.5 is complete.

\( \tilde{q} = 2 \). This means that \( \text{spt} \|V\| = \bigcup_{j=1}^2 \text{graph}(\tilde{u}_j) \), with \( \tilde{u}_1 \leq \tilde{u}_2 \) and \( \tilde{u}_1 < \tilde{u}_2 \) on some nonempty open set. The set \( \tilde{u}_1 = \tilde{u}_2 \) is the set of points of multiplicity \( q \) and away from it we have multiplicity \( \leq q - 1 \) and (inductively) \( C^2 \) regularity. The
set $\tilde{u}_1 = \tilde{u}_2$ is the set $C$ from Lemma 4.4 so we have that $\tilde{u}_1, \tilde{u}_2$ are $C^2$ on $B \setminus C$ and on each connected component of $B \setminus C$ the graph of $\tilde{u}_j$ is either completely minimal or completely PMC. We distinguish two cases:

(I) graph($\tilde{u}_2$) \ graph($\tilde{u}_1$) has some PMC $C^2$ embedded connected component;  
(II) graph($\tilde{u}_2$) \ graph($\tilde{u}_1$) has no PMC $C^2$ embedded connected component. 

We begin with (II). Then, by [RogTon08], graph($\tilde{u}_2$) \ graph($\tilde{u}_1$) is minimal and has even multiplicity, and moreover $\tilde{u}_\infty = -1$ on the supergraph of $\tilde{u}_2$ and $\tilde{u}_\infty = -1$ on the set $\{(x,y) \in B \times I : \tilde{u}_1(x) < y < \tilde{u}_2(x)\}$. By considering a ball $D$ in an arbitrary connected component of $B \setminus C$, chosen so that a point on the boundary of $D$ belongs to $C$, we use Hopf boundary point lemma in $D$ (see Remark 4.9) to rule out the possibility that $\tilde{u}_j$ is minimal on $D$, or PMC on $D$ with mean curvature $g$. The only possibility is therefore that $\tilde{u}_1$ is PMC on $D$ with mean curvature $-g$, and therefore PMC on $B \setminus C$ with mean curvature $-g$. This also forces, again by [RogTon08], that $u_\infty = +1$ on the subgraph of $\tilde{u}_1$ and multiplicity to be 1 at points in graph($\tilde{u}_1$) \ graph($\tilde{u}_2$). (We are thus in a case that can only occur when $q$ is odd.) 

We next check that the set graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) has vanishing $\mathcal{H}^q$-measure: if that were not the case, then we would contradict [RogTon08], since almost every point on graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) must lie in the interior of set $\{u_\infty = -1\}$, against our earlier conclusions that $u_\infty = +1$ on the subgraph of $\tilde{u}_1$ and $u_\infty = -1$ on the supergraph of $\tilde{u}_2$. (Or, almost every point on graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) must have even multiplicity, against the conclusion that $q$ is odd.) 

We now consider case (I). Let $B_c$ be a connected component of $B \setminus C$ on which the graph of $\tilde{u}_2$ is PMC. Then by considering a ball $D$ in $B_c$ chosen so that a point on the boundary of $D$ belongs to $C$, we use Hopf boundary point lemma (see Remark 4.9) to conclude that the mean curvature of graph($\tilde{u}_2$) on $D$ is $g$ (and not $-g$). We therefore have that each connected components of graph($\tilde{u}_2$) \ graph($\tilde{u}_1$) is either minimal or PMC with mean curvature $g$, and there is at least one PMC component. By [RogTon08] $u_\infty = +1$ on the supergraph of $\tilde{u}_2$ and $u_\infty = -1$ on the set $\{(x,y) \in B \times I : \tilde{u}_1(x) < y < \tilde{u}_2(x)\}$. We check now that the set graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) has vanishing $\mathcal{H}^q$-measure. If that were not the case, then by [RogTon08] we would have that almost every point on graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) must lie in the interior of set $\{u_\infty = -1\}$, against our earlier conclusion. Moreover, by using [RogTon08] in the same way, almost all points in graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) must have multiplicity 1 and mean curvature equal to $g$. We thus conclude that each connected components of graph($\tilde{u}_2$) \ graph($\tilde{u}_1$) is PMC with mean curvature $g$. We now consider any connected component of graph($\tilde{u}_1$) \ graph($\tilde{u}_2$). By using Hopf boundary point lemma as before, we conclude that each such connected component is either minimal or PMC with mean curvature $-g$. If one of them is minimal, then (arguing as above by means of [RogTon08]) $u_\infty = -1$ on the subgraph of $\tilde{u}_1$ and this rules out the fact that any connected component of graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) is PMC. In other words, either all connected components of graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) are minimal, or they are all PMC with mean curvature $-g$. 

In conclusion, for the case $q = 2$ we have two $C^{1,\alpha}$-functions $\tilde{u}_1 \leq \tilde{u}_2$, that coincide on a set $C$ with $\mathcal{H}^q(C) = 0$, are $C^2$ on $B \setminus C$ and exactly one of these possible configurations occurs: 

- graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) is minimal with even multiplicity, graph($\tilde{u}_2$) \ graph($\tilde{u}_1$) has mean curvature $g$ and multiplicity 1; 
- graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) has mean curvature $-g$ and multiplicity 1, graph($\tilde{u}_2$) \ graph($\tilde{u}_1$) has mean curvature $g$; 
- graph($\tilde{u}_1$) \ graph($\tilde{u}_2$) has mean curvature $-g$ and multiplicity 1, graph($\tilde{u}_2$) \ graph($\tilde{u}_1$) is minimal with even multiplicity. 

$q = 3$. This means that $\text{spt} \ ||V|| = \bigcup_{j=1}^3 \text{graph}(\tilde{u}_j)$, with $\tilde{u}_1 \leq \tilde{u}_2 \leq \tilde{u}_3$ and $\tilde{u}_1 < \tilde{u}_2 < \tilde{u}_3$ on some nonempty open set. The set $\tilde{u}_1 = \tilde{u}_2 = \tilde{u}_3$ is the set of points of multiplicity $q$ and away from it we have multiplicity $\leq q - 1$ and (inductively)
respectively PMC with mean curvature $-\mu$ respectively 1, $u$ of bottom graph has mean curvature $-\tilde{\mu}$ which has vanishing $H$.

We claim the following: if $\tilde{u}_1 < \tilde{u}_2 < \tilde{u}_3$ on an open subset of $B$ whose complement has vanishing $H^\alpha$ measure; the graphs of $\tilde{u}_1$, $\tilde{u}_2$, $\tilde{u}_3$ have, on this open set, multiplicity respectively 1, $q - 2$ even, 1; the graphs of $\tilde{u}_1$, $\tilde{u}_2$, $\tilde{u}_3$ are, on $B \setminus C$, respectively PMC with mean curvature $-g$, minimal, PMC with mean curvature $g$. (In particular, this case can only occur for $q$ even and $\geq 4$.)

For the proof of our claim, we consider $B_c$, a connected component of $B \setminus C$ on which $\tilde{u}_1 < \tilde{u}_2 < \tilde{u}_3$ and a ball $D$ in $B_c$ chosen so that a point on the boundary of $D$ belongs to $C$. Using Hopf boundary point lemma in $D$ as done in Remark 4.9 we obtain that, on $B_c$, the top graph has necessarily mean curvature $g$, the bottom graph has mean curvature $-g$, the middle one is minimal. This forces (by [RogTon08]) the fact that $u_\infty = +1$ on the supergraph of $\tilde{u}_3$ and on the subgraph of $\tilde{u}_1$, while $u_\infty = -1$ on $\{ (x,y) \in B \times I : \tilde{u}_3(x) < y < \tilde{u}_1(x) \}$. Always by [RogTon08], the minimal portions have even multiplicity and the PMC portions have multiplicity 1. Additionally, we also obtain that the set $Z$ of points where two of the three graphs agree is a set of vanishing $H^\alpha$-measure (for otherwise, almost everywhere on this set we would have multiplicity higher than 1, hence even because $g > 0$, and by [RogTon08] these points would have to belong to the interior of $\{ u_{\infty} = -1 \}$, against the previous conclusions). The $C^2$ regularity on $Z \setminus C$ completes the proof of the claim.

In the remainder of this section we analyse the possible configurations of $V$ for $\tilde{q} \in \{ 2, 3 \}$. We will show that, for each configuration, each $\tilde{u}_j$ is in $C^2(B)$ and graph $u_j$ is either completely minimal, or completely PMC with mean curvature $g$, or completely PMC with mean curvature $-g$. We consider $\tilde{u}_j$, and $\tilde{u}_{j_1}$ for $j_1 < j_2$ and let $v = \tilde{u}_{j_2} - \tilde{u}_{j_1}$. Then $v$ is a non-negative function and it is $C^2$ on $B \setminus C$. We compute the PDE satisfied by $v$ on $B \setminus C$, arguing as in Lemma 4.4 and keeping in mind the possible configurations. We obtain, for all $\zeta \in C^\infty_c(B \setminus C)$,

$$- \int (a_j D_j v + b_j v) D_j \zeta + \int (d_j D_j v + c v) \zeta = \begin{cases} \int g(x, u_{j_1}) \zeta \\ \int g(x, u_{j_2}) \zeta \\ \int (g(x, u_{j_2}) + g(x, u_{j_1})) \zeta \end{cases}, \quad (21)$$

where $a_j$, $b_j$, $c$, $d_j \in C^{0,\alpha} (B) \cap C^1 (B \setminus C)$ and $a_j$ is symmetric and positive definite. Which right-hand-side appears in (21) depends on the right-hand-side of the PDEs for $\tilde{u}_j$, and $\tilde{u}_{j_1}$ in (15). The first corresponds to the case $\tilde{u}_j$ minimal and $\tilde{u}_{j_1}$ PMC (with mean curvature $g$). The second corresponds to the case $\tilde{u}_j$ minimal and $\tilde{u}_{j_1}$ PMC (with mean curvature $-g$). The third corresponds to the case in which $\tilde{u}_j$ and $\tilde{u}_{j_1}$ are PMC with mean curvature respectively $-g$ and $g$.

We will next prove that the PDE (21) for $v$ extends (in its weak form) to the whole of $B$. Note that $v \in C^{1,\alpha} (B)$ by assumption and $v = 0$ on $C$ ($v \geq 0$ on $B$) and that the right-hand-side (in all three cases) is a $C^{1,\alpha}$ function on $B$.

All possible cases are treated similarly and the third option in (21) corresponds to the situation treated in [BelWic-2] (and in [BelWic-1] when $g \equiv \text{const}$). All cases follow a similar argument, that we carry out here only in the case corresponding to the first option in (21). Recall, from our conclusions on the possible configurations when $\tilde{q} \in \{ 2, 3 \}$, that $H^\alpha (C) = 0$.

We note, first of all, that inequality 6 provides an $L^2$-bound on $|D^2 \tilde{u}_j|$ on $B \setminus C$, because the generalized second fundamental form appearing in (6) agrees with the usual second fundamental form on gen-reg $V$, and thus controls the second derivatives of $\tilde{u}_j$ on $B \setminus C$. In other words $\| \tilde{u}_j \|_{W^{2,2} (B \setminus C)}$ is finite for each choice of $j$. This also implies that $\| v \|_{W^{2,2} (B \setminus C)}$ is finite. Moreover, $Dv \in C^{0,\alpha} (B)$ and $Dv = 0$ on $C$, with $Dv \in C^1 (B \setminus C)$. Using [BelWic-1] Lemma C.1 we obtain
that $Dv \in W^{1,2}_{\text{loc}}(B)$. We then adapt the argument in [BelWic-1 Section 7.5]. We consider the function $f = a_j D_i v + b_i v$ on $B$: $f \in C^{0,\alpha}(B) \cap C^1(B \setminus C)$, with $f = 0$ on $C$ (since $v$ and $Dv$ vanish there). Recalling the structure of $a_j$ and $b_i$, we find that $Df$ is in $L^1(B \setminus C)$, thanks to the $W^{2,\infty}$ estimate on $v$, in particular $\int_{B \setminus C} |Df| < \infty$. Applying again [BelWic-1 Lemma C.1] we conclude that $f \in W^{1,1}(B)$, with distributional derivative given by the $L^1$ function equal to $Df$ on $B \setminus C$ and 0 on $C$. Using this fact, and recalling that the PDE for $v$ is satisfied strongly on $B \setminus C$, we compute, for $\zeta \in C_0^\infty(B)$:

$$-\int_B (a_j D_i v + b_i v) D_j \zeta + (d_i D_i v + cv) \zeta = \int_B \underbrace{D_j (a_j D_i v + b_i v)}_{\text{extend from } B \setminus C} \zeta + (d_i D_i v + cv) \zeta$$

$$= \int_{B \setminus C} D_j (a_j D_i v + b_i v) \zeta + (d_i D_i v + cv) \zeta = \int_{B \setminus C} g(x, \tilde{u}_j) \zeta = \int_B g(x, \tilde{u}_j) \zeta,$$

where we used (last equality) the fact that $H^n(C) = 0$. Equality (22) says that the weak PDE for $v$ is valid on the whole of $B$.

With the knowledge that the (weak) PDEs for $\tilde{u}_j = \tilde{u}_j$ extend from $B \setminus C$ to $B$, we are now ready to prove that $\tilde{u}_j \in C^2(B)$ for $j \in \{1, 2, 3\}$. We will do so for the case $\tilde{q} = 3$ (thus with $q$ even); the case $\tilde{q} = 2$ can be treated analogously (and is more straightforward). We point out that the functions $F_j$ in (18) for $j \in \{1, \ldots, n\}$ are odd in the third variable: this follows by recalling that $F(x, \nu)$ is even in $\nu$ (since $F$ is the integrand of the area functional). The first variation formula for $V$ gives (recalling that $\|V\| = |\text{graph}(u_1)| + (q - 2)|\text{graph}(u_2)| + |\text{graph}(u_3)|$ with $q$ even, in the case $\tilde{q} = 3$), for any $\zeta \in C_0^\infty(B)$

$$\sum_{j=1}^3 \tilde{q}_j \left( -\int_B F_j(x, \tilde{u}_j, D\tilde{u}_j) D_j \zeta + \int F_{n+1}(x, \tilde{u}_j, D\tilde{u}_j) \zeta \right) = \int_B (g(x, \tilde{u}_3) - g(x, \tilde{u}_1)) \zeta,$$

with $\tilde{q}_1 = \tilde{q}_3 = 1, \tilde{q}_2 = q - 2$ with $q$ even. We rewrite (23) as follows

$$q \left[ \int_B -F_j(x, \tilde{u}_1, D\tilde{u}_1) D_j \zeta + F_{n+1}(x, \tilde{u}_1, D\tilde{u}_1) \zeta \right] + (q - 2) \left[ \int_B -F_j(x, \tilde{u}_2, D\tilde{u}_2) D_j \zeta + F_{n+1}(x, \tilde{u}_2, D\tilde{u}_2) \zeta \right] - (q - 2) \left[ \int_B -F_j(x, \tilde{u}_3, D\tilde{u}_3) D_j \zeta + F_{n+1}(x, \tilde{u}_3, D\tilde{u}_3) \zeta \right]$$

$$- \left[ \int_B -F_j(x, \tilde{u}_1, D\tilde{u}_1) D_j \zeta + F_{n+1}(x, \tilde{u}_1, D\tilde{u}_1) \zeta \right]$$

$$= \int_B (g(x, \tilde{u}_3) - g(x, \tilde{u}_1)) \zeta.$$
above (see (22)) the validity of these PDEs on $B$, so we obtain
\[
q \left[ \int_B - F_j(x, \tilde{u}_1, D\tilde{u}_1) D_j \zeta + F_{n+1}(x, \tilde{u}_1, D\tilde{u}_1) \zeta \right] \\
+ (q - 2) \int_B g(x, \tilde{u}_1) \zeta + \int_B (g(x, \tilde{u}_3) + g(x, \tilde{u}_1)) \zeta = \int_B (g(x, \tilde{u}_3) - g(x, \tilde{u}_1)) \zeta,
\]
which implies that (for every $\zeta \in C^\infty_c(B)$)
\[
\int_B - F_j(x, \tilde{u}_1, D\tilde{u}_1) D_j \zeta + F_{n+1}(x, \tilde{u}_1, D\tilde{u}_1) \zeta = - \int_B g(x, \tilde{u}_1) \zeta. \tag{24}
\]
Similarly we can prove (using the PDEs for $v_{32} = \tilde{u}_3 - \tilde{u}_2$ and $v_{31} = \tilde{u}_3 - \tilde{u}_1$) that for every $\zeta \in C^\infty_c(B)$
\[
\int_B - F_j(x, \tilde{u}_3, D\tilde{u}_3) D_j \zeta + F_{n+1}(x, \tilde{u}_3, D\tilde{u}_3) \zeta = \int_B g(x, \tilde{u}_3) \zeta. \tag{25}
\]
By standard elliptic theory the fulfilment (in weak sense) of the PDEs (24) and (25) implies that $\tilde{u}_1$ and $\tilde{u}_3$ are $C^2$ and solve these PDEs strongly. These facts, together with (23), also imply that $\tilde{u}_2$ is $C^2$ and solves
\[
D_j (F_j(x, \tilde{u}_2, D\tilde{u}_2)) + F_{n+1}(x, \tilde{u}_2, D\tilde{u}_2) = 0
\]
on $B$. This completes the proof of Theorem 4.5.

4.5 Regularity of stable limit $(g,0)$-varifolds, Part II: proofs of Theorem 4.2 and Theorem 4.3

Proof of Theorem 4.2. In view of (3), we may in the first instance argue as in the proof of ([TonWic12, Proposition 3.2]) to show that there exists a set $\Sigma \subset \text{spt} \|V\| \cap B_{\rho_0}(0)$ with dim$_H(\Sigma) \leq n - 2$ such that no tangent cone to $V$ at any point $y \in \text{spt} \|V\| \setminus \Sigma$ is supported on a union of three or more half-hyperplanes meeting along an $(n-1)$-dimensional subspace. Note that this argument is valid without the positivity assumption on $g$, i.e. for any $g \in C^{1,1}(\mathbb{N})$.

It follows from this in fact that no tangent cone to $V$ at any point $y \in \text{spt} \|V\| \cap B_{\rho_0}(0)$ can be supported on a union of three or more half-hyperplanes meeting along an $(n-1)$-dimensional subspace. To see this, suppose $C$ is a tangent cone of this type at some point $y$, so that $\eta_{y,\sigma_j} \# V \to C$ as varifolds for some sequence of positive numbers $\sigma_j \to 0$. Passing to a subsequence of $\{\varepsilon_j\}$ without relabelling, we may assume that $\tilde{\varepsilon}_j = \sigma_j^{-1} \varepsilon_j \to 0$. Letting $\tilde{u}_{\tilde{\varepsilon}_j}(X) = u_{\varepsilon_j}(Y + \sigma_j X)$, we see by the reasoning as in ([TonWic12, p. 200]) with obvious modifications, that $C$ is the limit varifold associated with the sequence $(\tilde{u}_{\tilde{\varepsilon}_j})$ in the same way that the limit varifold $V$ is associated with the sequence $(u_{\varepsilon_j})$. However, we have just shown that any such limit varifold has the property that the set of points where one of its tangent cones is supported on a union of three or more half-hyperplanes meeting along a common axis has Hausdorff dimension $\leq n - 2$, a property clearly violated by $C$. Thus no tangent cone to $V$ can be supported on three or more half-hyperplanes meeting along a common axis, and this is conclusion (i).

We next prove conclusion (ii), by arguing by induction on $q$. If $q = 1$ the conclusion holds by Allard’s regularity theorem. Suppose $q \geq 2$ and that the following induction hypothesis holds:

(*) For any stable limit $(g,0)$-varifold $V$, conclusion (ii) holds with $q'$ in place of $q$ for any $q' \in \{1, 2, \ldots, q - 1\}$. 

33
Let $\mathcal{V}$ be the set of all stable limit $(g,0)$-varifolds on $N$ for some fixed positive $g \in C^{1,1}(N)$ (and varying double-well potential functions associated with sequences of functionals $F_{\epsilon}^{g,\sigma}$ corresponding to stable limit $(g,0)$-varifolds. We first wish to apply Theorem 1.4 to $\mathcal{V}$. In view of Theorem 3.1 and conclusion (i), it is clear that $\mathcal{V}$ satisfies hypotheses (a), (b) of Theorem 1.4 so we only need to verify hypothesis (c) of Theorem 1.4 for some $\beta > 0$ (depending on $g$). Specifically, we need to show that there is a constant $\beta > 0$ such that, if $V \in \mathcal{V}$, $X \in N$, $\tilde{V} = (\Gamma \circ \exp_X)^* V \wedge \mathcal{N}_{\min,X}(X)$ for some orthogonal rotation $\Gamma : T_X N \cong \mathbb{R}^{n+1} \to T_X N$, $\sigma \in (0, \min\{1, \text{inj}_X N\}]$, $$(\omega, 2^n)^{-1} \|\eta_{0,\sigma} \# \tilde{V}\|(B_1^{n+1}(0)) \leq q + 1/2,$$ 

$$q - 1/2 \leq \omega_n^{-1} \|\eta_{0,\sigma} \# \tilde{V}\|((B_1^{n/2}(0) \times \mathbb{R}) \cap B_1^{n+1}(0)) \leq q + 1/2,$$ 

$$\Theta(\|\eta_{0,\sigma} \# \tilde{V}\|, Y) < q \text{ for each } Y \in B_1^{n+1}(0) \text{ and if}$$ 

$$\text{if } \int_{(B_1^{n/2}(0) \times \mathbb{R}) \cap B_1^{n+1}(0)} |x|^{n+1} |d\|\eta_{0,\sigma} \# \tilde{V}\|$$ 

$$+ \sigma \left( \int_{(B_1^{n/2}(0) \times \mathbb{R}) \cap B_1^{n+1}(0)} |H_V(\exp_X(\Gamma^{-1}(\sigma Y)))|^p d\|\eta_{0,\sigma} \# \tilde{V}\|(Y) \right)^{1/p} + \sigma < \beta,$$ 

then 

$$\eta_{0,\sigma} \# \tilde{V} \wedge ((B_1^{n/2}(0) \times \mathbb{R}) \cap B_1^{n+1}(0)) = \sum_{j=1}^q |\text{graph } u_j|$$ 

(26)

for some $u_j \in C^2(B_1^{n/2}(0))$, $j = 1, 2, \ldots, q$, with $u_1 \leq u_2 \leq \ldots \leq u_q$. To do this, note first that locally near any point $Z \in (\text{spt } \|V\| \setminus \text{sing } V) \cap \mathcal{N}_{\sigma}(X)$, we can write $\text{spt } \|V\|$ as the union of a (possibly empty) $C^2$ embedded minimal hypersurface $M_1$ and a (possibly empty) $C^2$ quasi-embedded oriented hypersurface $M_2$ with mean curvature $\nu$ for some choice of continuous unit normal $\nu$ on $M_2$. Letting $\hat{M}_i = \exp_X M_i$ and writing $M_i$ near $\tilde{Z} = \exp_X(Z)$ either as a single $C^2$ graph $G_1^{(0)}$ (for $i = 1$), or as a single $C^2$ graph $G_2^{(0)}$ or the union of two $C^2$ graphs $G_1^{(1)}, G_2^{(1)}$ (for $i = 2$) over the common tangent space to $\hat{M}_i$ at $\tilde{Z}$, and considering the separate Euler–Lagrange equations satisfied by the functions defining these graphs, we see (cf. proof of [SchSim51] inequality (1.16)) that for each $G_i^{(j)}$, $i \in \{1, 2\}$, $j \in \{0, 1, 2\}$, 

$$|H_{G_i^{(j)}(\tilde{Z})}| \leq C(\|\tilde{Z}\| |A_{G_i^{(j)}(\tilde{Z})}| + 1)$$

where $C = C(g, N)$ and $H_{G_i^{(j)}}$, $A_{G_i^{(j)}}$ denote the mean curvature and the second fundamental form of $G_i^{(j)}$ respectively. In view of this and Lemma 4.3 above, the assertion [26] will follow from [BelWic2 Theorem 5.1] provided we take $\beta = \varepsilon$ where $\varepsilon = \varepsilon(n, p, q, C)$ is as in [BelWic2 Theorem 5.1] and provided we can show, under the above hypotheses on $W$, that 

$$\text{dim}_H(\text{sing } V \wedge \mathcal{N}_{\sigma}(X)) \leq n - 7.$$ 

To verify this dimension bound, note first that since $\Theta(\|V\|, Y) < q$ for each $Y \in \mathcal{N}_{\sigma}(X)$, it follows from the induction hypothesis (s) that if $Y \in \text{spt } \|V\| \cap \mathcal{N}_{\sigma}(X)$ is such that a tangent cone to $V$ at $Y$ is supported on a hyperplane then $Y \not\in \text{sing } V$. In view of this and conclusion (i), we see that if $C$ is a tangent cone to $V$ at a point $Y \in \text{sing } V \cap \mathcal{N}_{\sigma}(X)$, and if $d$ is the dimension of the largest subspace along which $C$ is translation invariant, then $d \leq n - 2$. Now consider
repeatedly taking tangent cones to tangent cones at singular points away from the origin. Note that the density at the origin of any such cone is \( < \Theta(||C||,0) = \Theta(||V||,Y) < q \). So in view of the induction hypothesis (\(*)\) and the fact that \( C \) and all successive tangent cones in this process, being limit varifolds associated with sequences of stable critical points of \( F_{\varepsilon_j,g_j} \) for some sequences \( \varepsilon_j \to 0 \) and \( g_j \to 0 \), are stationary cones with stable regular (embedded) parts. Hence this process produces a cone that after a rotation can be written as \( C_0 \times \mathbb{R}^k \) where \( C_0 \) is an \( n-k \) dimensional stable cone in \( \mathbb{R}^{n-k+1} \) with sing \( C_0 = \{0\} \). By conclusion (i) we have that \( k \leq n-2 \), so it follows from Simons’ theorem (\[Sim68\]) applied to \( C_0 \) that \( k \leq n-7 \). Since \( k \geq d \), we have that \( d \leq n-7 \). Thus we have shown that the dimension of the largest subspace along which any tangent cone to \( W \) is translation invariant is \( \leq n-7 \). It follows from Almgren’s generalised stratification theorem that \( \dim_{\mu} (\text{sing} \bigcup N_{\varepsilon}(X)) \leq n-7 \). This concludes the verification of hypothesis (c) of Theorem 4.4 for the class \( V \).

Now if \( Y \in \text{spt} ||V|| \) and if \( q|P| \in \text{VarTan}_Y \) for some \( n \)-dimensional subspace \( P \), then we may apply the conclusion of Theorem 4.4 to see that in an appropriately small neighborhood around the origin, the varifold \( \exp_{V}^{-1} \) is supported on three or more half-hyperplanes meeting along a common \((n-1)\)-dimensional subspace.

Proof of Theorem 4.5. Let \( G \) be the set of \( g \in C^{1,1}(N) \) such that \( \sup_N |g| \leq 1 \). Let \( V \) be the collection of all stable limit \((g,0)\)-varifolds on \( N \) over all \( g \in G \) and all double-well potentials \( W : \mathbb{R} \to [0,\infty) \) of class \( C^2 \) having precisely two non-degenerate minima at \( \pm 1 \) with value 0. By Theorem 4.2, part (i), no tangent cone to a varifold in \( V \) is supported on three or more half-hyperplanes meeting along a common \((n-1)\)-dimensional subspace.

We first prove part (ii). For this note first that there is a constant \( \mu = \mu(N) \) such that if \( y \in N \), \( \rho \in (0,\text{inj}_y N) \), \( \mu \rho < 1 \) and if \( \bar{V} \) is any varifold on \( B^n_{\rho}(0) \subset T_y N \approx \mathbb{R}^{n+1} \) such that \( \bar{V} = \exp^{-1}_{y}(V \bigcup N_{\rho}(y)) \) for some \( V \in \mathcal{V} \), then the generalised mean curvature of \( \bar{V} \) (with respect to the Euclidean metric on \( B^n_{\rho}(0) \)) satisfies \( |H_{\bar{V}}| \leq C \), where \( C = C(\Gamma) \). Fix \( y \in N \) and \( \rho \in (0,\text{inj}_y N) \) such that \( \rho \mu \rho < 1 \), and let \( \bar{V} \) be the set of all varifolds \( \bar{V} \) on \( B^n_{\rho}(0) \) such that \( \bar{V} = \exp^{-1}_{y}(V \bigcup N_{\rho}(y)) \) for some \( V \in \mathcal{V} \), some \( z \in B^n_{\rho}(0) \) and some \( \sigma \in (0,1] \). Then no tangent cone to a varifold in \( \mathcal{V} \) is supported on three or more half-hyperplanes meeting along a common \((n-1)\)-dimensional subspace, since this is so for varifolds in \( V \). By the Allard regularity theorem and the \( C^{2,\alpha} \) Schauder theory for uniformly elliptic equations, we have the validity of the assertion in part (ii) in the case \( q = 1 \). Let \( q \geq 2 \), and assume by induction that part (ii) is valid with \( q' \) in place of \( q \) for any \( q' \in \{1,2,\ldots,q-1\} \). By arguing as in the proof of Theorem 4.2, part (i), using the present induction hypothesis in place of the induction hypothesis (\(*)\) therein, we see (by applying Theorem 4.4) that the conclusion of part (ii) of the current theorem holds, in the first instance, with \( C^{1,\alpha} \) functions \( u_j \), \( j = 1,2,\ldots,q \). By Theorem 4.5, we see that these \( u_j \) are in fact of class \( C^{2,\alpha} \). Finally, the desired \( C^{2,\alpha} \) estimate in part (ii) follows from standard Schauder estimates.

For part (i), note that by Theorem 4.2 we have that any varifold in \( \bar{V} \) is quasi-embedded away from a closed set of Hausdorff dimension at most \( n-7 \). The assertion in part (i) now follows arguing as in \[SchSim81\] pp. 785–787] (the first part of the proof of \[SchSim81\, Theorem 2]).
5 A min-max construction of Allen–Cahn solutions

For notational convenience, we will work, in Sections 5 and 6, with the functional \( F_\varepsilon = F_{\varepsilon,g} \). This means that, given \( g > 0 \) in \( C^{1,1}(N) \), we will establish the existence of a quasi-embedded immersed hypersurface with mean curvature \( \frac{g}{\varepsilon} \).

This is of course equivalent to proving Theorem 1.1 for \( g > 0 \), \( g \in C^{1,1}(N) \).

In this section we will choose two functions \( a_\varepsilon \) and \( b_\varepsilon \) as valley points for the functional \( F_\varepsilon \) and prove that the class of continuous paths in \( W^{1,2}(N) \) that joins \( a_\varepsilon \) to \( b_\varepsilon \) satisfies a suitable “wall condition” to produce minmax critical points. The functions \( a_\varepsilon \) and \( b_\varepsilon \) converge uniformly on \( N \) as \( \varepsilon \to 0 \) respectively to the constants \(-1\) and \(+1\). The value of the functional at \(-1\) and \(+1\) is respectively \( F_\varepsilon(-1) = \int_N g \) and \( F_\varepsilon(+1) = -\int_N g \). We will consider in the next lemma affine subspaces of the form

\[
\Pi_\delta = \left\{ u \in W^{1,2}(N) : \int_N g u = -\int_N g + \delta \right\}.
\]

In Proposition 5.1 we will see that an affine subspace of this type provides a suitable “wall”.

**Lemma 5.1.** There exists \( \delta \in (0, 2\int_N g) \) such that the following is true. For any \( \varepsilon_j \to 0 \) there exists \( \delta_j \to \delta \) such that

\[
\liminf_{j \to \infty} \left( \inf_{u \in \Pi_{\delta_j}} F_{\varepsilon_j}(u) \right) > \int_N g \left( = F_{\varepsilon_0}(-1) \geq F_{\varepsilon_0}(+1) \right).
\]

**Proof.** Within this proof we will write \( \mathbb{I}_A \) for the characteristic function of \( A \), and \( |A| \) for \( \mathcal{H}^{n+1}(A) \). Pick a point \( z \) where \( g \) achieves its maximum \( g_M \) and consider a geodesic ball \( B \) centred at \( z \) and such that \( g > g_M/2 \) on \( B \). Then \( \mathbb{I}_B = \mathbb{I}_{N \setminus B} \). We set \( \delta_B = 2\int_B g \), therefore \( \int_B g \to -\int_B g + \delta_B \) and \( |B|g_M \leq \delta_B \leq 2|B|g_M \). We will prove that for a sufficiently small choice of \( B \) (if \( g > g_M/2 \) is true on a ball centred at \( z \), it is true on any ball centred at \( z \) and contained in the first), the Lemma holds with \( \delta = \delta_B \) and with \( \delta_j = \int_N g \varepsilon_j + \int_N g \) (for an arbitrary given sequence \( \varepsilon_j \to 0 \)). The choice of \( \delta_j \) is made to ensure \( \varepsilon_j \in \Pi_{\delta_j} \), and \( \delta_j \to \delta \).

Note that, for each \( \varepsilon_j \),

\[
\inf_{u \in \Pi_{\delta_j}} F_{\varepsilon_j}(u) = \left( \inf_{u \in \Pi_{\delta_j}} \mathcal{E}_{\varepsilon_j}(u) \right) - \delta_j + \int_N g
\]

and the infimizing sequences are the same for \( \inf_{u \in \Pi_{\delta_j}} F_{\varepsilon_j}(u) \) and for \( \inf_{u \in \Pi_{\delta_j}} \mathcal{E}_{\varepsilon_j}(u) \) (because the two functionals differ by the constant \(-\delta_j + \int_N g \) on \( \Pi_{\delta_j} \)). Since \( \mathcal{E}_{\varepsilon_j}(v) \to 2\sigma\mathcal{H}^n(\partial B) \) as \( \varepsilon \to 0 \) and \( v_{\varepsilon_j} \in \Pi_{\delta_j} \) by construction, we can see that there exists an upper bound for \( \inf_{u \in \Pi_{\delta_j}} F_{\varepsilon_j}(u) \) (and for \( \inf_{u \in \Pi_{\delta_j}} \mathcal{E}_{\varepsilon_j}(u) \)) that is independent of \( j \). Since \( \mathcal{E}_{\varepsilon} \geq 0 \), we also have a lower bound for \( \inf_{u \in \Pi_{\delta_j}} F_{\varepsilon_j}(u) \) independently of \( j \).

Pick \( u_j \) such that \( F_{\varepsilon_j}(u_j) - \inf_{u \in \Pi_{\delta_j}} F_{\varepsilon_j}(u) \) converges to 0 (as \( j \to \infty \)). Then \( F_{\varepsilon_j}(u_j) \) is uniformly bounded above and therefore so is \( \mathcal{E}_{\varepsilon_j}(u_j) \). The latter condition guarantees, thanks to a standard argument ([HutToni00], [MosMor77]) that we now recall, that \( u_j \) converges in \( L^1 \) to a \( BV \) function \( u_\infty \) that takes only the values \( \pm 1 \).
The uniform bound on $\mathcal{E}_{\varepsilon_j}$ and the fact that $W(x) \geq \kappa(x-1)^2$ for $x \geq 2$ imply that $\|u\|_{L^2}$ is uniformly bounded. Set $\Phi(s) = \int_0^s \sqrt{\frac{W(s)}{2} ds}$ and $\sigma = \int_1^s \frac{W(s)}{2} ds$ (so that $\Phi(\pm 1) = \pm \sigma/2$) and let $w_j = \Phi(u_j)$. Then $\|w_j\|_{BV}$ is uniformly bounded (with $\int_N |\nabla w_j| \leq \frac{\mathcal{E}_{\varepsilon_j}(u_k)}{2}$) and therefore (upon passing to a subsequence that we do not relabel) there exists $w_\infty \in BV(N)$ such that $w_j \rightharpoonup w_\infty$. In particular, $w_j \to w_\infty$ and $w_j \to w_\infty$ a.e., so that $u_j \to u_\infty := \Phi^{-1}(w_\infty)$ a.e. The uniform bound on $\mathcal{E}_{\varepsilon_j}$ and Fatou’s lemma imply that $\int_N W(u_\infty) \leq \liminf_{j \to \infty} \int_N W(u_j) = 0$, from which it follows that $w_\infty = \pm 1$ a.e. Then $\frac{2}{\varepsilon} w_\infty = w_\infty \in BV(N)$.

The convergence $w_j \rightharpoonup w_\infty$ implies in particular that $w_j \to w_\infty$ in measure. Since $\Phi^{-1}$ is uniformly continuous, for any $s > 0$ we can choose $\varepsilon > 0$ such that $|w_j - w_\infty| \leq \varepsilon$ implies $|u_j - u_\infty| \leq s$. Therefore for every $s > 0$ we have $|\{ |u_j - u_\infty| \geq s \}| \to 0$, i.e. $u_j \to u_\infty$ in measure. We can then prove $u_j \rightharpoonup u_\infty$ as follows.

$$\int_N |u_j - u_\infty| = \int_{|u_j - u_\infty| \geq s} |u_j - u_\infty| + \int_{|u_j - u_\infty| < s} |u_j - u_\infty| \leq \int_{|u_j - u_\infty| \geq s} |u_j| + \int_{|u_j - u_\infty| < s} |u_\infty| + \int_N s \leq |\{ |u_j - u_\infty| \geq s \}|^{1/2} \left( \int_N |u_j|^2 \right)^{1/2} + |\{ |u_j - u_\infty| \geq s \}| + s|N|$$

and all three terms go to 0 as $s \to 0$.

Since $u_j \in \Pi_{\varepsilon_j}$, we conclude that $\int_N g u_\infty = -\int_N g + \delta$. As $u_\infty = \pm 1$ a.e. and is $BV$, we must have that there exists a set $D \subset N$ with finite perimeter such that $\mathbb{I}_D - \mathbb{I}_{N \setminus D} = u_\infty$ and $\int_D g = \int_D g$. This implies that $|D| \geq \frac{1}{2}|B|$ (by the choice of $B$). Denoting by $\overline{w_j}$ the average of $w_j$ and recalling the Sobolev–Poincaré inequality we have:

$$\frac{\mathcal{E}_{\varepsilon_j}(u_j)}{2} \geq \int_N |\nabla w_j| = \int_N |\nabla (w_j - \overline{w_j})| \geq C_{SP} \left( \int_N |w_j - \overline{w_j}|^{\frac{n+1}{n}} \right)^\frac{n}{n+1}.$$

By Fatou’s lemma (and the $L^1$ and a.e. convergence $w_j \to w_\infty = \mathbb{I}_D - \mathbb{I}_{N \setminus D}$, which also gives $\overline{w_j} \to \frac{1}{|N|}(2|D| - |N|)$), we get

$$\left( \int_N \frac{|w_j - \overline{w_j}|^{\frac{n+1}{n}}}{|N|} \right)^\frac{n}{n+1} \leq \liminf_{j \to \infty} \left( \int_N |w_j - \overline{w_j}|^{\frac{n+1}{n}} \right)^\frac{n}{n+1}.$$

Computing the left-hand-side (giving up the integral over $N \setminus D$) we get

$$\left( 2 - \frac{2|D|}{|N|} \right) \frac{|D|}{|N|} \leq \frac{1}{2C_{SP}} \liminf_{j \to \infty} \mathcal{E}_{\varepsilon_j}(u_j). \quad (27)$$

Assume that the conclusion of the lemma fails for a certain $\delta$ (i.e. for a certain choice of $B$). In other words, we assume that for some $\varepsilon_j \to 0$ we have (for the choice of $\delta_j$ specified above) $\liminf_{j \to \infty} \left( \inf_{u \in \Pi_{\varepsilon_j}} \mathcal{F}_{\varepsilon_j}(u) \right) \leq \int_N g$. This means that there exists a subsequence (not relabeled) $\varepsilon_j \to 0$ such that (for $u_j$ chosen above) $\lim_{j \to \infty} \mathcal{F}_{\varepsilon_j}(u_j) \leq \int_N g$. Then

$$\int_N g \geq \liminf_{j \to \infty} \mathcal{F}_{\varepsilon_j}(u_j) = \liminf_{j \to \infty} \mathcal{E}_{\varepsilon_j}(u_j) + \int_N g - \delta,$$
bounded in $\{\mbox{sequences}, i.e. that for any sequence $(i)\quad \mbox{We show that the Palais–Smale condition is satisfied (at fixed}

\begin{align*}
\liminf_{j \to \infty} E_{\varepsilon_j}(u_j) \leq \delta.
\end{align*}

Recalling $|D| \geq \frac{1}{2}|B|$ and $g_M |B| \leq \delta < 2g_M |B|$, \cite{27} and \cite{28} give that $\frac{\delta}{\varepsilon^{\frac{1}{2}}} \lesssim \delta$, a contradiction if $B$ is chosen small enough (to make $\delta$ suitably small).

Choice of valley points $a_\varepsilon$ and $b_\varepsilon$. There exist two functions $a_\varepsilon$ and $b_\varepsilon$ on $N$ that solve $F'_\varepsilon = 0$ with $-1 < a_\varepsilon < -1 + \varepsilon$ and $b_\varepsilon > +1$ and $a_\varepsilon \to -1$ and $b_\varepsilon \to +1$ uniformly on $\varepsilon \to 0$, where $a > 0$ depends on $W$ and on the maximum of $g$. To see this, consider the constant $-1$ and evaluate $-F'_\varepsilon(-1) = \Delta(-1) - \frac{W'(-1)}{\varepsilon} + g = g > 0$. For the constant $(-1 + \varepsilon)$ on the other hand we have $-F'\varepsilon(-1 + \varepsilon) = -\frac{W'(-1 + \varepsilon)}{\varepsilon} + g$. Recall that $W'(-1 + \varepsilon) \approx cC_W \varepsilon$ ($W$ is quadratic around $-1$); choosing $c$ sufficiently large (depending only on $W$ and $g$) we can ensure that $g < cC_W$, and therefore that $-F'\varepsilon(-1 + \varepsilon) < 0$. Therefore, by considering the negative gradient flow of $F_\varepsilon$ with initial condition given by the constant $-1$, we obtain a (stable) solution $a_\varepsilon$ to $F'_\varepsilon = 0$ that lies between $-1$ and $-1 + \varepsilon$ (the latter acts as an upper barrier by the maximum principle). Similarly, computing $-F'_\varepsilon(+1) = g > 0$ and $-F'_\varepsilon(1 + \varepsilon) = -\frac{W'(1 + \varepsilon)}{\varepsilon} + g < 0$, we obtain that there is a (stable) solution $b_\varepsilon$ to $F'_\varepsilon = 0$ that lies between $1$ and $1 + \varepsilon$ and we can obtain $b_\varepsilon$ via negative gradient flow of $F_\varepsilon$ with initial condition given by the constant $+1$. We will use the functions $a_\varepsilon$ and $b_\varepsilon$ as valley points for the class of admissible paths.

Proposition 5.1 (Existence of a mountain pass solution). For $\varepsilon > 0$ let $\Gamma$ denote the collection of all continuous paths $\gamma : [-1, 1] \to W^{1,2}(N)$ such that $\gamma(-1) = a_\varepsilon$ and $\gamma(1) = b_\varepsilon$. Then there exists $\varepsilon_0 > 0$ such that for each $\varepsilon < \varepsilon_0$

$$
\inf_{\gamma \in \Gamma} \sup_{u \in \gamma([-1,1])} F_\varepsilon(u) = \beta_\varepsilon
$$

is a critical value, i.e. there exists $u_\varepsilon \in W^{1,2}(N)$ that is a critical point of $F_\varepsilon$ and has Morse index $\leq 1$.

Proof. (i) We show that $\beta_\varepsilon > F_\varepsilon(a_\varepsilon)$ and $\beta_\varepsilon > F_\varepsilon(b_\varepsilon)$. By the choice of $a_\varepsilon$ and $b_\varepsilon$, we have $F_\varepsilon(a_\varepsilon) < F_\varepsilon(-1) = \int_N g$ and $F_\varepsilon(b_\varepsilon) < F_\varepsilon(+1) = -\int_N g$, so it will suffice to prove that for all sufficiently small $\varepsilon$ we have $\beta_\varepsilon > \int_N g$. Observe that $\int_N \alpha g < -\int_N g + \varepsilon \|\alpha\|_{L^\infty}$. The choice of $\delta \in (0, \int_N g)$ in Lemma \cite{5.3} can be made independently of $\varepsilon$, in particular $\delta > c\varepsilon\|\alpha\|_{L^\infty}$ for $\varepsilon$ sufficiently small. Then the continuity of $u \to \int_N u$ in $W^{1,2}(N)$ guarantees that any continuous path joining $a_\varepsilon$ to $b_\varepsilon$ must cross $\Pi_{\delta_j}$ (for $j$ large enough), thus ensuring the mountain pass condition.

(ii) We show that the Palais–Smale condition is satisfied (at fixed $\varepsilon$) on $F_\varepsilon$-bounded sequences, i.e. that for any sequence $\{u_m\}_{m=1}^\infty$ such that $F_\varepsilon(u_m) \to 0$ (as elements of the dual of $W^{1,2}(N)$) there exists a subsequence of $u_m$ converging strongly in $W^{1,2}(N)$. Note that for $|u| \geq 2$ we have $W(u) - gu \geq k\varepsilon u^2 - \|g\|_{L^\infty} |u|$ for some $k > 0$ and thus for $|u| \geq C_{g,W}$ we have $W(u) - gu \geq \frac{\varepsilon}{2}u^2$. Therefore the assumption $F_\varepsilon(u_m) \leq K$ implies that $\|\nabla u_m\|_{L^2(N)} \leq \frac{K}{\varepsilon}$ and that $\|u_m\|_{L^2(N)} \leq C_{g,W,K}$. Rellich-Kondrachov theorem provides a subsequence (not relabeled) that converges weakly in $W^{1,2}(N)$ to a function $u$. Recall that $F'_\varepsilon(u_m) = \int_N \nabla u_m \nabla \psi + \frac{W'\varepsilon(u_m)}{\varepsilon} \psi - g \psi$. By the $L^2$-convergence of $u_m$ to $u$ and the fact that $W'$ is linear at $\pm \infty$ we obtain that $W'(u_m) \to W'(u)$ in $L^2$. Using the weak convergence $u_m \overset{W^{1,2}}{\to} u$ we get $\lim_{m \to \infty} F'_\varepsilon(u_m) = F'_\varepsilon(u)$. On the other hand, the assumption on $F'_\varepsilon(u_m)$ gives $F'_\varepsilon(u) = 0$, i.e. $u$ is a critical point of $F_\varepsilon$. The boundedness of $u_m - u$ in $W^{1,2}$ gives that $F'_\varepsilon(u)(u_m - u) \to 0$ and therefore $\frac{F'_\varepsilon(u_m)(u_m - u) - F'_\varepsilon(u)(u_m - u)}{\varepsilon} \to 0$.
$u) = \int_N \varepsilon |\nabla (u_m - u)|^2 + \int_N \varepsilon \frac{1}{2} (W'(u_m) - W'(u))(u_m - u) - \int_N g(u_m - u) \to 0$. The second and third integrals go to 0 by the strong $L^2$-convergence $u_m \to u$, therefore $\int_N |\nabla (u_m - u)|^2 \to 0$, concluding that $u_m \to u$ in $W^{1,2}$. (strongly). (iii) The proposition now follows from standard minmax theory since the class of continuous paths $\gamma : [-1, 1] \to W^{1,2}(\Omega)$ such that $\gamma(-1) = a_\varepsilon$ and $\gamma(1) = b_\varepsilon$ is invariant under the flow induced by the negative gradient of $F_\varepsilon$ (see e.g. [Str]).

What is left to do is to make sure that the energy $E_\varepsilon(u_\varepsilon)$ associated to the mountain pass solution provided by Proposition 5.1 stays uniformly (in $\varepsilon$) bounded above and away from zero. This will guarantee that, as $\varepsilon \to 0$, the energy distribution of $u_\varepsilon$ gives rise to a non-trivial varifold (with finite mass). In Lemma 5.2 we discuss the upper bound. The lower bound (Lemma 5.3 will be immediate from Lemma 5.1.

**Lemma 5.2.** There exist $\varepsilon_0 > 0$ and $K > 0$ such that $E_\varepsilon(u_\varepsilon) \leq K$ for every $\varepsilon < \varepsilon_0$, where $u_\varepsilon$ is as in Proposition 5.1.

**Proof.** Step 1. We produce, for every $\varepsilon$ sufficiently small, a continuous path joining $v_1$ to $v_2$ and such that the maximum of $F_\varepsilon$ on the path is attained with a value that is bounded above independently of $\varepsilon$. There is a standard way to do this, see e.g. [Gua18], with the extra care that for the path produced in [Gua18] there is a uniform upper and lower bound for the term $\int_N gu_\varepsilon$, independently of $\varepsilon$. This implies a uniform upper bound for $F_\varepsilon(u_\varepsilon)$, by the minmax characterization of $u_\varepsilon$.

**Step 2.** A critical point $u_\varepsilon \in W^{1,2}(\Omega)$ of $F_\varepsilon$ solves the weak formulation of the semilinear elliptic PDE $\varepsilon \Delta u - W'(|u|) = g$. By elliptic theory $u_\varepsilon$ is $C^{2,\alpha}$. Note (arguing as we did when we chose $a_\varepsilon$ and $b_\varepsilon$) that any constant smaller than $-1$ is a lower barrier and any constant larger than $1 + \varepsilon$ is an upper barrier. The maximum principle therefore implies that any solution of the PDE (in particular $u_\varepsilon$) is bounded between $-1$ and $1 + \varepsilon$, which gives an $L^\infty$ bound $\|u_\varepsilon\|_{L^\infty(\Omega)} \leq 2$ uniformly in $\varepsilon$.

**Step 3.** Since $E_\varepsilon(u_\varepsilon) \leq F_\varepsilon(u_\varepsilon) + |N||g|\|u_\varepsilon\|_\infty$, we conclude from steps 1 and 2 a uniform upper bound for $E_\varepsilon(u_\varepsilon)$.

**Lemma 5.3.** There exist $\varepsilon_0 > 0$ and $L > 0$ such that $E_\varepsilon(u_\varepsilon) \geq L$ for every $\varepsilon < \varepsilon_0$, where $u_\varepsilon$ is as in Proposition 5.1.

**Proof.** We have $\liminf_{\varepsilon \to 0} \beta_\varepsilon > \varepsilon_0$ by definition of $\beta_\varepsilon = F_\varepsilon(u_\varepsilon)$ and by Lemma 5.1. Then $E_\varepsilon(u_\varepsilon) = \beta_\varepsilon + \int_N gu_\varepsilon$ and we have (see the proof of Lemma 5.3) that $u_\varepsilon > -1$, so $\int_N gu_\varepsilon \geq -\varepsilon_0$. Then we conclude that $\liminf_{\varepsilon \to 0} E_\varepsilon(u_\varepsilon) > 0$.

### 6 Proof of the existence theorem

**Remark 6.1.** Recall that, given $g > 0$, $g \in C^{1,1}(\Omega)$, we will prove the existence conclusion of Theorem 4.1 with $\frac{\sigma}{2}$ in place of $g$, because (for notational convenience) we work with the functional $F_{\varepsilon,g}$.

**First part of the proof of Theorem 4.1.** If there exists a sequence $\varepsilon_i \to 0$ such that the minmax critical points $u_{\varepsilon_i}$ (provided by Proposition 5.1) for $\varepsilon_i \to 0$ have the property that $u_{\varepsilon_i} \rightharpoonup u_\infty$ in $L^1(\Omega)$ and $u_\infty$ is not identically $-1$, then Theorem 4.1 implies that $\partial \{u_\infty = +1\}$ is a closed hypersurface that is quasi-embedded away from a possible singular set $\Sigma$ of dimension $\leq n - 7$ and with mean curvature given on $\partial \{u_\infty = +1\} \setminus \Sigma$ by $\frac{\sigma}{2} \nu$, where $\nu$ is the inward pointing normal to $\{u_\infty = +1\}$. In particular, $\partial \{u_\infty = +1\}$ is, possibly away from a set $\Sigma$ of dimension $n - 7$, the image of a two-sided $C^2$ immersion, with unit normal $\nu$ and mean curvature $\frac{\sigma}{2} \nu$. Theorem 4.1 is thus proved in the case in which $u_\infty \neq -1$ for some $u_{\varepsilon_i}$.
In order to establish Theorem 1.1, we only need to address the case in which \( u_\infty \equiv -1 \) for any sequence of minmax critical points \( u_{\varepsilon_i} \). In fact, to complete the proof, it will suffice to consider a single sequence \( u_{\varepsilon_i} \) for which \( u_\infty \equiv -1 \). By possibly passing to a subsequence, we may also assume that \( V^{\varepsilon_i} \to V \) as varifolds (with \( V \) stationary). Always by Theorem 4.1 and \( \varepsilon_i \to 0 \), \( V^{\varepsilon_i} \to V \) is, away from a singular set of dimension \( \leq n-7 \), a smooth minimal hypersurface \( M \) with locally constant even multiplicity. We will write \( M \) for the closure of \( M \); note that \( \overline{M} \) coincides with \( \text{spt} [\|V\|] \) and \( M \setminus \text{sing} \) is the singular set of \( V \). Recall that we have, in this situation, \( \limsup \varepsilon_i \to 0 \frac{1}{\varepsilon_i} \int_{\varepsilon_i} F_{\varepsilon_i}(u_{\varepsilon_i}) = \|V\|(N) + \frac{1}{\varepsilon_i} \int_N g \geq 2H^n(M) + \frac{1}{\varepsilon_i} \int_N g \). In the forthcoming sections, our goal will be to prove the following.

**Proposition 6.1.** Let \( u_{\varepsilon_i} \) be as in the preceding paragraph. Then there exist \( v_{\varepsilon_i} : N \to \mathbb{R} \) that solve \( F'_{\varepsilon_i}(v_{\varepsilon_i}) = 0 \) and \( F''_{\varepsilon_i}(v_{\varepsilon_i}) \geq 0 \) (i.e. stable critical points of \( F_{\varepsilon_i} \)) with \( \lim \inf v_{\varepsilon_i} \to 0 \mathcal{E}_{\varepsilon_i}(u_{\varepsilon_i}) > 0 \) and \( \lim \sup v_{\varepsilon_i} \to 0 \mathcal{E}_{\varepsilon_i}(v_{\varepsilon_i}) < \infty \); moreover, there exists a (fixed) non-empty open set that is contained in \( \{ v_{\varepsilon_i} > \frac{3}{4} \} \) for all \( \varepsilon_i \).

Second part of the proof of Theorem 1.1, assuming Proposition 6.1. Let \( u_{\varepsilon_i} \) be as in the paragraph preceding Proposition 6.1 and \( v_{\varepsilon_i} \) be the functions given by Proposition 6.1. Thanks to the condition that \( \{ v_{\varepsilon_i} > \frac{3}{4} \} \) contains a fixed non-empty open set, we obtain that any (subsequential) \( L^1 \)-limit \( v_\infty \) of \( v_{\varepsilon_i} \) equals +1 on this open set. Thanks to the upper and lower bounds on \( \mathcal{E}_{\varepsilon_i}(v_{\varepsilon_i}) \) we consider the associated varifolds \( V^{\varepsilon_i} \) as \( \varepsilon_i \to 0 \) and apply Theorem 4.1 to any subsequential limit. We obtain that the multiplicity-1 varifold associated to the reduced boundary of \( \{ v_\infty = +1 \} \) is non-trivial and provides the prescribed-mean-curvature hypersurface needed to complete the proof of Theorem 1.1.

The proof of Proposition 6.1 will be achieved, in the forthcoming sections, by exploiting the minmax characterisation of \( u_{\varepsilon_i} \) and the geometric fact that a minimal hypersurface with multiplicity 2 is not a stationary point for the geometric functional \( A - \text{Vol}_2 \), when viewed as an immersion from its double cover: the term \( A \) measures the hypersurface area and the term \( \text{Vol}_2 \) is the enclosed \( \frac{2}{\varepsilon} \)-volume, which is \( \int_E \frac{2}{\varepsilon} \) when the hypersurface in question is the reduced boundary of a Caccioppoli set \( E \subset N \). (See BelWic-2 for the more general definition that applies also to the case in which the hypersurface is not a boundary. This is the natural functional whose critical points are hypersurfaces with scalar mean curvature prescribed by \( \frac{2}{\varepsilon} \).) More specifically, we will proceed as follows. Recall that, under the assumptions made, we have a sequence \( u_{\varepsilon_i} \) of minmax critical points whose associated varifolds converge to a minimal hypersurface \( M \) endowed with locally constant even multiplicity, and \( \limsup \varepsilon_i \to 0 \frac{1}{\varepsilon_i} \int_{\varepsilon_i} F_{\varepsilon_i}(u_{\varepsilon_i}) \geq 2H^n(M) + \frac{1}{\varepsilon_i} \int_N g \). We will exhibit, for all sufficiently small \( \varepsilon = \varepsilon_i \), a continuous path \( \gamma : [-1, 1] \to W^{1,2}(N) \) with \( \gamma(-1) = a_{\varepsilon_i} \), with the second endpoint \( \gamma(+1) \) a stable solution to \( F_{\varepsilon_i} = 0 \), and with an energy bound \( \mathcal{F}_{\varepsilon}(\gamma(t)) \leq 2(2\sigma)H^n(M) - c_M + \int_N g \) for all \( t \in [-1, 1] \) and for some \( c_M > 0 \) independent of \( \varepsilon \) (\( c_M \) will depend only on \( M \subset N \)). Then the minmax characterisation of \( u_{\varepsilon_i} \) will imply that the second endpoint \( \gamma(+1) \) is a function \( v_\varepsilon \) that cannot be \( b_\varepsilon \). Owing to the way in which we will construct the path, \( v_\varepsilon \) will satisfy the remaining conditions in Proposition 6.1.

### 6.1 Preliminaries

**One-dimensional profiles.** We denote by \( H : \mathbb{R} \to \mathbb{R} \) the monotonically increasing solution to the Allen–Cahn ODE \( u'' - W'(u) = 0 \) such that \( \lim_{r \to \pm \infty} H(r) = \pm 1 \), with \( H(0) = 0 \). (In the case of the standard potential \( \frac{1}{2}(1-s^2)^2 \), we have \( H(r) = \tan \left( \frac{r}{\sqrt{2}} \right) \).) The rescaled function \( H_{\varepsilon}(r) = H \left( \frac{r}{\varepsilon} \right) \) solves the ODE \( \varepsilon u'' - \frac{W'(u)}{\varepsilon} = 0 \). We will need a truncated version \( \overline{H} \) of \( H \) in our construction: this
approximate solution is set to be constant \((\pm 1)\) away from \([-6 \log \varepsilon, 6 \log \varepsilon]\). This is convenient for the construction of an “Allen–Cahn approximation” of a hypersurface, i.e., a function on \(N\) that takes on large sets the values \(\pm 1\) and presents transitions between these two values along the hypersurface in question, and such that its Allen–Cahn energy \(\mathcal{E}_\varepsilon\) approximates the area of the hypersurface. Similar truncations are used and explained in detail in [Bel], so we do not elaborate on them here. For \(\Lambda = 3\|\log \varepsilon\|\) define

\[
\mathfrak{H}(r) = \chi(\Lambda^{-1}r - 1)\mathfrak{t}(r) \pm (1 - \chi(\Lambda^{-1}|r| - 1)),
\]

where \(+1\) or \(-1\) is chosen respectively on \(r > 0, r < 0\) and \(\chi\) is a smooth bump function that is \(+1\) on \((-1, 1)\) and has support equal to \([-2, 2]\). With this definition, \(\mathfrak{H} = \mathfrak{H}\) on \((-\Lambda, \Lambda), \mathfrak{H} = -1\) on \((-\infty, -2\Lambda)\), \(\mathfrak{H} = +1\) on \([2\Lambda, \infty)\). Moreover the function \(\mathfrak{H}\) solves \(\|\mathfrak{H}^{\prime\prime} - W'(\mathfrak{H})\|_{C^2(\mathbb{R})} \leq C\varepsilon^3\), for \(C > 0\) independent of \(\varepsilon\). (Note also that \(\mathfrak{H}^{\prime\prime} - W'(\mathfrak{H}) = 0\) away from \((-2\Lambda, -\Lambda) \cup (\Lambda, 2\Lambda)\).)

For \(\varepsilon < 1\), we rescale these truncated solutions and let \(\mathfrak{H}^\varepsilon(\cdot) = \mathfrak{H}(\varepsilon \cdot)\). Note that \(\mathfrak{H}^\varepsilon\) solves \(\|\varepsilon \mathfrak{H}^{\prime\prime} - W'(\varepsilon \mathfrak{H})\|_{C^2(\mathbb{R})} \leq C\varepsilon^2\) and \(\varepsilon \mathfrak{H}^{\prime\prime} - \frac{W'(\varepsilon \mathfrak{H})}{\varepsilon} = 0\) on \((-\varepsilon \Lambda, \varepsilon \Lambda)\), \(\mathfrak{H}^{\varepsilon} = +1\) on \((2\varepsilon, \infty)\), \(\mathfrak{H}^{\varepsilon} = -1\) on \((-\infty, -2\varepsilon)\).

Using these facts and recalling that \(\mathcal{E}_\varepsilon(\mathfrak{H}^\varepsilon) = 2\sigma\) we get \(\mathcal{E}_\varepsilon(\mathfrak{H}^\varepsilon) = 2\sigma + O(\varepsilon^2)\). (The function \(O(\varepsilon^2)\) is bounded by \(C\varepsilon^2\) for all \(\varepsilon\) sufficiently small, with \(C\) independent of \(\varepsilon\).)

The above profile will be needed to write Allen–Cahn approximations of multiplicity-1 hypersurfaces. In order to deal with multiplicity-2 portions we define \(\Psi : \mathbb{R} \to \mathbb{R}\)

\[
\Psi(r) = \begin{cases} 
\mathfrak{H}(r + 2\varepsilon\Lambda) & r \leq 0 \\
\mathfrak{H}^{\varepsilon}(r - 2\varepsilon\Lambda) & r > 0 
\end{cases}
\]  

(29)

(This function is smooth, since all derivatives of \(\mathfrak{H}^\varepsilon\) vanish at \(\pm 2\varepsilon\Lambda\).) We have \(\mathcal{E}_\varepsilon(\Psi) = 2(2\sigma) + O(\varepsilon^2)\). Additionally, we will need a continuous family of one-dimensional profiles that will be employed to replicate, for functions, the geometric operation of continuously changing the weight of a hypersurface; in a similar spirit, this family will also be used to produce Allen–Cahn approximations of hypersurfaces-with-boundary. In view of this we define, for \(t \in [0, \infty)\):

\[
\Psi_t(r) := \begin{cases} 
\mathfrak{H}(r + 2\varepsilon\Lambda - t) & r \leq 0 \\
\mathfrak{H}^{\varepsilon}(r - 2\varepsilon\Lambda - t) & r > 0 
\end{cases}
\]  

(30)

Note that \(\Psi_0 = \Psi\) and \(\Psi_t \equiv -1\) for \(t \geq 4\varepsilon\Lambda\). For \(t \in (0, 4\varepsilon\Lambda)\) the function \(\Psi_t\) is equal to \(-1\) for \(r\) such that \(|r| \geq 4\varepsilon\Lambda - t\). For each \(t\) the function \(\Psi_t\) is even and Lipschitz. The energy \(\mathcal{E}_\varepsilon(\Psi_t)\) is decreasing in \(t\): indeed we have \(\mathcal{E}_\varepsilon(\Psi_t) = \mathcal{E}_\varepsilon(\Psi) - \int_{-t}^t \varepsilon|\Psi'|^2 + \frac{W'(\Psi)}{\varepsilon} dt\).

**Distance to \(\overline{M}\).** Let \(M\) be as in the beginning of Section 6 (just before the statement of Proposition [6,1]). We denote by \(d_{\overline{M}} : N \to \mathbb{R}\) the Lipschitz function \(d_{\overline{M}}(x) = \text{dist}(x, \overline{M})\), where dist is the Riemannian distance. By Hopf–Rinow theorem, \(d_{\overline{M}}(x)\) is always realized by at least one geodesic from \(x\) to a point in \(\overline{M}\); in our case, the endpoint of such a geodesic will always belong to \(M\), see [Bel] Lemma 3.1. We let \(\omega \in (0, \text{inj}(N))\) and consider the open set \(T_\omega = \{x : d_{\overline{M}}(x) < \omega\}\). We will restrict our analysis to this neighbourhood of \(\overline{M}\).

By the analysis in [Bel] Section 3 (see also [ManMen02]), the function \(\nabla d_{\overline{M}}\) is in \(SBV(T_\omega \setminus \overline{M})\), i.e., it is a \(BV\)-function whose distributional derivatives are Radon measures with no Cantor part. More precisely, we have that (see [Bel] Lemma 3.2) the distributional Laplacian \(\Delta d_{\overline{M}}\) restricted to \(T_\omega \setminus \overline{M}\) is a Radon measure whose singular part is a negative. Moreover (see [ManMen02] and [Bel]...
Proposition 3.1), always restricting to \( T_\omega \setminus \bar{M} \), the support of the singular part of \( \Delta d_{\bar{M}} \) is countably \( n \)-rectifiable and agrees with the so-called cut-locus of \( M \), denoted by \( \text{Cut}(M) \); away from the cut-locus of \( M \), the Laplacian \( \Delta d_{\bar{M}} \) is smooth. For \( x \in T_\omega \setminus \text{Cut}(M) \setminus \bar{M} \) we have that there exists a unique geodesic from \( x \) to \( \bar{M} \) whose length realizes \( d_{\bar{M}}(x) \). This geodesic is completely contained (except for its endpoint, that lies in \( M \)) in the open set \( T_\omega \setminus \text{Cut}(M) \setminus \bar{M} \). This yields a retraction of \( T_\omega \setminus \text{Cut}(M) \) onto \( M \), see [Bel, Remark 3.2], with points moving towards \( M \) at unit speed along the unique geodesic connecting them to \( M \). Arguing as in [Bel, Lemma 3.3] by means of Riccati’s equation [Gra, Corollary 3.6], and replacing the condition \( \text{Ric}_N > 0 \) (valid in [Bel, Lemma 3.3]) with \( \text{Ric}_N > -C \) for some \( C > 0 \) (valid on our compact manifold \( N \)), we obtain that \( \Delta d_{\bar{M}} \leq C d_{\bar{M}}(x) \) on \( T_\omega \setminus \text{Cut}(M) \setminus \bar{M} \) (recall that \( d_{\bar{M}} \) is smooth on this open set). In fact, since \( d_{\bar{M}} \) is smooth on \( M \) and \( M \) is minimal, we have \( \Delta d_{\bar{M}} = 0 \) on \( M \), so \( \Delta d_{\bar{M}}(x) \leq C d_{\bar{M}}(x) \) on \( T_\omega \setminus \text{Cut}(M) \setminus (\bar{M} \setminus M) \).

**Proposition 6.2.** Let \( N \) be a compact \((n+1)\)-dimensional Riemannian manifold and \( M \) a smooth minimal hypersurface as in the beginning of Section 7. Denote by \( d_{\bar{M}} \) the distance function to \( \bar{M} \) and by \( T_\omega = \{ x \in N : d_{\bar{M}}(x) < \omega \} \), where \( \omega \in (0, \text{inj}(N)) \). Then the distributional Laplacian \( \Delta d_{\bar{M}} \) on \( T_\omega \) is a Radon measure that satisfies \( \Delta d_{\bar{M}} \leq C d_{\bar{M}} \), for \( C = -\min N \text{Ric}_N \).

**Proof.** We consider the distribution \( \Delta d_{\bar{M}} - C d_{\bar{M}} \), defined by its action on \( v \in C^\infty_c(T_\omega) \) by

\[
(\Delta d_{\bar{M}} - C d_{\bar{M}})(v) = -\int \nabla d_{\bar{M}} \cdot \nabla v - C \int_N d_{\bar{M}}v.
\]

Note that this is a distribution on \( T_\omega \) of order at most 1, because \( \nabla d_{\bar{M}} \in L^\infty(T_\omega) \) with \( |\nabla d_{\bar{M}}| \leq 1 \) and hence we get \( |(\Delta d_{\bar{M}} - C d_{\bar{M}})(v)| \leq H^{n+1}(N)(C\omega + 1)\|v\|_{C^1(T_\omega)} \).

Putting together \( \Delta d_{\bar{M}}(x) \leq C d_{\bar{M}}(x) \) on \( T_\omega \setminus \text{Cut}(M) \setminus (\bar{M} \setminus M) \) with the sign condition on the singular part of \( \Delta d_{\bar{M}} \) obtained in [Bel, Lemma 3.2] (and discussed above), we conclude that the restriction of the distribution \( \Delta d_{\bar{M}} - C d_{\bar{M}} \) to \( T_\omega \setminus (\bar{M} \setminus M) \) is \( \leq 0 \). This restriction is therefore a (negative) Radon measure on \( T_\omega \setminus (\bar{M} \setminus M) \).

We argue via a capacity argument similarly to [Bel, Proposition 3.2]. For any \( \delta > 0 \) we choose (see [EvaGar, 4.7]) \( \chi \in C^\infty_c(T_\omega) \) to be a function that takes values in \([0,1]\), is identically 1 in an open neighbourhood of \( \bar{M} \setminus M \), identically 0 away from a (larger) neighbourhood of \( \bar{M} \setminus M \) and such that \( \int_{T_\omega} |\nabla \chi| < \delta \). For \( v \in C^\infty_c(T_\omega) \), \( v \geq 0 \), we get

\[
\int_{T_\omega} (\Delta d_{\bar{M}} - C d_{\bar{M}})v = \int_{T_\omega} (\Delta d_{\bar{M}} - C d_{\bar{M}})(1 - \chi)v + \int_{T_\omega} \Delta d_{\bar{M}} \chi v - \int_{T_\omega} C d_{\bar{M}}\chi v =
\]

\[
= \int_{T_\omega} (\Delta d_{\bar{M}} - C d_{\bar{M}})(1 - \chi)v - \int_{T_\omega} \nabla d_{\bar{M}} \nabla \chi v - \int_{T_\omega} \nabla d_{\bar{M}} \nabla v \chi - \int_{T_\omega} C d_{\bar{M}}\chi v. \tag{31}
\]

The second, third and fourth terms in the last identity tend to 0 as \( \delta \to 0 \), because \( \|\chi\|_{W^{1,1}(T_\omega)} \to 0 \) as \( \delta \to 0 \) and \( |\nabla d_{\bar{M}}| \leq 1 \). The first term is \( \leq 0 \) for any \( \delta \), because \( (1 - \chi)v \geq 0 \) and we saw that the distribution \( \Delta d_{\bar{M}} - C d_{\bar{M}} \) is a negative Radon measure.

\[\text{This should be interpreted as an inequality between measures: the function } C d_{\bar{M}} \text{ is identified with the measure } C d_{\bar{M}} H^{n+1} \mathbb{1}_{T_\omega}, \text{ and we know already that the distributional Laplacian } \Delta d_{\bar{M}} \text{ is a Radon measure, therefore the inequality means that } C d_{\bar{M}} H^{n+1} - \Delta d_{\bar{M}} \text{ is a positive (Radon) measure on } T_\omega. \]

Also recall that a distribution is said to be \( \leq 0 \) if for every non-negative test function the result is \( \leq 0 \) (similarly for \( \geq 0 \)). A distribution that is \( \geq 0 \) or \( \leq 0 \) is necessarily a Radon measure, see e.g. [EvaGar, Theorem 1.39].
measure on the support of \((1-\chi)v\) (for any \(\delta\)). Taking the limit in (31) as \(\delta \to 0\) we therefore obtain that \(\Delta d_{\overline{M}} - C\overline{d}_{\overline{M}}\) is a negative distribution on \(T_{\omega}\), and therefore it is a (negative) Radon measure on \(T_{\omega}\).

We will denote by \(\tilde{M}\) the oriented double cover of \(M\). For \(q = (y,v) \in \tilde{M}\), for \(y \in M\) and \(v\) a choice of unit normal to \(M\) at \(y\), the geodesic \(s \in (0,\text{inj}(N)) \to \exp_y(sv)\) leaves \(M\) orthogonally and is minimizing, between \(y\) and the point \(\exp_y(tv)\), as long as \(t\) is sufficiently small. We denote by \(\sigma_{(y,v)}\) the (positive) number such that this geodesic is minimizing between \(y\) and the point \(\exp_y(te)\) for all \(t \leq \sigma_{(y,v)}\) and is no longer minimizing if \(t > \sigma_{(y,v)}\). As we are restricting to \(T_{\omega}\), we truncate \(\sigma_{(y,v)}\) using the convention that \(\sigma_{y,v} = \omega\) if the geodesic is minimizing for some \(t > \omega\). The function \(\sigma_{(y,v)}\) on \(\tilde{M}\) is continuous (see [Bel, Section 3] and [ManMen02] for details); moreover, we have a smooth diffeomorphism \(F\)

\[
F : \{(y,v), s : (y,v) \in \tilde{M}, s \in (0, \sigma_{(y,v)})\} \to T_{\omega} \setminus \text{Cut}(M) \setminus \overline{M}
\]

that extends by continuity to a map from \(V_{\tilde{M}} = \{(y,v), s : (y,v) \in \tilde{M}, s \in [0, \sigma_{(y,v)})\} \to T_{\omega} \setminus (\text{Cut}(M) \cup (\overline{M} \setminus M))\). This extended map will be still denoted by \(F\) and is \(2-1\) on \(\tilde{M} \times \{0\}\). Since \(\sigma_{(y,v)} > 0\), on any compact subset of \(M\) there is a positive lower bound for \(\sigma_{(y,v)}\) and therefore the map \(F\) provides, around any compact set of \(M\), a system of Fermi coordinates (tubular neighbourhood system).

We next collect certain properties of the level sets \(\Gamma_t = \{x \in N : d_{\overline{M}}(x) = t\}\) for \(t \in [0,\omega)\). For \(t \in (0,\omega)\) we have that \(\Gamma_t \setminus \text{Cut}(M)\) is smooth, with scalar mean curvature at \(x\) given by \(-d_{\overline{M}}(x)\). For \(H^1\)-a.e. \(t \in (0,\omega)\) we have \(H^n(\Gamma_t \cap \text{Cut}(M)) = 0\), since \(\text{Cut}(M)\) has dimension \(n\). Therefore we have that \(H^1\)-a.e. level set is \(H^n\)-a.e. smooth. For these level sets we can therefore compute the \(H^n\)-measure by computing the measure of their smooth part. For this we argue as in [Bel, Lemma 4.1] (to which we refer for further details). We use [Gra, Theorem 3.11] to compute the distortion of the area element as we move along a geodesic \(s \in (0, \sigma_{(y,v)}) \to \exp_y(sv)\), for \((y,v) \in \tilde{M}\), the oriented double cover of \(M\) (in other words, \(y \in M\) and \(v\) is one of the two choices of unit normal to \(M\) at \(y\)). The distortion of the area element is ruled by the ODE \(\frac{\partial}{\partial s} \log \theta_s = -\tilde{H}(y,s) \cdot \frac{\partial}{\partial s}\), where \(\tilde{H}(y,s)\) is the mean curvature of the level set at distance \(s\) evaluated at the point \((y,s) = \exp_y(sv)\). Using Riccati’s equation [Gra, Corollary 3.6], and the bound \(\text{Ric}_N \geq -C\) on \(N\), we obtain that \(\tilde{H}(y,s) \geq -Cs\), where \(\tilde{H}(y,s) = \tilde{H}(y,s) \cdot \frac{\partial}{\partial s}\) is the scalar mean curvature of the level set \(\Gamma_s\) at the point \((y,s)\), with respect to the unit normal that points away from \(M\). We thus have \(\frac{\partial}{\partial s} \log \theta_s \leq Cs\). Integrating this inequality we obtain that, with coordinates chosen so that \(\theta_0(x) = 1\), the area element evolves with the bound \(\theta_s(x) \leq e^{Cs^2/2}\). Therefore

\[
H^n(\Gamma_t) \leq 2H^n(M)e^{Ct^2/2}\text{ for almost every } t \in (0,\omega)
\]

(level sets of \(d_{\overline{M}}\) are double covers of \(M\), since \(d_{\overline{M}}\) is unsigned). Recall also that the scalar mean curvature of \(\Gamma_t\) at the point \(x \in \Gamma_t \setminus \text{Cut}(M)\) agrees with \(-\Delta d_{\overline{M}}(x)\).

The following estimate is implicit in the previous discussion:

\[
x \in \Gamma_t \setminus \text{Cut}(M) \Rightarrow \Delta d_{\overline{M}}(x) \leq Ct.
\]

### 6.2 Allen–Cahn approximation of \(2|M|\)

In this section we will produce for all sufficiently small \(\varepsilon\), a function \(G^\varepsilon_0 : N \to \mathbb{R}\) whose Allen–Cahn energy is\(^3\) approximately \(2H^n(M)\). While this function is not

\(^3\)We will only be interested in obtaining a control from above of the energy by the area, up to a small error term. It is however true, as can be seen by computations similar to those that we give in this section, that a control from below is also valid.
part of the path $\gamma$ that we aim to construct (see the discussion that precedes Section 6.1), by suitably deforming it we will construct two functions through which the path $\gamma$ will pass.

Let $\varepsilon > 0$ be sufficiently small to ensure $12 \varepsilon |\log \varepsilon | < \min \{\omega, 1\}$. The function $G_0^\varepsilon$ has level sets coinciding with the level sets of $d_{\overline{\mathcal{M}}}$ and the one-dimensional profiles normal to these level sets are dictated by $\mathcal{H}^1$ (see [29] for the definition of $\Psi$):

$$G_0^\varepsilon(x) = \left\{ \begin{array}{ll} -1 & \text{for } x \in N \setminus T_\omega \\
\Psi(d_{\overline{\mathcal{M}}}(x)) & \text{for } x \in T_\omega \end{array} \right. \quad (35)$$

Equivalently, for $x \in T_\omega$ we have $G_0^\varepsilon(x) = \mathcal{H}^1(- d_{\overline{\mathcal{M}}}(x) + 2 \varepsilon \Lambda)$. We will now compute $\mathcal{E}_\varepsilon (G_0^\varepsilon)$ and the first variation $\mathcal{E}'_\varepsilon (G_0^\varepsilon)$.

We use the shorthand notation $\Lambda = 3 |\log \varepsilon |$. By definition we have, on $T_\omega$, that $\nabla G_0^\varepsilon = \Psi'(d_{\overline{\mathcal{M}}}) \nabla d_{\overline{\mathcal{M}}}$, while the energy of $G_0^\varepsilon$ is 0 in $N \setminus \{x \in N : d_{\overline{\mathcal{M}}}(x) \leq 4 \varepsilon \Lambda\}$; we use the coarea formula for the Lipschitz function $d_{\overline{\mathcal{M}}}$ (for which $|\nabla d_{\overline{\mathcal{M}}} | = 1$) to get

$$\mathcal{E}_\varepsilon (G_0^\varepsilon) = \int_{T_\omega} \frac{|\nabla G_0^\varepsilon|^2}{2} + \frac{W(G_0^\varepsilon)}{\varepsilon} = \int_{0}^{\omega} \left( \int_{T_x} \frac{|\nabla G_0^\varepsilon|^2}{2} + \frac{W(G_0^\varepsilon)}{\varepsilon} \right) ds =$$

$$= \int_{-2 \varepsilon \Lambda}^{2 \varepsilon \Lambda} \left( \int_{T_{2 \varepsilon \Lambda - s}} \frac{\mathcal{H}^1(\varepsilon s)^2}{2} + \frac{W(\mathcal{H}^1(\varepsilon s))}{\varepsilon} \right) ds \leq 2 e^{\varepsilon (12 \varepsilon |\log \varepsilon |)^2} \mathcal{H}^n(M) \left( \int_{\mathbb{R}} \frac{e^{\varepsilon (\mathcal{H}^1(\varepsilon s))^2}}{2} + \frac{W(\mathcal{H}^1(\varepsilon s))}{\varepsilon} \right) \leq 2(2\sigma) \mathcal{H}^n(M) + O(\varepsilon |\log \varepsilon |),$$

for $\varepsilon$ in the chosen range. The Allen–Cahn first variation of $G_0^\varepsilon$ (which is clearly 0 outside $\{x \in N : d_{\overline{\mathcal{M}}}(x) \leq 4 \varepsilon \Lambda\}$) can be computed in $N \setminus \{x \in N : d_{\overline{\mathcal{M}}}(x) \leq 4 \varepsilon \Lambda\}$ as follows. The Radon measure $\Delta d_{\overline{\mathcal{M}}}$ satisfies Proposition 6.2; recall moreover the error by which $\mathcal{H}^n$ fails to solve the Allen–Cahn ODE (Section 6.1). Then, in the distributional sense, we have

$$-\mathcal{E}'_\varepsilon (G_0^\varepsilon) = \varepsilon \Delta G_0^\varepsilon - \frac{W'(G_0^\varepsilon)}{\varepsilon} = \quad (36)$$

$$= \varepsilon \mathcal{H}''(\varepsilon (- d_{\overline{\mathcal{M}}} + 2 \varepsilon \Lambda)) \nabla d_{\overline{\mathcal{M}}} - \mathcal{H}''(- d_{\overline{\mathcal{M}}} + 2 \varepsilon \Lambda) \Delta d_{\overline{\mathcal{M}}} - \frac{W'(\mathcal{H}^1(- d_{\overline{\mathcal{M}}} + 2 \varepsilon \Lambda))}{\varepsilon} =$$

$$= \varepsilon \mathcal{H}''(\varepsilon (- d_{\overline{\mathcal{M}}} + 2 \varepsilon \Lambda)) - \mathcal{H}''(- d_{\overline{\mathcal{M}}} + 2 \varepsilon \Lambda) \Delta d_{\overline{\mathcal{M}}} = \quad (37)$$

Here $-\mathcal{E}'_\varepsilon (G_0^\varepsilon)$ and $\Delta d_{\overline{\mathcal{M}}}$ are a Radon measures. The term $O(\varepsilon^2)$ in the last line is a Lipschitz function that we interpret as a density with respect to $\mathcal{H}^{n+1}$ and the last term is the measure $\Delta d_{\overline{\mathcal{M}}}$ multiplied by a bounded Lipschitz function (we have used $0 \leq (\mathcal{H}^1)' \leq 3/\varepsilon$). Note that $d_{\overline{\mathcal{M}}} \leq 4 \varepsilon \Lambda$ on the relevant domain. Therefore there exist $\varepsilon N \in (0, 1)$ and $C_0$ (only depending on $N$) such that for all $\varepsilon < \varepsilon_N$

$$-\mathcal{E}'_\varepsilon (G_0^\varepsilon) \geq -C_0 \varepsilon |\log \varepsilon |,$$
6.3 Immersions and signed distance

We will need a certain notion of signed distance. Fix a (non-empty) compact set $K \subset \hat{M}$; this set will be kept fixed throughout the construction in the coming sections. It is convenient to choose $K$ to be even, i.e. $K = \iota^{-1}(F(K))$. Let $\tilde{\phi} \in C^\infty_c(\hat{M})$, $\tilde{\phi} \geq 0$, such that $\text{supp}\tilde{\phi} \subset \text{Int}(K)$. The continuous function $\sigma_{(y,v)}$ (see (32) and the discussion preceding it) has a strictly positive minimum on $K$ and we choose $\sigma_K > 0$ strictly smaller than this minimum. Consider, for $c \in (0, \sigma_K/3)$ and $t \in [0, \frac{a_K}{3 \max \tilde{\phi}}]$, the following immersions:

$$p = (y, v) \in \text{Int}(K) \to \exp_{(\iota^{-1}(F(K)))}(c + t\tilde{\phi}(y))v.$$

The image of this immersion is a smooth embedded hypersurface. Note that the immersion extends smoothly up to the boundary, because in a neighbourhood of $\partial K$ we have $\tilde{\phi} = 0$. We will denote by $K_{c,t,\tilde{\phi}}$ the image of $K$ via the immersion. Note that the image of $K \setminus \text{supp}\tilde{\phi}$ in contained in the level set $\Gamma_c$.

For a point $(q, s) \in \text{Int}(K) \times (0, \sigma_K)$ we define its signed distance to $\text{graph}(c + t\tilde{\phi})$ to be non-negative for $s < (c + t\tilde{\phi})(q)$, positive for $s > (c + t\tilde{\phi})(q)$ and vanishing on $\text{graph}(c + t\tilde{\phi})$, with absolute value equal to the Riemannian distance of $(q, s)$ to $\text{graph}(c + t\tilde{\phi})$, where the Riemannian distance is the one induced by the pull-back via $F$ of the metric on $N$. This distance descends to a well-defined (smooth) signed distance $\text{dist}_{K_{c,t,\tilde{\phi}}}$ on $F(\text{Int}(K) \times [0, \sigma_K]) \setminus M$.

Since $c + t\tilde{\phi}$ is smooth up to $\partial K$ (and extends smoothly to an open neighbourhood of $K$ with value $c$), there exists a tubular neighbourhood of $K_{c,t,\tilde{\phi}}$ in which the nearest point projection onto $K_{c,t,\tilde{\phi}}$ is well-defined. We denote this projection by $\Pi_{c,t}$. Upon choosing the tubular neighbourhood sufficiently small, we also ensure that in the tubular neighbourhood of $F((K \setminus \text{supp}\tilde{\phi}) \times \{c\})$ the projection $\Pi_{c,t}$ agrees with the nearest point projection onto $\Gamma_c$, denoted by $\Pi_c$.

We choose $c_1 > 0$, $t_1 > 0$ such that for all $c \in (0, c_1]$ and all $t \in [0, t_1]$ there exists a tubular neighbourhood of $K_{c,t,\tilde{\phi}}$ of semi-width $c$ in which the nearest point projection $\Pi_{c,t}$ is well-defined, it coincides with $\Pi_c$ in the tubular neighbourhood of $F((K \setminus \text{supp}\tilde{\phi}) \times \{c\})$, and moreover the following bounds hold. There exists $\kappa_K > 0$ such that for all $x$ in the tubular neighbourhood of $K_{c,t,\tilde{\phi}}$ of semi-width $c$

$$|\frac{1}{|\Pi_{c,t}(x)|} - 1| \leq \kappa_K s$$

where $|\Pi_{c,t}| = \sqrt{(\Pi_{c,t}(x))}'(\Pi_{c,t})$ and $s$ is the distance of $x$ to $K_{c,t,\tilde{\phi}}$.

The way in which we will use these properties (in the forthcoming sections) is that for each fixed $\varepsilon$ (sufficiently small) we will work with $c = 4\varepsilon \Lambda$ and with variable $t$, and with tubular neighbourhoods of semi-width $4\varepsilon \Lambda$. The choice is made so that we can fit one-dimensional Allen-Cahn profiles in the normal bundle to $K_{c,t,\tilde{\phi}}$. The estimates in (38) guarantee that the Allen–Cahn energy of the resulting function (defined in the tubular neighbourhood of $K_{c,t,\tilde{\phi}}$) is very close to the area of $K_{c,t,\tilde{\phi}}$ (up to the usual multiplicative constant $2\sigma$). The fact that $K_{c,t,\tilde{\phi}}$ agrees with $\Gamma_c$ on its boundary will give that the function just constructed can be extended to a Lipschitz function on $N$, thanks to the properties of $d_{\Pi_{c,t}}$.

We will additionally make use of the following fact. There exists a constant $\kappa_K > 0$ (that we can assume is the same as the one appearing in (38)) depending only on $K \subset N$ such that the nearest point projection $\Pi_K : K \times [0, \sigma_K) \to K$

---

4When $\overline{M} = M$, e.g. for $n \leq 6$, one can choose $K = \overline{M}$ and the whole construction presented in the remaining sections can be shortened considerably.
satisfies
\[ |J\Pi_K| (x) - 1| \leq \kappa_K s \quad \text{and} \quad \frac{1}{|J\Pi_K| (x)} - 1 \leq \kappa_K s, \]
where \(|J\Pi_K| = \sqrt{(\nabla \Pi_K)(\nabla \Pi_K)^T}\) and \(x = (q, s)\). The same notation \(\Pi_K\) will be used to denote the nearest point projection from \(F(K \times [0, \sigma_K])\) onto \(F(K) \subset M\) (recall that \(s\) is the distance of \(F(x)\) to \(F(K)\)).

6.4 Choice of \(\varepsilon\) and geometric quantities involved

Recall that our final aim, in order to achieve the proof of Proposition 6.1, is to produce, for each sufficiently small \(\varepsilon\), a continuous path \(\gamma\) (with values in \(W^{1,2}\)) that joins \(a\) to another stable solution (see the discussion after the statement of Proposition 6.1). The path itself will be exhibited for each \(\varepsilon < \varepsilon_1\), for a certain \(\varepsilon_1 > 0\). Estimates on the energy \(\mathcal{F}_\varepsilon\) along the path will be obtained in terms of certain (fixed) geometric quantities (e.g. \(\mathcal{H}^n(M)\)), that are independent of \(\varepsilon\), and error terms. For all \(\varepsilon \leq \varepsilon_2\), for a certain \(\varepsilon_2 \in (0, \varepsilon_1]\), these error terms will be of the type \(O(\varepsilon \log \varepsilon)\), i.e. they will be bounded, in absolute value, by \(C\varepsilon \log \varepsilon\) with \(C > 0\) independent of \(\varepsilon \in (0, \varepsilon_2]\). Finally (this will only happen in Section 6.10), for a certain \(\varepsilon_3 \in (0, \varepsilon_2]\), the terms \(O(\varepsilon \log \varepsilon)\) will be absorbed in the geometric quantities, leading to an effective estimate on the energy \(\mathcal{F}_\varepsilon\) along the path, i.e. an estimate that only depends on geometric quantities.

Rather than picking \(\varepsilon_1\) here, we will require a smallness condition on it repeatedly (finitely many times) throughout the forthcoming sections. One smallness requirement was made in Section 6.2: \(12\varepsilon_1 \log \varepsilon_1 < \min\{\omega, 1\}\). At every new requirement, we will implicitly assume that all those previously imposed remain valid. Similarly, upon estimating \(\mathcal{F}_n\) for the functions that are constructed for \(\varepsilon < \varepsilon_1\), we will (finitely many times) write the error terms in the form \(O(\varepsilon \log \varepsilon)\) for \(\varepsilon \in (0, \varepsilon_2]\); each time we specify a new \(\varepsilon_2\) we will implicitly assume that we pick the smallest \(\varepsilon_2\) among all those identified until that moment. An initial requirement is \(\varepsilon_2 \leq \varepsilon_N\), for the \(\varepsilon_N\) chosen in Section 6.2. At the end (Section 6.10) we will choose \(\varepsilon_3 \in (0, \varepsilon_2]\) and restrict to \(\varepsilon \in (0, \varepsilon_3]\); for this range, the energy estimates become effective and allow us to conclude the proof.

From now on we shall let the sequence \(u_{\varepsilon_j}\) be as in the beginning of Section 6. Note however that until Section 6.10, all we need to know about \(u_{\varepsilon_j}\) is that the corresponding sequence of varifolds \(V^{u_{\varepsilon_j}}\) converges to \(q|\mathcal{M}|\) where \(q\) is a locally constant function on \(\mathcal{M}\) taking even integer values and \(|\mathcal{M}|\) is stationary (i.e. zero mean curvature); the fact that \(\mathcal{F}_{\varepsilon_j}(u_{\varepsilon_j})\) are min-max values does not play a role until Section 6.10. The minimal hypersurface \(\mathcal{M}\) however plays a key role throughout. We fix the compact set \(K\) (chosen in Section 6.3) and two geodesic balls \(D_1, D_2 \subset \mathcal{M}\) whose double covers have compact closures contained in \(K\) and such that the respective concentric balls of half the radius (that will be denoted by \(B_j \subset D_j\) for \(j \in \{1, 2\}\)) have equal areas. Geometric quantities depending on \(D_1, D_2, \mathcal{M}, N\) will appear in the energy estimates. To avoid burdening the notation, we will often drop the explicit dependence on \(\varepsilon\) for the functions, and paths of functions, that we construct. With reference to the upcoming sections, \(\varepsilon, D_{1,2}, t, \gamma_t, f_t, h_t\) all implicitly depend on \(\varepsilon\).

6.5 One-parameter deformation, bumping outwards

Let \(\iota : \tilde{M} \to N\) be the minimal immersion of the oriented double cover \(\tilde{M}\) of \(M\) induced by the standard projection \((y, v) \in \tilde{M} \to y \in M\). (The image of \(\iota\) is \(M\) and is covered twice.) Let \(D \subset M\) be a geodesic ball, with radius denoted by \(2R > 0\), with closure contained in \(\text{Int}(K)\) (where \(K\) is the compact set fixed at the beginning of Section 6.3); here \(D\) is either of the geodesic balls \(D_1, D_2\).
one-parameter (one-sided) smooth family of immersions

\[
\{ \tilde{\iota}_t \}_{t \in \mathbb{R}}
\]

\[D \times [0, \sigma_K] \] (recall the choice of \( \sigma_K \) at the beginning of Section 6.3):

\[p = (y, v) \in \tilde{D} \rightarrow \iota_{D,t}(p) = \exp_y(t \chi_{\tilde{D}}(p) v). \quad (40)
\]

We have \( \iota_{D,t} = \iota_{\tilde{D}} \) on \( \tilde{D} \setminus \text{supp} \chi_{\tilde{D}} \) for every \( t \), and \( \iota_{0,D} = \iota_{\tilde{D}} \).

The key property of interest to us for this (one-sided) deformation of \( \iota_{\tilde{D}} \) is going to be the value of the functional \( J_{g}(t) = J_{\tilde{g}}(\iota_{t,D}) \), where \( J_{\tilde{g}} = A - \text{Vol}_{\tilde{g}} \). The evaluation of \( \text{Vol}_{\tilde{g}} \) requires the preliminary choice of an oriented open set \( O \subset N \) (that must contain the images of the immersions); in our case we choose the cylinder \( F(\tilde{D} \times [0, \sigma_K]) \), with \( F \) defined in (42).

The value of \( \text{Vol}_{\tilde{g}}(t) \) is then given simply by

\[
\int_{F(S_{t})} \frac{2}{\sigma^{n_D}} dH^{n+1},
\]

where \( S_t = \{ (y, s) \in \tilde{D} \times [0, \sigma_K) : s < \tau_{\tilde{g}}(y) \} \). The term \( A(t) \) is the \( n \)-dimensional area of the immersion \( \iota_{t,D} \). We refer to \([BelWic-2]\) for the general definition of \( J_{g} \).

It is readily checked that \( J_{\tilde{g}}(t) \) decreases on \( t \in [0, t_0] \), for some \( t_0 \in (0, 1/ \max \chi_{D}) \).

This can be seen by considering the first variation \( J_{\tilde{g}}'(0) \). The first variation of area is 0 because \( \iota \) is minimal, while the first variation of \( \text{Vol}_{\tilde{g}} \) is given by integrating on \( \tilde{D} \) the product of \( \frac{2}{\sigma^{n_D}} \) with the normal speed \( \left( \frac{\partial}{\partial t} \right)_{t=0} \cdot v = \chi_{\tilde{D}} \).

We thus obtain \( J_{\tilde{g}}'(0) = -\int_{\tilde{D}} \frac{2}{\sigma^{n_D}} \chi_{\tilde{D}} < 0 \). Therefore there exists \( t_0 > 0 \) such that \( J_{\tilde{g}}(t) \) decreases on \( t \in [0, t_0] \).

**Remark 6.2** (smallness of \( t_0 \)). We will need, in view of Section 6.9, to possibly make \( t_0 \) smaller. Namely, we require that the mean curvature of immersion (40) is, in absolute value, smaller than \( \min_{x \in \tilde{D}} \sigma_x \) for all \( t \in [0, t_0] \). This is possible, since \( \min_{x \in \tilde{D}} \sigma_x > 0 \), the values of the mean curvature change continuously in \( t \) and at \( t = 0 \) the immersion is minimal. In the next remark we impose another condition, which may require to make \( t_0 \) smaller one more time.

**Remark 6.3** (choice of \( \tau_B \)). We will denote by \( \tau_B > 0 \) the value \( J_{g}(0) - J_{g}(t_0) \). By possibly choosing a smaller \( t_0 > 0 \), we ensure that \( \tau_B < 2H^n(B) \). This will be used in Section 6.10.

We point out, for future reference, that \( \iota_{t,D} \) and \( \iota \) coincide in a neighbourhood of \( \partial \tilde{D} \) in \( \tilde{D} \). We now consider, for \( c \in [0, \frac{2}{\sqrt{n}}] \), \( t \in \left[ 0, \frac{\sigma_K}{\max \chi_{D}} \right] \), the immersion from \( \tilde{D} \) into \( N \) given by

\[
p = (y, v) \in \tilde{D} \rightarrow \exp_y((c + t\chi_{\tilde{D}}(p)) v).
\]

(For \( c > 0 \) the image of such an immersion is embedded.) We denote by \( J_{\tilde{g}}(c, t) \) the evaluation of \( J_{\tilde{g}} \) on the immersion (41), with reference to the oriented open set \( F(\tilde{D} \times [0, \sigma_K]) \).

We recall that \( J_{g}(c, t) = A(c, t) - \text{Vol}_{\tilde{g}}(c, t) \), where \( A(c, t) \) is the area of the immersion (41) and \( \text{Vol}_{\tilde{g}}(c, t) = \int_{F(S_{c,t})} \frac{2}{\sigma^{n_D}} dH^{n+1} \), with \( S_{c,t} = \{ (y, s) \in \tilde{D} \times [0, \sigma_K) : s < c + t\chi_{\tilde{D}}(y) \} \). (Note that \( J_{\tilde{g}}(0, t) \) agrees with \( J_{\tilde{g}}(t) \) used above.) By continuity in \( c \) and \( t \), and by the decreasing property of \( J_{\tilde{g}}(0, t) \) on \( [0, t_0] \) obtained above, there exist \( c_0 \in \left( 0, \frac{2}{\sqrt{n}} \right) \) such that for all \( c \in [0, c_0] \) and for all \( t \in [0, t_0] \)

\[
J_{g}(c, t) \leq J_{g}(0, 0) + \frac{\tau_B}{2}, \quad J_{g}(c, t_0) \leq J_{g}(0, 0) - \tau_B,
\]

(42)
where \( \tau_B > 0 \) was chosen above. Note that \( J_2(0,0) = 2\mathcal{H}^{n+1}(D) \), since the term \( \text{Vol}_g \) vanishes for \( t_0, D = \partial \tilde{D} \).

Remark 6.4. By making \( c_0 \) and \( t_0 \) smaller if necessary, we also assume that \( c_0 < c_1 \) and \( t_0 < t_1 \), where \( t_1 \) and \( c_1 \) are chosen as in Section 6.3 with \( \phi \) replaced by the function \( \tilde{\chi}_D \), extended to be in \( C_0^\infty(K) \), by setting it equal to 0 in the complement of \( \tilde{D} \).

The geometric deformation of \( \partial \mathcal{D} \) just described will now be replicated with a family of functions in \( W^{1,2}(F(\tilde{D} \times [0, \sigma_K])) \), with bounds on \( \mathcal{F}_\varepsilon \) that will replace (and be deduced from) the estimates in \( \{12\} \). We define, initially, a one-parameter family of functions on \( \tilde{D} \times [0, \sigma_K] \subset V_{\mathcal{M}} \), using coordinates \((q,s)\) and then we pass the definition to \( F(\tilde{D} \times [0, \sigma_K]) \). In the following we assume \( \varepsilon < \varepsilon_1 \), with \( 6\varepsilon_1 \log \varepsilon_1 < c_0 \). We use the shorthand notation \( 2\varepsilon \Lambda = 6\varepsilon \log \varepsilon \). We set, for \( t \in [0,t_0] \) and \((q,s) \in \tilde{D} \times (0, \sigma_K) \) (and for each \( \varepsilon \) in the specified range)

\[
\varpi_{\varepsilon, t}(q,s) = \frac{1}{\varepsilon} \left( -\text{dist}_{\text{graph}(2\varepsilon \Lambda + t \chi_{\tilde{D}})}(q,s) \right).
\]

Here we are using the signed distance to graph\((2\varepsilon \Lambda + t \chi_{\tilde{D}})\) that was discussed in Section 6.3. Note that although the distance is not defined for \( s = 0 \), the definition in \( \{43\} \) extends continuously from \( s > 0 \) to \( s \geq 0 \) with value 1 at \( s = 0 \); this follows upon observing that \( \lim_{s \to 0} \text{dist}_{\text{graph}(2\varepsilon \Lambda + t \chi_{\tilde{D}})}(q,s) = -2\varepsilon \Lambda \) and that \( \mathcal{H}^n \) has value 1 on \((2\varepsilon \Lambda, \infty)\). More precisely, \( \mathcal{H}^n \) has vanishing derivative at \( 2\varepsilon \Lambda \).

This guarantees that the function defined in \( \{43\} \), extended to \( s = 0 \) with value 1, passes to the quotient as a \( C^1 \) function on the open set \( F(D \times (0, \sigma_K)) \) \( (\text{this is a tubular neighbourhood of } D) \).

With slight abuses of notation, we use the notation \( \varpi_{\varepsilon, t} \) also to denote this quotient and we write \( \varpi_{\varepsilon, t} : F(D \times (0, \sigma_K)) \to \mathbb{R} \).

Here \( K_{2\varepsilon \Lambda, t, \chi_{\tilde{D}}} \) denotes, as in Section 6.3, the set \( F(\text{graph}(2\varepsilon \Lambda + t \chi_{\tilde{D}})) \). As before, the function \( \varpi_{\varepsilon, t} \) is extended in a \( C^1 \) fashion across \( D \), with value 1 on \( D \).

We remark that for \( t = 0 \) we have \( \varpi_{\varepsilon, 0} = G_{6}^\varepsilon|_{(\tilde{D} \times (0, \sigma_K))} \). Moreover, we point out that (for each \( \varepsilon \) considered) the assignment \( t \in [0,t_0] \to \varpi_{\varepsilon, t} \in W^{1,2}(F(\tilde{D} \times (0, \sigma_K)) \) is continuous.

The Allen–Cahn energy of \( \varpi \) can be computed using the coarea formula (either with respect to the distance, for which the Jacobian factor is 1, or with respect to the nearest point projection, using \( \{38\} \)) in a tubular neighbourhood of \( K_{2\varepsilon \Lambda, t, \chi_{\tilde{D}}} \cap F(D \times (0, \sigma_K)) \) of semi-width \( 2\varepsilon \Lambda \), away from this tubular neighbourhood, \( \varpi_{\varepsilon, t} \) is constantly \( \pm 1 \), hence there is no energy contribution. The coarea formula gives that there exists \( \varepsilon_2 > 0 \) such that for all \( \varepsilon \leq \varepsilon_2 \) the following bound holds:

\[
\mathcal{E}_\varepsilon(\varpi_{\varepsilon, t}) \leq (2\varepsilon)\mathcal{H}^n(K_{2\varepsilon \Lambda, t, \chi_{\tilde{D}}} \cap F(D \times (0, \sigma_K))) + O(\varepsilon | \log \varepsilon |).
\]

The set \( K_{2\varepsilon \Lambda, t, \chi_{\tilde{D}}} \cap F(D \times (0, \sigma_K)) \) is the image of the immersion in \( \{41\} \) (for which we obtained the estimates in \( \{12\} \) — we will use these to obtain \( \{16\} \) below).

In order to get an estimate for \( \mathcal{F}_\varepsilon(\varpi_{\varepsilon, t}) \) we now consider \( \int_{\tilde{D} \times (0, \sigma_K)} g \varpi_{\varepsilon, t} \, d\mathcal{H}^{n+1} \) and relate this quantity to \( \text{Vol}_g(2\varepsilon \Lambda, t, \Lambda) \). As in \( \{12\} \), \( \text{Vol}_g(c,t) \) is the evaluation of

\[\text{To be coherent with Section 6.3, } \chi_{\tilde{D}} \text{ must be extended to a function in } C_0^\infty(K), \text{ by setting it equal to 0 in the complement of } \tilde{D} \text{ and the graph in question must be considered as a graph on } K. \text{ Then the resulting distance, defined in Section 6.3 on } K \times (0, \sigma_K), \text{ has to be restricted to } \tilde{D} \times (0, \sigma_K).\]

\[\text{As in the previous footnote, the graph is taken over } K \text{ and } \chi_{\tilde{D}} \text{ is extended to a function in } C_0^\infty(K).\]
Vol$\gamma$ on the immersion \([41]\) and we are choosing \(c = 2\varepsilon\Lambda\). Let \(g\) denote the function that is \(+1\) on \(\{ (q, s) \in \tilde{D} \times [0, \sigma_K) : \text{dist}_{\text{graph}(2\varepsilon\Lambda + t\chi_D)}((q, s)) \leq 0 \}\) and \(-1\) on \(\{ (q, s) \in \tilde{D} \times [0, \sigma_K) : \text{dist}_{\text{graph}(2\varepsilon\Lambda + t\chi_D)}((q, s)) > 0 \}\). Then \(\int_{\tilde{D} \times [0, \sigma_K]} g \, d\mathcal{H}^{n+1} = 2\int_{\{ t\varepsilon = 1 \}} g \, d\mathcal{H}^{n+1} = 2\int_{\tilde{D} \times [0, \sigma_K]} g \, d\mathcal{H}^{n+1} = 2\text{Vol}_\gamma(2\varepsilon\Lambda, t) - \int_{\tilde{D} \times [0, \sigma_K]} g \, d\mathcal{H}^{n+1}\). We let

\[
U_{1,0} = \{ (q, s) \in \tilde{D} \times [0, \sigma_K) : -2\varepsilon\Lambda \leq \text{dist}_{\text{graph}(2\varepsilon\Lambda + t\chi_D)}((q, s)) \leq 0 \},
\]

\[
U_{0,-1} = \{ (q, s) \in \tilde{D} \times [0, \sigma_K) : 0 \leq \text{dist}_{\text{graph}(2\varepsilon\Lambda + t\chi_D)}((q, s)) \leq 2\varepsilon\Lambda \};
\]

the function \(\varpi_{D,t}\) decreases, on these two sets, respectively from \(1\) to \(0\) and from \(0\) to \(-1\) (as the distance increases respectively from \(-2\varepsilon\Lambda\) to \(0\) and from \(0\) to \(2\varepsilon\Lambda\)), so that \(\varpi_{D,t} - g\| \leq 1\) on \(U_{0,-1} \cup U_{1,0}\). On \(\tilde{D} \times [0, \sigma_K) \setminus (U_{0,-1} \cup U_{1,0})\), on the other hand, we have \(\varpi_{D,t} = g\). We then have

\[
\int_{\tilde{D} \times [0, \sigma_K]} g \, d\varpi_{D,t} \, d\mathcal{H}^{n+1} \geq 2\text{Vol}_\gamma(2\varepsilon\Lambda, t) - \int_{\tilde{D} \times [0, \sigma_K]} g \, d\mathcal{H}^{n+1} - \mathcal{H}^{n+1}(U_{1,0} \cup U_{0,-1}).
\]

Using the coarea formula (with respect to \(\text{dist}_{\text{graph}(2\varepsilon\Lambda + t\chi_D)}\) or with respect to \(\Pi_{t,\varepsilon}\), recalling \([38]\)), we obtain that \(\mathcal{H}^{n+1}(U_{1,0})\) and \(\mathcal{H}^{n+1}(U_{0,-1})\) are bounded above by \(2\varepsilon\Lambda \mathcal{H}^n(\text{graph}(2\varepsilon\Lambda + t\chi_D)_{\tilde{D}}) + O(\varepsilon |\log \varepsilon|)\). Moreover, from the area formula we obtain \(\mathcal{H}^n(\text{graph}(2\varepsilon\Lambda + t\chi_D)_{\tilde{D}}) \leq \mathcal{H}^n(\tilde{D})(1 + C_{\chi_D}(\omega, t, n))\). In conclusion, there exists \(\varepsilon_2 > 0\) such that for all \(\varepsilon \leq \varepsilon_2\) we have

\[
\int_{\tilde{D} \times [0, \sigma_K]} g \, d\varpi_{D,t} \, d\mathcal{H}^{n+1} \geq 2\text{Vol}_\gamma(2\varepsilon\Lambda, t) - \int_{\tilde{D} \times [0, \sigma_K]} g \, d\mathcal{H}^{n+1} - |O(\varepsilon |\log \varepsilon|)|.
\]

Putting this together with \([45]\) and \([42]\) we obtain that there exists \(\varepsilon_2 > 0\) such that for all \(\varepsilon \leq \varepsilon_2\)

\[
\frac{1}{2\sigma} \mathcal{F}_\varepsilon(\varpi_{D,t}) \leq J_\varepsilon(0, 0) + \int_{\tilde{D} \times [0, \sigma_K]} \frac{g}{2\sigma} \, d\mathcal{H}^{n+1} + \frac{T_B}{2} + O(\varepsilon |\log \varepsilon|) \quad \text{for all } t \in [0, t_0],
\]

\[
\frac{1}{2\sigma} \mathcal{F}_\varepsilon(\varpi_{D,t_0}) \leq J_\varepsilon(0, 0) + \int_{\tilde{D} \times [0, \sigma_K]} \frac{g}{2\sigma} \, d\mathcal{H}^{n+1} - T_B + O(\varepsilon |\log \varepsilon|). \quad \text{(46)}
\]

### 6.6 One-parameter deformation, bumping downwards

In this section \(\tilde{D}\) will be as in Section 6.5 and we will construct a deformation that continues the one provided in Section 6.5 to \(t \leq 0\). More precisely, we will construct, for each \(\varepsilon \in (0, \varepsilon_1]\) (for \(\varepsilon_1\) as in Section 6.5) a one-parameter deformation

\[t \in [-4\varepsilon\Lambda, 0] \rightarrow \varpi_{D,t} \in W^{1,2}(\tilde{D} \times [0, \omega])\].

At \(t = 0\) the resulting function will agree with \(G_{\tilde{D}}^0\) and therefore with the function defined by \([43]\) for \(t = 0\) (justifying the notation). We will then check that the functions \(\varpi_{D,t}\) (for \(t \in [-4\varepsilon\Lambda, 0]\)) pass to the quotient as \(W^{1,2}\) (actually \(W^{1,\infty}\)) functions on \(F(\tilde{D} \times [0, \omega])\) and that the resulting deformation \([-4\varepsilon\Lambda, 0] \rightarrow W^{1,2}(\tilde{D} \times [0, \omega])\) is continuous in \(t\).
Let \( \chi_{\tilde{D}} : \tilde{D} \to [0,1] \) be as in Section 6.5 and \( \Psi_t : \mathbb{R} \to \mathbb{R} \) as in (30). We define, for \( t \in [-4 \varepsilon, 0] \) and \( (q,s) \in \tilde{D} \times [0, \sigma_K) \),

\[
\varpi_{D,t}(q,s) = \Psi_{-t \chi_D(q)}(s).
\]

(47)

Since \( \chi_{\tilde{D}} \) is even on \( \tilde{D} \) by construction, the function \( \varpi_{D,t}(q,s) \) in (47) passes to the quotient in \( F \left( \tilde{D} \times [0, \sigma_K) \right) \). We will now check that it is in fact Lipschitz on \( F \left( \tilde{D} \times [0, \sigma_K) \right) \). Note that at \( t = 0 \) the function agrees with \( G_0 |_{F(\tilde{D} \times [0, \sigma_K])} \).

We only need to check the Lipschitz property locally around a point \( x \in D \), since the function is smooth in \( F \left( \tilde{D} \times [0, \sigma_K) \right) \). Let \( (y,a) \in B_p(x) \times (-\sigma_K, \sigma_K) \) denote Fermi coordinates centred at geodesic ball around \( x \) in \( D \). Then the expression of the function obtained by passing \( \varpi_{D,t} \) to the quotient is \( \Psi_{-t \chi_D(q)}(a) \), because \( \Psi_a \) is even for all \( s \geq 0 \). Moreover, \( \Psi_a \) is Lipschitz for all \( s \geq 0 \); this implies that the function \( \Psi_{-t \chi_D(q)}(a) \) is Lipschitz with respect to the product metric on \( B_p(x) \times (-\sigma_K, \sigma_K) \). The distortion factor between the Riemannian metric on \( N \) and this product metric is bounded by a geometric constant (depending only on the geometry of \( F(K \times [0, \sigma_K]) \)), hence the function \( \varpi_{D,t} \) passes to the quotient as a Lipschitz function. We also point out that there exists a neighbourhood of \( F \left( \tilde{D} \times [0, \sigma_K) \right) \) in which the function agrees with \( G_0^t \) for every \( t \in [-4 \varepsilon, 0] \).

Next we estimate the Allen–Cahn energy of \( \varpi_{D,t} \). Denote by \( \nabla_q \) the metric gradient in \( \tilde{D} \times [0, \sigma_K) \) projected onto the level set \( \{s = \text{cnst}\} \). Then

\[
\nabla_q \varpi_{D,t}(q,s) = -t \frac{d}{da} \Psi_a(s) \bigg|_{a = -t \chi_D(q)} \nabla_q \chi_{\tilde{D}}(q,s),
\]

where we think momentarily of \( \chi_{\tilde{D}}(q,s) = \chi_D(q,s) \) as a function on \( \tilde{D} \times [0, \sigma_K) \) that only depends on the variable \( q \). Recalling (30), we have \( \frac{d}{ds} \Psi_a(s) = |\Psi'(a + |s|)| \leq \frac{3}{2} \). Moreover we ensured \( |\nabla \chi_{\tilde{D}}| \leq \frac{3}{2} \), as a function on \( \tilde{D} \), and therefore \( |\nabla_q \chi_{\tilde{D}}(q,s)| \leq \frac{9 C_K}{R} \), for a constant \( C_K \) that depends only on the Riemannian metric on \( K \times [0, \sigma_K) \). Therefore

\[
\frac{\varepsilon |\nabla_q \varpi_{D,t}(q,s)|^2}{\varepsilon R^2} \leq \frac{9 C_K^2 \varepsilon}{R^2} \leq \frac{9 \cdot 16 C_K^2 \varepsilon}{R^2} \varepsilon A^2 = C \varepsilon \log \varepsilon^2,
\]

for a constant \( C > 0 \) that depends only on the choices of \( K \) and \( D \) (in particular, it does not depend on \( \varepsilon \)). Recalling that (for the Riemannian metric) the unit vectors \( \frac{d}{ds} \) are orthogonal to the level sets \( \{s = \text{cnst}\} \) we can write \( |\nabla q \varpi_{D,t}|^2 = |\nabla_q \varpi_{D,t}|^2 + \frac{1}{\varepsilon} |\varpi_{D,t}|^2 \). We then compute the Allen–Cahn energy of \( \varpi_{D,t} \) by using the coarea formula with respect to \( \Pi_K \), recalling that \( \nabla_q \varpi_{D,t} = 0 \) on \( \tilde{B} \times [0, \sigma_K) \) (since \( \chi_{\tilde{D}} = 1 \) on \( \tilde{B} \)):

\[
\int_{\tilde{D} \times (0, \sigma_K)} \frac{\varepsilon |\nabla \varpi_{D,t}|^2}{\varepsilon R^2} + \frac{W(\varpi_{D,t})}{\varepsilon} = \int_{\tilde{B}} \int_0^\sigma K \frac{1}{|\Pi_K|} (q,s) \left( \frac{d}{ds} \Psi_{|t|}^\sigma(s) \right)^2 + \frac{W(\Psi_{|t|}^\sigma(s))}{\varepsilon} ds \right) dq + \int_{\tilde{D} \setminus \tilde{B}} \int_0^\sigma K \frac{1}{|\Pi_K|} \left( \frac{d}{ds} \Psi_{-t \chi_D(q)}^\sigma(s) \right)^2 + \frac{W(\Psi_{-t \chi_D(q)}^\sigma(s))}{\varepsilon} ds \right) dq \]
We consider the right-hand-side. The first term bounded by \((1+k_K \in \Lambda)\mathcal{H}^n(\tilde{D})(2\sigma)\), by the observation following \((30)\) and by the bounds on \(\Pi_K\) in \((30)\). For the same reason, the second term is bounded by \((1+k_K \in \Lambda)\mathcal{H}^n(\tilde{D} \setminus B)(2\sigma)\). In view of \((45)\) the third term is \(O(\varepsilon |\log \varepsilon|)\) (in fact, it is \(O(\varepsilon^2 |\log \varepsilon|^2)\), by noticing that the integrand vanishes on \((\tilde{D} \setminus B) \times (4 \varepsilon \Lambda, \sigma_K))\). We thus have that there exists \(\varepsilon_2 \in (0, \varepsilon_1]\) such that, for all \(\varepsilon \in \varepsilon_2\), the following bound holds independently of \(\varepsilon\):

\[
E(\omega_{D,t}) \leq (2\sigma)2\mathcal{H}^n(D) + O(\varepsilon |\log \varepsilon|). \tag{50}
\]

One can easily obtain a finer bound by estimating the first term more precisely. We only need to do so for \(t = -4 \varepsilon \Lambda\); in this case the first term vanishes, therefore

\[
E(\omega_{D,-4 \varepsilon \Lambda}) \leq (2\sigma)2\mathcal{H}^n(D \setminus B) + O(\varepsilon |\log \varepsilon|). \tag{51}
\]

To conclude this section, we estimate the energy \(F_{\varepsilon}\) of the functions \(\omega_{D,t}\). A very rough estimate will suffice for our purposes: since \(\omega_{D,\varepsilon} \geq -1\) we have

\[
\int_{\tilde{D} \times (0, \sigma_K)} \omega_{D,t} \, g \, d\mathcal{H}^{n+1} \geq -\int_{\tilde{D} \times (0, \sigma_K)} g \, d\mathcal{H}^{n+1};
\]

together with \((50)\) and \((51)\) this gives, for all \(\varepsilon \leq \varepsilon_2\),

\[
\frac{1}{2\sigma} F_{\varepsilon}(\omega_{D,-4 \varepsilon \Lambda}) \leq 2\mathcal{H}^n(D) - 2\mathcal{H}^n(B) + \int_{\tilde{D} \times (0, \sigma_K)} g \, d\mathcal{H}^{n+1} + O(\varepsilon |\log \varepsilon|), \tag{52}
\]

\[
\frac{1}{2\sigma} F_{\varepsilon}(\omega_{D,t}) \leq 2\mathcal{H}^n(D) + \int_{\tilde{D} \times (0, \sigma_K)} g \, d\mathcal{H}^{n+1} + O(\varepsilon |\log \varepsilon|) \text{ for all } t \in [-4 \varepsilon \Lambda, 0].
\]

6.7 Avoiding the peak

Given a geodesic ball \(D \subset M\) (and denoting by \(\tilde{D} \subset K\) its double cover), in Sections 6.5 and 6.6 we produced, for \(t_0 > 0\) (depending on \(D\)) and for every sufficiently small \(\varepsilon \leq \varepsilon_2\), a continuous family \(t \in [-4 \varepsilon \Lambda, t_0] \to W^{1,2}\left(\tilde{D} \times (0, \sigma_K)\right)\).

This family descends to a continuous one \(t \in [-4 \varepsilon \Lambda, t_0] \to \omega_{D,t} \in W^{1,2}\left(\tilde{D} \times (0, \sigma_K)\right)\); more precisely, the functions in the image of this curve are in \(W^{1,\infty}\left(\tilde{D} \times (0, \sigma_K)\right)\).

At \(t = 0\) we have that \(\omega_{D,0}\) agrees with \(G_0^\varepsilon\). In other words, we have a continuous two-sided deformation of \(G_0^\varepsilon\) in \(F(\tilde{D} \times (0, \sigma_K))\). By \((46)\) and \((52)\) the energy \(\frac{1}{2\sigma} F_{\varepsilon}\) stays below

\[
J_{\frac{1}{2\sigma}}(0,0) + \int_{\tilde{D} \times (0, \sigma_K)} g \, d\mathcal{H}^{n+1} + \frac{\tau_B}{2} + O(\varepsilon |\log \varepsilon|)
\]

for all \(t \in [-4 \varepsilon \Lambda, t_0]\) and, moreover, at the endpoints \(t = -4 \varepsilon \Lambda\) and \(t = t_0\) we have that the energy \(\frac{1}{2\sigma} F_{\varepsilon}\) is at most

\[
J_{\frac{1}{2\sigma}}(0,0) - \min\{2\mathcal{H}^n(B), \tau_B\} + \int_{\tilde{D} \times (0, \sigma_K)} g \, d\mathcal{H}^{n+1} + O(\varepsilon |\log \varepsilon|).
\]

In this section we choose two distinct geodesic balls and, around each of them, we produce a two-sided deformation of \(G_0^\varepsilon\) as we did above. By suitably combining them, and extending to \(N\), we will produce a continuous curve into \(W^{1,2}(\mathcal{N})\) (the functions are moreover Lipschitz on \(\mathcal{N}\) with the key property that the energy \(\frac{1}{2\sigma} F_{\varepsilon}\) stays always below the value \(2\mathcal{H}^n(M) - \varepsilon + \int_{\mathcal{N}} g \, d\mathcal{H}^{n+1} + O(\varepsilon |\log \varepsilon|)\), for
some \( \epsilon > 0 \) that only depends on the geometric properties of \( \varepsilon \) and on the choice of the two geodesic balls; in particular, the energy bound holds independently of \( \varepsilon \), for all sufficiently small \( \varepsilon \).

Let \( D_1 \subset \subset M \) and \( D_2 \subset \subset M \) be the geodesic balls chosen in Section 6.4. We denote respectively by \( B_1 \) and \( B_2 \) the concentric geodesic balls with half the radius. The balls are chosen so that \( \mathcal{H}^n(B_1) = \mathcal{H}^n(B_2) \). We denote the respective \( \varepsilon \) double covers by \( \tilde{D}_1, \tilde{D}_2, \tilde{B}_1, \tilde{B}_2 \) and we assumed that \( \tilde{D}_j \subset \subset K \) for \( j \in \{1, 2\} \). For each \( D_j \) we can repeat the construction in Sections 6.5, 6.6. We let \( \tau \) denote the smallest of the two \( \tau_1 \) identified for \( j \in \{1, 2\} \), and \( \tau_0 \) the final time of the deformation identified for \( D_j \) (which was denoted by \( t_0 \) for \( D \)). We thus obtain, for each \( \varepsilon \leq \varepsilon_1 \) two continuous curves into \( W^{1,2}(\tilde{F} \times [0, \sigma_K]) \), respectively for \( j = 1, 2 \):

\[
t \in [-4 \varepsilon \Lambda, \tau_0^{(1)}] \rightarrow \varpi_{D_{1,t}}, \quad t \in [-4 \varepsilon \Lambda, \tau_0^{(2)}] \rightarrow \varpi_{D_{2,t}}.
\]

We define, for \( t \in [-4 \varepsilon \Lambda - \tau_0^{(1)}, 4 \varepsilon \Lambda + \tau_0^{(2)}] \), the following continuous one-parameter family of \( W^{1,2} \) functions on \( F(\tilde{D}_1 \times [0, \sigma_K]) \cup F(\tilde{D}_2 \times [0, \sigma_K]) \):

\[
\gamma_t = \begin{cases}
\varpi_{D_1,t+\tau_0^{(1)}} & \text{on } F(\tilde{D}_1 \times [0, \sigma_K]) \text{ for } t \in [-4 \varepsilon \Lambda - \tau_0^{(1)}, -\tau_0^{(1)}] \\
\varpi_{D_2,-4 \varepsilon \Lambda} & \text{on } F(\tilde{D}_2 \times [0, \sigma_K]) \text{ for } t \in [-4 \varepsilon \Lambda - \tau_0^{(1)}, -\tau_0^{(1)}] \\
\varpi_{D_1,t+\tau_0^{(1)}} & \text{on } F(\tilde{D}_1 \times [0, \sigma_K]) \text{ for } t \in [0, \tau_0^{(1)}] \\
\varpi_{D_2,-4 \varepsilon \Lambda} & \text{on } F(\tilde{D}_2 \times [0, \sigma_K]) \text{ for } t \in [0, \tau_0^{(1)}] \\
\varpi_{D_1,t_0^{(1)}} & \text{on } F(\tilde{D}_1 \times [0, \sigma_K]) \text{ for } t \in [4 \varepsilon \Lambda, 4 \varepsilon \Lambda + t_0^{(2)}] \\
\varpi_{D_2,-4 \varepsilon \Lambda} & \text{on } F(\tilde{D}_2 \times [0, \sigma_K]) \text{ for } t \in [4 \varepsilon \Lambda, 4 \varepsilon \Lambda + t_0^{(2)}] \\
\end{cases}
\]

The idea is that for every \( t \), in one of the two subdomains \( F(\tilde{D}_1 \times [0, \sigma_K]) \cup F(\tilde{D}_2 \times [0, \sigma_K]) \), the function agrees with one of the endpoints of the curve \( \varpi_{D_{1,t}} \). Recall that \( \tau_{B_j} < \mathcal{H}^n(B_j) \) by the choice made in Section 6.5, moreover we ensured \( \mathcal{H}^n(B_1) = \mathcal{H}^n(B_2) \). Thanks to (54) and (52) we can estimate, for \( \varepsilon \leq \varepsilon_2 \), the energy \( \frac{1}{\varepsilon^2} \int F(\gamma_t) \) on the set \( F(\tilde{D}_1 \times [0, \sigma_K]) \cup F(\tilde{D}_1 \times [0, \sigma_K]) \) from above with the quantity

\[
2\mathcal{H}^n(D_1) + 2\mathcal{H}^n(D_2) + \int_{\tilde{D}_1 \times [0, \sigma_K]} \frac{\partial}{2} d\mathcal{H}^n+1 + \int_{\tilde{D}_2 \times [0, \sigma_K]} \frac{\partial}{2} d\mathcal{H}^n+1

- \min \left\{ \mathcal{H}^n(B_j), \frac{\tau_{B_1}}{2}, \frac{\tau_{B_2}}{2} \right\} + O(\varepsilon | \log \varepsilon |).
\]

Each \( \gamma_t \) is in fact, on \( F(\tilde{D}_1 \times [0, \sigma_K]) \cup F(\tilde{D}_1 \times [0, \sigma_K]) \), a Lipschitz function. Our next aim is to extend \( \gamma_t \), for each \( t \), to a Lipschitz function on \( N \), obtaining a continuous curve from \([-4 \varepsilon \Lambda - \tau_0^{(1)}, 4 \varepsilon \Lambda + t_0^{(2)}] \) into \( W^{1,2}(N) \). In order to do that, we recall that there exists a neighbourhood of \( F(\tilde{D}_1 \times [0, \sigma_K]) \cup F(\tilde{D}_1 \times [0, \sigma_K]) \) in which \( \gamma_t \) agrees with \( G_0^t \) for every \( t \). This implies that we

---

[Continuity is immediate from the continuity of each \( \varpi_{D_{1,t}} \) in their respective (disjoint) domains.]
Figure 3: Schematic picture representing the $\varepsilon \to 0^+$ limit of the path $\gamma_t$ defined above. The order (i) $\to$ (ii) $\to$ (iii) $\to$ (iv) $\to$ (v) $\to$ (vi) represents increasing $t$. The deformation (i) $\to$ (ii), for instance, represents the $\varepsilon \to 0^+$ limit of the top two lines of the piecewise definition of $\gamma_t$; likewise, (ii) $\to$ (iii) corresponds to the third and fourth lines of the definition, etc. The deformations (i) $\to$ (ii), and (iii) $\to$ (iv) are instantaneous “jumps” in the $\varepsilon \to 0^+$ limit (and therefore discontinuous), but the corresponding deformation $t \mapsto \gamma_t$ are continuous in $t$ for fixed $\varepsilon = \varepsilon_j > 0$. The evolution $t \mapsto \gamma_t$ (for fixed $\varepsilon = \varepsilon_j > 0$) corresponding to “(vi) onwards” in the picture is carried out by the negative gradient flow of $F_{\varepsilon_j}$, and converges to the stable solution $v_{\varepsilon_j}$, as described in Section 6.9 below.

can define a Lipschitz function on $N$ that extends $\gamma_t$ by setting

$$
\gamma_t(x) = \begin{cases} 
\gamma_t(x) & \text{for } x \in F \left( \tilde{D}_1 \times [0, \sigma_K] \right) \cup F \left( \tilde{D}_2 \times [0, \sigma_K] \right), \\
G_0^\varepsilon(x) & \text{for } x \in N \setminus \left( F \left( \tilde{D}_1 \times [0, \sigma_K] \right) \cup F \left( \tilde{D}_2 \times [0, \sigma_K] \right) \right).
\end{cases}
$$

(55)

The continuity of $t \in [-4 \varepsilon \Lambda - t_0^{(1)}, 4 \varepsilon \Lambda + t_0^{(2)}]$ $\to \gamma_t \in W^{1,2}(N)$ is also immediate from the continuity of $\gamma_t$.

We next estimate the Allen–Cahn energy of $\gamma_t$. For that purpose, we first give a lower bound on the Allen–Cahn energy of $G_0^\varepsilon$ on the sets $F \left( \tilde{D}_1 \times [0, \sigma_K] \right)$ and $F \left( \tilde{D}_2 \times [0, \sigma_K] \right)$. We follow the argument in Section 6.1 that led to (33), this time using the bound $\text{Ric}_N \leq C$; this is true for some $C > 0$ because $N$ is compact. Integrating Riccati’s equation we get $H_t \leq \sqrt{C} \tan(\sqrt{C}t)$, where $H_t$ denotes the scalar mean curvature of the level set of $d_M$ at distance $t$ (computed with respect to the unit normal that points away from $M$, equivalently with respect to $\partial_s$ in $K \times [0, \sigma_K]$). The ODE for the area element $\theta_s$ then leads to the following bound for the evolution of $\theta_s$ along a geodesic orthogonal to $\tilde{M}$: $\theta_s \geq \theta_0 \left( 1 - \frac{C^2}{2} \right)$. 

53
Therefore
$$\mathcal{H}^{n} \left( \tilde{D}_{j} \times \{s\} \right) \geq \left( 1 - \frac{C_{n}}{2} \right) \mathcal{H}^{n} \left( \tilde{D}_{j} \times \{0\} \right).$$

Then the coarea formula, used for the function $d_{\mathcal{M}}$ gives for $j \in \{1, 2\}$ (similarly to Section 6.2):

$$\int_{F(\tilde{D}_{j} \times [0, \sigma_{K}])} \varepsilon \frac{|\nabla G_{0}|^2}{2} + W(G_{0}) \, ds = \int_{0}^{\sigma_{K}} \left( \int_{\tilde{D}_{j} \times \{s\}} \varepsilon \frac{|\nabla G_{0}|^2}{2} + W(G_{0}) \, ds \right) \, ds = \int_{-2 \varepsilon \Lambda}^{2 \varepsilon \Lambda} \left( \frac{\varepsilon}{2} \left( \frac{H_{\varepsilon}^r(s)}{2} + W(H_{\varepsilon}^r(s)) \right) \right) \, ds \geq 2 \left( 1 - \frac{C(4 \varepsilon \Lambda)^{2}}{2} \right) \mathcal{H}^{n}(D_{j}) \left( \int_{-2 \varepsilon \Lambda}^{2 \varepsilon \Lambda} \frac{\varepsilon}{2} \left( \frac{H_{\varepsilon}^r(s)}{2} + W(H_{\varepsilon}^r(s)) \right) \right) \geq 2(2\sigma)\mathcal{H}^{n}(D_{j}) - |O(\varepsilon| \log \varepsilon)|.
$$

The last inequality holds for $\varepsilon \leq \varepsilon_2$ for a suitable choice of $\varepsilon_2$. The bound for $E_{\varepsilon}(G_{0})$ obtained in Section 6.2, together with (56) gives an upper bound for $G_{0}$ on the set $N \setminus \left( F \left( \tilde{D}_{1} \times [0, \sigma_{K}] \right) \cup F \left( \tilde{D}_{2} \times [0, \sigma_{K}] \right) \right)$: the Allen–Cahn energy of $G_{0}$ on this set is at most

$$2(2\sigma)\mathcal{H}^{n}(M) - 2(2\sigma)\mathcal{H}^{n}(D_{1}) - 2(2\sigma)\mathcal{H}^{n}(D_{2}) + O(\varepsilon| \log \varepsilon|).$$

In order to estimate $F_{\varepsilon}$ on this same set we note that $G_{0} \geq -1$ and therefore

$$\int_{N \setminus \left( F \left( \tilde{D}_{1} \times [0, \sigma_{K}] \right) \cup F \left( \tilde{D}_{2} \times [0, \sigma_{K}] \right) \right)} G_{0} g \, d\mathcal{H}^{n+1} \geq -\int_{N \setminus \left( F \left( \tilde{D}_{1} \times [0, \sigma_{K}] \right) \cup F \left( \tilde{D}_{2} \times [0, \sigma_{K}] \right) \right)} g \, d\mathcal{H}^{n+1}.$$

Recalling (55) and the bound (54), the last two estimates imply that there exists $\varepsilon_2 > 0$ such that for all $\varepsilon \leq \varepsilon_2$ the family $t \in [-4 \varepsilon \Lambda - t_{0}^{(1)}, 4 \varepsilon \Lambda + t_{0}^{(2)}] \rightarrow \gamma_{t} \in W^{1, 2}(N)$ (recall that $\gamma_{t} = \gamma_{t}^{r}$ is constructed for each fixed $\varepsilon$ in the chosen range) satisfies

$$\frac{1}{2\sigma} F_{\varepsilon}(\gamma_{t}) \leq 2\mathcal{H}^{n}(M) + \int_{N} \frac{g}{2\sigma} \, d\mathcal{H}^{n+1} - \min \left\{ \mathcal{H}^{n}(B_{j}), \frac{T_{B_{j}}}{2}, \frac{T_{B_{2}}}{2} \right\} + O(\varepsilon| \log \varepsilon|)$$

for all $t \in [-4 \varepsilon \Lambda - t_{0}^{(1)}, 4 \varepsilon \Lambda + t_{0}^{(2)}]$, as claimed in the beginning of this section.

### 6.8 Path to $a_{\varepsilon}$

In this section we exhibit, for each sufficiently small $\varepsilon$, a continuous path in $W^{1, 2}(N)$ that connects the function $a_{\varepsilon}$ (the first endpoint of the class of admissible paths, see Section 6.1) to the function $\gamma_{-4 \varepsilon \Lambda - t_{0}^{(1)}}$ obtained in Section 6.7 ensuring that $F_{\varepsilon}$ along this path remains bounded by the right-hand-side of (57). To ease notation, in this section we will denote simply by $f_{0} : N \rightarrow \mathbb{R}$ the function $\gamma_{-4 \varepsilon \Lambda - t_{0}^{(1)}}$. We rewrite the definition of $f_{0}$ as follows, recalling (47), (53) and (55):

$$f_{0}(x) = \begin{cases} \Psi_{\varepsilon \Lambda \chi_{\tilde{D}_{j}}}(0)(s) & \text{for } x = F(q, s), (q, s) \in \tilde{D}_{j} \times [0, \sigma_{K}], j \in \{1, 2\} \\ \Psi_{0}(d_{\mathcal{M}}(x)) & \text{for } x \in N \setminus \left( F(\tilde{D}_{1} \times [0, \sigma_{K}]) \cup F(\tilde{D}_{2} \times [0, \sigma_{K}]) \right) \end{cases}. \quad (58)$$
We define, for $r \in [0, 4 \varepsilon \Lambda]$

$$f_r(x) = \begin{cases} 
\Psi_{\varepsilon \Lambda \chi_{D_j}(q) + r}(s) & \text{for } x = F(q, s), (q, s) \in \tilde{D}_j \times [0, \sigma_K), \ j \in \{1, 2\} \\
\Psi_r(d_{\tilde{\Sigma}}(x)) & \text{for } x \in N \setminus \left(F(\tilde{D}_1 \times [0, \sigma_K)) \cup F(\tilde{D}_2 \times [0, \sigma_K])\right) 
\end{cases}$$

(59)

The first line is well-defined thanks to the fact that $\chi_{D_j}$ is even on $\tilde{M}$; note that for $r = 0$ the definition in (59) agrees with the one in (58), justifying the notation. Since $\chi_{D_j} \in C^\infty_c(\tilde{D}_j)$ for $j \in \{1, 2\}$, the definition in the first line agrees with $\Psi_r(d_{\tilde{\Sigma}}(x))$ when $(q, s)$ is in a neighbourhood of $\partial \tilde{D}_j \times [0, \sigma_K)$. Thanks to the fact that $\Psi_r$ is Lipschitz, one can check that each $f_r$ is a Lipschitz function on $N$. Moreover, the mapping $r \in [0, 4 \varepsilon \Lambda] \rightarrow f_r \in W^{1,2}(N)$ is continuous. Note that $f_{4 \varepsilon \Lambda} \equiv \setminus 1$.

To compute $E_x(f_r)$ we employ the coarea formula, with respect to the function $\Pi_K$ on $\bigcup_{j=1}^2 F(\tilde{D}_j \times [0, 4 \varepsilon \Lambda])$ and with respect to $d_{\tilde{\Sigma}}$ on the complementary domain in $\{d_{\tilde{\Sigma}} < 4 \varepsilon \Lambda\}$, since $f_r = -1$ there. The key observation is that the one-dimensional profile that appears for $f_r$ (with $r > 0$) in the normal bundle to $M$ carries less energy than the one that appears for $f_0$. Indeed, (58) and (59) show that the profile $\Psi_0$ for $f_0$ is replaced by $\Psi_r$ for $f_r$, and the profile $\Psi_{\varepsilon \Lambda \chi_{D_j}(q)}$ for $f_0$ is replaced by $\Psi_{\varepsilon \Lambda \chi_{D_j}(q) + r}$ for $f_r$ (recall the observation that follows (59)). Moreover, we note that $f_0 \geq -1$ and therefore $\int_N f_r g d\mathcal{H}^{n+1} \geq -\int_N g d\mathcal{H}^{n+1}$. Therefore we obtain that there exists $\varepsilon > 2$ such that for all $\varepsilon \leq \varepsilon_2$ the following bound holds

$$\frac{1}{2\sigma} F_r(f_r) \leq 2 \mathcal{H}^n(M) - 2 \mathcal{H}^n(B_1) - 2 \mathcal{H}^n(B_2) + \int_N g d\mathcal{H}^{n+1} + O(\varepsilon |\log \varepsilon|)$$

(60)

for all $r \in [0, 4 \varepsilon \Lambda]$. (It should be kept in mind that $f_r = f_r^\varepsilon$ is built for each fixed $\varepsilon$ in the chosen range.)

We next connect $f_{4 \varepsilon \Lambda} \equiv -1$ to $a_\varepsilon$, continuously in $W^{1,2}$. For this it suffices to recall that $a_\varepsilon$ was defined as limit of the negative $F_\varepsilon$-gradient flow with initial condition $-1$; we denote by $[0, T_\varepsilon]$ the time interval on which this flow is defined. The same flow (translated by $4 \varepsilon \Lambda$) provides the continuation of $f_r$ to an interval $[0, 4 \varepsilon \Lambda + T_\varepsilon]$ with the property that the family $r \in [0, 4 \varepsilon \Lambda + T_\varepsilon] \rightarrow f_r \in W^{1,2}(N)$ is a continuous path that satisfies the bound in (60) for all $r$ and such that $f_0 = \gamma_{-4 \varepsilon \Lambda - t_0(1)}$ and $f_{4 \varepsilon \Lambda + T_\varepsilon} = a_\varepsilon$.

### 6.9 Flow to a stable solution

In this section we produce a continuous path in $W^{1,2}(N)$ that starts at the function $\gamma_{4 \varepsilon \Lambda + t_0(2)}$ obtained in Section 6.7 and ends at a stable solution $v_\varepsilon$ of $F_\varepsilon = 0$.

To ease notation, we denote by $h$ the function $\gamma_{4 \varepsilon \Lambda + t_0(2)}$. Recalling (44) and (55), $h$ is given by

$$h(x) = \begin{cases} 
\mathbb{P}(-\text{dist}_K 2 \varepsilon \Lambda_{t_0(n)} \chi_{\tilde{D}_j}(x)) & \text{for } x = F(q, s), (q, s) \in \tilde{D}_j \times [0, \sigma_K), \ j \in \{1, 2\}, \\
\Psi_0(d_{\tilde{\Sigma}}(x)) & \text{for } x \in N \setminus (F(\tilde{D}_1 \times [0, \sigma_K)) \cup F(\tilde{D}_2 \times [0, \sigma_K])). 
\end{cases}$$

(61)

*Although irrelevant for our construction, it is easy to check that $T_n$ tends to 0 as $\varepsilon \rightarrow 0$.  

55
The first line is well-defined thanks to the fact that $\chi_{\bar{D}_j}$ is even on $\bar{M}$. Since $\chi_{\bar{D}_j} \in C_0^{\infty}(\bar{D}_j)$ for $j \in \{1, 2\}$, the definition in the first line agrees with $\Psi_0(d_{\mathcal{M}}(x))$ (and thus with $C_0^\circ$) when $(q, s)$ is in a neighbourhood of $\partial \bar{D}_j \times [0, \sigma_K]$. In view of this, we will compute separately the first variation of $h$ with respect to $\mathcal{E}_{\epsilon}$ in the domains that appear in the first and second line of (61). For the second line, it suffices to recall (47). For the first, we need to estimate, for any sufficiently small $\varepsilon$, the mean curvature of the embedded hypersurfaces given by level sets of the distance to $\tilde{F}$ graph(2 $\varepsilon$ $\Lambda$ + $t_0^{(1)}$ $\chi_{\bar{D}_1} + t_0^{(2)}$ $\chi_{\bar{D}_2})$, where the graph is intended over Int($K$). Recall Remark 6.2 and the fact that $M$ is minimal. Then, choosing $\varepsilon_2$ sufficiently small, we can ensure that for all $\varepsilon \leq \varepsilon_2$ the mean curvature of the embedded hypersurfaces $\{\text{dist}_{K_{\Lambda \epsilon \sigma d}^\epsilon, N_{\bar{D}_j}} = d\}$ for $d \in [-2 \varepsilon \Lambda, 2 \varepsilon \Lambda]$ are, in absolute value, smaller than $\frac{\min N g}{3}$. With a computation similar to (36), with the distance $\text{dist}_{K_{\Lambda \epsilon \sigma d}^\epsilon, N_{\bar{D}_j}}$ in place of $d_{\mathcal{M}}$, we obtain that, on the set $F(\bar{D}_1 \times [0, \sigma_K)) \cup F(\bar{D}_2 \times [0, \sigma_K])$, and for all $\varepsilon \leq \varepsilon_2$,

$$
\mathcal{E}_\epsilon'(h) \geq -\frac{\min N g}{3} - |O(\varepsilon | \log \varepsilon)|.
$$

Together with (47), this implies that $\mathcal{E}_\epsilon'(h) \geq \frac{2 \min N g}{3} - |O(\varepsilon | \log \varepsilon)|$ on $N$ for all $\varepsilon < \varepsilon_1$ and therefore there exists $\varepsilon_2 > 0$ such that for all $\varepsilon \leq \varepsilon_2$

$$
\mathcal{F}_\epsilon'(h) \geq \frac{\min N g}{2} > 0 \text{ on } N. \quad (62)
$$

(As in (37), this is understood to be an inequality between Radon measures.)

We will now produce the path mentioned at the beginning of this section, by means of a negative gradient flow (with respect to the functional $\mathcal{F}_\epsilon$). If $n \leq 6$, then the function $h$ is smooth (because so is $d_{\mathcal{M}}$ in a tubular neighbourhood of $M = \bar{M}$) and we can use it as initial condition for the flow. For the general case, we first need a smoothing of $h$ with the key property that it preserves the positivity of the first variation (62). We refer to [Bel, Appendix A] for the smoothing operation by means of mollifiers $\rho_\delta : N \times N \to \mathbb{R}$. For each sufficiently small $\varepsilon$ there exists $\delta > 0$ such that the convolution $h_\delta = h * \rho_\delta$ is a smooth function on $N$ and satisfies

$$
-\mathcal{F}_\epsilon'(h_\delta) = \varepsilon \Delta h_\delta - \frac{\mathcal{F}_\epsilon'(h_\delta)}{\varepsilon} + g > 0 \quad (63)
$$

(see [Bel] Lemma A.2), this uses that $h$ is Lipschitz). Moreover, $r \in (0, \delta] \to r_h = h * \rho_\delta$ is continuous in $W^{1,2}(N)$ and extends by continuity at $r = 0$ with $h_0 = h$ (see [Bel] Lemma A.1). Still by continuity in $r$ of the convolution (Bel Lemma A.1), upon choosing $\delta$ sufficiently small we can also ensure that a bound of the form (57) continues to hold for $h_r$, i.e.

$$
\frac{1}{2 \sigma} \mathcal{F}_\epsilon(h_r) \leq 2H^n(M) + \int_N \frac{g}{2 \sigma} dH^{n+1} - \frac{2}{3} \varepsilon + O(\varepsilon | \log \varepsilon |) \quad (64)
$$

for all $r \in [0, \delta]$.

We need to ensure two more conditions on $h_\delta$, for which we may have to make $\delta$ still smaller (the choice of $\delta$ is allowed to depend on $\varepsilon$). Recall that by construction $h = 1$ on the fixed non-empty open set $F\left(\{(q, s) : q \in \bar{D}_1 \cup \bar{D}_2, 0 \leq s < t_0^{(1)}(\chi_{\bar{D}_1} + t_0^{(2)}(\chi_{\bar{D}_2})\right)$. By continuity, if $\delta$ is chosen sufficiently small, then $h * \rho_\delta > 3/4$ on the same open set. By construction, $h \leq 1$. Recall that the stable solution $b_\delta$ to $\mathcal{F}_\epsilon' = 0$ lies (strictly) above 1. Therefore, with a suitable small choice of $\delta$, $h * \rho_\delta < b_\delta$. 56
We will use the function $h_\epsilon$ just identified as initial condition for the negative gradient flow (with respect to $\mathcal{F}_\epsilon$), defined for $t \geq \delta$ by

$$
\begin{cases}
\varepsilon \partial_t h_t = \varepsilon \Delta h_t - \frac{W'(h_t)}{\varepsilon} + g

h_t|_{t=0} = h \ast \rho_\delta = h_\delta
\end{cases}
$$

(65)

This flow is well-defined and smooth for all times by standard semi-linear parabolic theory. Moreover, it is mean-convex with respect to $\mathcal{F}_\epsilon$, i.e.

$$
\mathcal{F}_\epsilon'(h_t) > 0 \text{ for all } t \in [\delta, \infty).
$$

To see this, notice that $F_t = \varepsilon \Delta h_t - \frac{W'(h_t)}{\varepsilon} + g$ is smooth on $N$ for all $t \geq \delta$ and $F_\delta > 0$ by (63). Since $h_t$ solves (65), then the PDE $\partial_t U_t = \Delta U_t - \frac{W'(h_t)}{\varepsilon} U_t$ is solved by $U_t = F_t$. Another solution is given by $U_t \equiv 0$. By the parabolic maximum principle, the condition $F_t > 0$ holds for all $t \geq \delta$, since it holds for $t = \delta$.

The mean-convexity guarantees, in a first instance, that the functions $h_t : N \to \mathbb{R}$ are increasing in $t$ and the limit $h_\infty$ for $t \to \infty$ is well-defined. The function $h_\infty$ is a solution of the elliptic PDE $\mathcal{F}_\epsilon'(h_\infty) = 0$; the mean-convexity further gives the following.

**Lemma 6.1.** The function $h_\infty = \lim_{t \to \infty} h_t$ is a stable solution of $\mathcal{F}_\epsilon' = 0$. Moreover, there exists a fixed non-empty open set (independent of $\epsilon$) that is contained in $\{h_\infty > \frac{3}{2}\}$.

**Proof.** To prove stability, we recall that the second variation of $\mathcal{F}_\epsilon$ at $h_\infty$ is given by the quadratic form $Q(\phi, \phi) = \int_N \varepsilon |\nabla \phi|^2 - \frac{W'(\theta_\infty)\phi^2}{\varepsilon}$ and that the associated Jacobi operator is $-\varepsilon \Delta \phi + \frac{W'(\theta_\infty)}{\varepsilon} \phi$. Letting $\rho_1$ denote the first eigenfunction and $\lambda_1$ the associated eigenvalue, if $\lambda_1 < 0$ we can find $s > 0$ sufficiently small so that $\mathcal{F}_\epsilon'(h_\infty - sp_1) < 0$ on $N$. Since $h_\delta < h_\infty$, we can also ensure that $h_\delta < h_\infty - sp_1$.

We then let the flow evolve from $h_\delta$ until the first time $T$ such that $h_T = h_\infty - sp_1$ at a point $x \in N$. At this point we must then have $\Delta h_T \leq \Delta (h_\infty - sp_1)$ and $W'(h_T) = W'(h_\infty - sp_1)$, therefore $\mathcal{F}_\epsilon'(h_\infty - sp_1) \geq \mathcal{F}_\epsilon'(h_\delta) > 0$ at $x$. This contradicts $\lambda_1 < 0$.

To prove the second statement, we recall that $h_\delta > 3/4$ on the (fixed) non-empty open set $F = \{(q, s) : q \in \tilde{D}_1 \cup \tilde{D}_2, 0 \leq s < a_0^{(1)} \chi_{D_1} + a_0^{(2)} \chi_{D_2}\}$. The condition that $h_t$ increases along the flow guarantees the conclusion. \qed

**Remark 6.5.** As $\mathcal{F}_\epsilon$ decreases along the flow, (64) holds for all $r \geq 0$.

**Remark 6.6.** By construction $h_\delta < b_\epsilon$, so that $b_\epsilon$ acts as a barrier for the flow $h_t$. In particular, $h_\infty \leq b_\epsilon$.

### 6.10 Concluding argument for the proof of Proposition 6.1

By reversing the path $f_\epsilon$ in Section 6.8 and composing it with the one produced in Section 6.7 and the one in Section 6.9, we obtain a continuous path in $W^{1,2}(N)$ that starts at $a_\epsilon$ and ends at a stable critical point $h_\infty$ of $\mathcal{F}_\epsilon$. This path can be produced for all $\epsilon$ sufficiently small. Moreover, from (67), (60), (64) and Remark 6.5 there exists $\varepsilon_2 > 0$ such that, for all $\epsilon < \varepsilon_2$ we have, all along this path, an upper bound of the type $\frac{1}{2\sigma} \mathcal{F}_\epsilon \leq 2\mathcal{H}''(M) + \int_N \frac{g}{2\sigma} d\mathcal{H}^{n+1} - \frac{\varepsilon}{2} + O(\varepsilon |\log \varepsilon|)$, with $\zeta > 0$ independent of $\varepsilon$. By choosing a suitable $\varepsilon_3 \leq \varepsilon_2$, we then have, for all $\epsilon \leq \varepsilon_3$, and all along the path,

$$
\frac{1}{2\sigma} \mathcal{F}_\epsilon \leq 2\mathcal{H}''(M) + \int_N \frac{g}{2\sigma} d\mathcal{H}^{n+1} - \frac{\zeta}{2}.
$$

(66)
We now proceed to conclude the proof of Proposition 6.1. Recalling the assumption set up at the beginning of Section 6, the minmax solutions \( u_\varepsilon \) (obtained in Proposition 5.1) have the property that \( \lim_{\varepsilon \to 0} u_\varepsilon = u_\infty \equiv -1 \) and

\[
\frac{1}{2}\mathcal{E}_\varepsilon(u_\varepsilon) \to 2\mathcal{H}^n(M) + \int_N \frac{g}{2\sigma} d\mathcal{H}^{n+1}.
\]

This implies that, for \( \varepsilon = b_\varepsilon \); the function \( h_\infty \) cannot be \( b_\varepsilon \); otherwise, the continuous path in \( W^{1,2}(N) \) that starts at \( a_\varepsilon \) and ends at \( b_\varepsilon \) and satisfies (66) would contradict the minmax characterization of \( u_\varepsilon \). Therefore \( h_\infty \) is a stable solution to \( \mathcal{E}_\varepsilon = 0 \) that does not coincide with \( b_\varepsilon \). We now make the dependence on \( \varepsilon \) explicit, and denote \( h_\infty \) by \( v_\varepsilon \); we check next that \( v_\varepsilon \) are the solution sought to complete the proof.

The stable solutions \( v_\varepsilon \) have a uniform upper bound on \( \mathcal{E}_\varepsilon(v_\varepsilon) \), because

\[
\frac{1}{2}\mathcal{F}_\varepsilon(v_\varepsilon) \leq 2\mathcal{H}^n(M) + \int_N \frac{g}{2\sigma} d\mathcal{H}^{n+1} - \frac{\varepsilon}{2}
\]

and \(-1 \leq v_\varepsilon \leq b_\varepsilon \leq 1 + c_W \varepsilon \). The condition in Lemma 6.1 gives a fixed non-empty open set on which \( \varepsilon_i > \frac{3}{4} \) for all \( \varepsilon_i \).

It only remains to prove that \( \mathcal{E}_{\varepsilon_i}(v_{\varepsilon_i}) \) is bounded from above. To see this, we will prove first of all that if \( \mathcal{E}_{\varepsilon_i}(v_{\varepsilon_i}) \to 0 \) with \( \varepsilon_i \to 0 \) for a sequence \( w_i \in W^{1,2}(N) \) satisfying \( \mathcal{F}_{\varepsilon_i}(w_i) = 0 \) and \( \lim sup_{i \to \infty} sup_N |w_i| < \infty \), then we must have that \( \{ w_i = 0 \} = \emptyset \) for all sufficiently large \( i \). If this is false, then passing to a subsequence without relabeling, we find points \( y_i \) such that \( w_i(y_i) = 0 \). Let \( r_0 \in (0, inf(N)) \). Let \( \tilde{w}_i : B_{\frac{r_0}{\varepsilon_i}}(0) \to \mathbb{R} \) be defined by \( \tilde{w}_i(x) = w_i(exp_{y_i}(\varepsilon_i x)) \).

Then \( \tilde{w}_i \) solves the PDE \( \Delta \tilde{w}_i - W'(\tilde{w}_i) = -\varepsilon_i g \) on \( B_{\frac{r_0}{\varepsilon_i}}(0) \), where \( \Delta \) is the Laplace-Beltrami operator with respect to the metric obtained by pulling-back the Riemannian metric to \( B_{\frac{r_0}{\varepsilon_i}}(0) \) via the map \( x \to exp_{y_i}(\varepsilon_i x) \). Note that \( \tilde{w}_i(0) = 0 \) for all \( i \) and that for any compact set \( K \subset \mathbb{R}^{n+1} \) we have that \( \int_K \frac{1}{2} |\nabla \tilde{w}_i|^2 + W(\tilde{w}_i) \to 0 \) as \( i \to \infty \). Since \( w_i \) is bounded, it follows from the De Giorgi–Nash–Moser estimates that \( \tilde{w}_i \) is locally uniformly Hölder continuous on \( \mathbb{R}^{n+1} \).

By Schauder theory, we then have in fact that \( \tilde{w}_i \) is locally uniformly bounded in \( C^{2,\alpha} \) for any \( \alpha \in (0, 1) \). Using a diagonal argument, we then obtain an entire \( C^{2,\alpha} \) solution \( \tilde{w} \) to \( \Delta \tilde{w} - W'(\tilde{w}) = 0 \) on \( \mathbb{R}^{n+1} \) such that \( \int_{\mathbb{R}^{n+1}} \frac{1}{2} |\nabla \tilde{w}|^2 + W(\tilde{w}) = 0 \). This forces \( \tilde{w} \equiv 1 \) or \( \tilde{w} \equiv -1 \). But this is impossible since we must also have \( \tilde{w}(0) = 0 \). Thus we have \( \{ w_i = 0 \} = \emptyset \) as asserted. Returning to the sequence \( (v_{\varepsilon_i}) \), if \( \lim inf_{i \to \infty} \mathcal{E}_{\varepsilon_i}(v_{\varepsilon_i}) = 0 \), then passing to subsequence and taking \( w_i = v_{\varepsilon_i} \) in the preceding discussion, we see that since \( v_{\varepsilon_i} \) is somewhere positive by Lemma 6.1, we must have that \( \inf_N v_{\varepsilon_i} > 0 \) for all sufficiently large \( i \). Since \( v_{\varepsilon_i} \) satisfies \( \varepsilon_i^2 \Delta v_{\varepsilon_i} - \varepsilon_i W'(v_{\varepsilon_i}) + g = 0 \) on \( N \), evaluating at any \( z \in N \) with \( v_{\varepsilon_i}(z) = \inf_N v_{\varepsilon_i} \), we see that \( W'(\inf_N v_{\varepsilon_i}) = \varepsilon_i^2 \Delta v_{\varepsilon_i}(z) + \varepsilon_i g(z) > 0 \). Since \( W'(t) \leq t \) for \( t \in [0, 1] \) and \( \inf_N v_{\varepsilon_i} > 0 \) for all sufficiently large \( i \), this implies that \( \inf_N v_{\varepsilon_i} > 1 \) for all sufficiently large \( i \). To complete the argument, note that the negative gradient flow of \( \mathcal{E}_{\varepsilon_i} \) with initial condition \( 1 \) tends to \( b_{\varepsilon_i} \) (by the definition of the latter). We consider the negative gradient flow for \( \mathcal{E}_{\varepsilon_i} \) with initial condition \( v_{\varepsilon_i} \); this flow is time-independent. Recalling that \( 1 \leq v_{\varepsilon_i} \leq b_{\varepsilon_i} \), the parabolic maximum principle then implies that \( v_{\varepsilon_i} = b_{\varepsilon_i} \).

We have thus shown that if \( \lim \inf_{i \to \infty} \mathcal{E}_{\varepsilon_i}(v_{\varepsilon_i}) = 0 \) then along a subsequence we have \( v_{\varepsilon_i} = b_{\varepsilon_i} \), a possibility we have excluded earlier in this section. Therefore we must have that \( \lim \inf_{i \to \infty} \mathcal{E}_{\varepsilon_i}(v_{\varepsilon_i}) > 0 \). This concludes the proof of Proposition 6.1.

Remark 6.7 (The case \( \text{Ric}_N > 0, g \equiv cnsf \)). In the case in which \( N \) has positive Ricci curvature and \( g \equiv \lambda \in (0, \infty) \), it follows from the expression of the second
variation of $F_{\epsilon,\lambda}$ that stable solutions to $F_{\epsilon,\lambda} = 0$ must be constant functions on $N$ (see e.g. [Bel, Proposition 7.1]). It is then easy to check that there are only two stable solutions to $F_{\epsilon,\lambda} = 0$, one of which is a constant close to $-1$ and the other is a constant close to $+1$. Then the first solution has to be the function that we denoted by $u_\epsilon$, while the second is the function that we denoted by $b_\epsilon$. In particular, in the argument given at the beginning of this section, the function $h_\infty = v_\epsilon$ had to agree with $b_\epsilon$, giving a contradiction. We therefore conclude that when $N$ has positive Ricci curvature and $\epsilon \equiv \lambda \in (0, \infty)$ the minmax solutions $u_\epsilon$ produced by Proposition 5.1 cannot yield (through their associated varifolds) a completely minimal hypersurface, in other words the open set $\{u_\infty = +1\}$ is non-empty.

7 Extension to the case of non-negative $g$ of class $C^{0,\alpha}$

Having completed (in Section 6) the proof of the existence result in the case in which $g \in C^{1,1}(N)$ and $g > 0$, we address, in this section, the case in which $g \in C^{0,\alpha}(N)$ and $g \geq 0$, thereby establishing Theorem 1.1 in full. We thus let $g : N \to [0, \infty)$ be of class $C^{0,\alpha}$ for some $\alpha \in (0, 1)$. We approximate $g$ by $g_j \in C^{\infty}(N)$ such that $g_j \to g$ in $C^{0,\alpha}(N)$ and $g_j > 0$. (To this end, it suffices to consider $g + \delta_j$, for a sequence of positive real numbers $(\delta_j)$ with $\delta_j \to 0^+$ as $j \to \infty$ and then define $g_j$ to be the convolution of $g + \delta_j$ with a standard mollifier.)

For each $j$ we apply the results obtained in Sections 5 and 6. More precisely, we know that for every $j$ there exists a sequence $(\epsilon_j^k)$ with $\epsilon_j^k \to 0^+$ as $k \to \infty$ and a sequence of critical points $u_{\epsilon_j^k} : N \to \mathbb{R}$ of $F_{\epsilon_j^k, \sigma g_j}$ with Morse index at most 1 with the following properties: the associated varifolds $V_j^k = V_{\epsilon_j^k}$ satisfy $V_j^k \to \sigma V_j$ as $k \to \infty$, where $V_j$ is a non-zero integral varifold that can be written in the form $V_j = V_0^j + V_{g_j}$ such that the conclusions of Theorem 4.1 hold with $V_0^j$, $V_{g_j}$, $\sigma V_j$ in place of $V_0$, $V_g$, $V$ respectively, and (crucially) with $V_{g_j} \neq 0$. (For each $j$, we have either (a) for each $k$, $u_{\epsilon_j^k}$ is a min-max critical point of $F_{\epsilon_j^k, \sigma g_j}$ given by Proposition 5.1 taken with $\epsilon = \epsilon_j^k$, or (b) for each $k$, $u_{\epsilon_j^k} = v_{\epsilon_j^k}$ where $v_{\epsilon_j^k}$ is the stable critical point of $F_{\epsilon_j^k, \sigma g_j}$ given by Proposition 6.1 taken with $\epsilon_j^k$ in place of $\epsilon_j$.)

A natural idea would be to take the varifold limit of $V_{g_j}^j$ as $j \to \infty$ in order to produce the desired hypersurface whose mean curvature is prescribed by $g$. However, in the absence of an explicit Morse index estimate on gen-reg $V_{g_j}$ (for the functional $A - \text{Vol}_{g_j}$), we cannot directly infer regularity of $\lim_{j \to \infty} V_{g_j}^j$. Therefore we will proceed as in the following outline: we take the varifold limit of $V_j^j$ (including the possible minimal portions), exploit the index information on $u_{\epsilon_j^k}$ to infer optimal regularity of $V = \lim_{j \to \infty} V_j^j$ (namely that, away from a genuine singular set of dimension $\leq n - 7$, $\text{spt} \|V\|$ is locally given by the union of a finite number of $C^2$ embedded graphs), and that the convergence $V_j \to V$ is smooth (locally graphical) away possibly from one point (“neck region”) and from a set of dimension $\leq n - 7$ (the genuine singular set of $V$). Broadly described, the argument for the optimal regularity of $V$ relies on a tangent cone analysis on $V$ and on the application of (ii) of Theorem 4.3 to the sequence $V_j^j$ uniformly in $j$, in balls centred at points of $V$ at which at least one tangent cone to $V$ is supported on a hyperplane. This argument simultaneously yields the statement on the convergence $V_j \to V$. Once the smooth convergence $V_j \to V$ is established, in the sense described above, one deduces that also $V_{g_j}^j \to V_g$ smoothly away from a set of dimension $\leq n - 7$, and that $\text{spt} \|V_g\|$ provides the sought immersed hypersurface.
part (ii) of Theorem 4.3 are valid for not specifying the pushforward of $V$. We recall that (as in Theorem 4.3) we are using normal coordinates at $p$ and the open ball $B_p = B_{2r}(p) \subset N$ such that the varifold $V \varphi B_{2r}(p)$ satisfies (working in an exponential chart around $p$, as):

(A) the mass of $V \varphi B_{2r}(p)$ is between the values $(\omega_n 2^n) \Theta(|V||p|(2r)^n - \frac{1}{2} (\omega_n 2^n)(2r)^n$ and $(\omega_n 2^n) \Theta(|V||p|(2r)^n + \frac{1}{2} (\omega_n 2^n)(2r)^n$;

(B) the mass of $V \varphi B_{2r}(p)$ in the cylinder $(B_{2r}(0) \times \mathbb{R}) \cap B_{2r+1}(0)$ is between the values $\omega_n \Theta(|V||p|r^n - \frac{1}{2} \omega_n r^n$ and $\omega_n \Theta(|V||p|r^n + \frac{1}{2} \omega_n r^n$;

(C) the coarse excess $E_{2r}$ of $V \varphi B_{2r}(p)$ in the cylinder $(B_{2r}(0) \times \mathbb{R}) \cap B_{2r+1}(0)$ is smaller than $\varepsilon_0/2$, where $\varepsilon_0$ is identified in part (ii) of Theorem 4.3.

We recall that (as in Theorem 4.3) we are using normal coordinates at $p$ the first factor in $B_{2r}(0) \times \mathbb{R}$ is in the chosen tangent plane to $V \varphi B_{2r}(p)$ at $p$, and the coarse excess $E_{2r}$ is defined in Theorem 4.4. By abuse of notation we are not specifying the pushforward of $V$ via the inverse of the exponential map. By varifold convergence, for each ball $B_p$ the conditions on mass and excess stated in part (ii) of Theorem 4.3 are valid for $V \varphi B_{2r}(p)$ for all sufficiently large $j$; indeed, varifold convergence implies that for all sufficiently large $j$ we have the validity of (A) and (B) for $V \varphi B_{2r}(p)$ with $\varepsilon$ replaced by $\varepsilon/2$, and the validity of (C) for $V \varphi B_{2r}(p)$ with $\varepsilon_0/2$ replaced by $\varepsilon_0/2$.

Recall that we know the regularity of $V$ already, and $|H_{V^j}| \leq \max g + 1$ for all $j$. We now consider the following two possibilities, one of which must be true:

(a) there exists a subsequence $j_k \to \infty$ such that for every $p$ the restrictions $V^j \varphi B_{p}$ satisfy the Schoen–Tonegawa inequality (with respect to $D_p$);

(b) there is $j_0 \in \mathbb{N}$ such that for every $j \geq j_0$ there exists $p = p_j$ such that the Schoen–Tonegawa inequality (with respect to $D_p$) fails in $B_p$ for $V \varphi B_{2r}(p)$.

**Remark 7.1.** If the Schoen–Tonegawa inequality fails in $B_p$ for $V \varphi B_{2r}(p)$ (for a certain $j$), then $u_{j_k} = u$ cannot be stable in $B_p$ for all sufficiently large $k$ (for otherwise, we would conclude the validity of the Schoen–Tonegawa inequality in $B_p$ by arguing as in Section 4.3). The fact that $u_{j_k}$ has Morse index $\leq 1$ in $N$, together with the fact that $u_{j_k}$ is not stable in $B_p$, then implies that $u_{j_k}$ is stable in $N \setminus B_p$ for all sufficiently large $k$.

Let us assume that case (b) occurs. Then we consider $B_p$, as $j \to \infty$. As $N$ is compact and the radii are all bounded by $R$, one can find $p_\infty \in N$ such that the ball $B_{2R}(p_\infty)$ contains $B_p$, for all sufficiently large $j$, say $j \geq j_1$. Then, by Remark 7.1, we have that, for every $j \geq j_1$, $u_{j_k}$ is stable in $N \setminus B_{2R}(p_\infty)$ for all sufficiently large $k$. This implies the validity of the Schoen–Tonegawa inequality for $V \varphi B$ for every geodesic ball $B$ contained in $N \setminus B_{2R}(p_\infty)$ and for every $j \geq j_1$. (Once again, this follows as in Section 4.3) In particular, the Schoen–Tonegawa inequality holds for $V \varphi B$ when $B$ is one of the balls chosen at the beginning and $B_p$ additionally satisfies $B_p \subset N \setminus B_{2R}(p_\infty)$. This allows us to use part (ii) of Theorem 4.3 to conclude that for all $j \geq j_1$ the restriction $V^j \varphi B_p$ is made of $\Theta(|V||p|)$ graphs (over $D_p$) of class $C^2$ and satisfying a $C^2,\theta$ estimate uniformly in $j$ for some $\theta \in (0, 1)$. (To this end, note that the constant on the right-hand-side of the estimate in (ii) of Theorem 4.3 only depends on the $L^\infty$ norm of $g_j$, and by construction $\|g_j\|_{L^\infty(N)}$ is uniformly bounded in $j$.)
The absence of classical singularities for \( V \subset (N \setminus B_{2\mathcal{R}}(p_\infty)) \) follows from part (i) of Theorem 4.3.

Remark 7.2. A standard argument, that we quickly recall, shows that any tangent cone to \( V \) (which is a stationary varifold) must have stable regular part. Let \( p \in \text{spt} \|V\| \) and consider a sequence of dilations at \( p \) that lead to a tangent cone \( T_p \) (with tip at the origin in \( \mathbb{R}^{n+1} \)). By suitable translations/dilations, and taking a diagonal sequence, we can construct a sequence \( \tilde{u}_{\epsilon_j} \) such that the associated varifolds converge to \( T_p \) as \( j \to \infty \). Here each \( \tilde{u}_{\epsilon_j} \) is defined on \( B_1(0) \) by taking a suitable translation/dilation of \( u_{\epsilon_j} \). For \( \rho > 0 \) arbitrary, consider the ball \( B_{\rho}^{n+1}(0) \) and the annulus \( B_{2\rho}^{n+1}(0) \setminus \overline{B}_{\rho}^{n+1}(0) \). Then, by the condition that each \( \tilde{u}_{\epsilon_j} \) has Morse index at most 1, we obtain that either there exists a subsequence \( \tilde{u}_{\epsilon_{j+1}} \) such that \( \tilde{u}_{\epsilon_{j+1}} \) is stable in \( B_1(0) \) for each \( j \), or \( \tilde{u}_{\epsilon_{j+1}} \) is stable in \( B_{2\rho}^{n+1}(0) \setminus \overline{B}_{\rho}^{n+1}(0) \) for all sufficiently large \( j \). If the former occurs for some \( \rho > 0 \) then the stability of \( T_p \) is immediate. On the other hand, if for every \( \rho \) the latter option occurs, we obtain the stability of \( T_p \) in \( B_1(0) \setminus \{0\} \), which can be extended to \( B_1(0) \) by a capacity argument (since \( n \geq 2 \)).

Combining the information just acquired with a standard tangent cone analysis (as done also in Section 1.3), we obtain that \( \text{spt} \|V\| \cap (N \setminus B_{2\mathcal{R}}(p_\infty)) \) is, away from a set of dimension \( \leq n - 7 \) that we denote by \( \text{sing} V \) (whose points are characterised by the condition that no tangent cone to \( V \) is supported on a hyperplane), locally given by the ordered union of \( C^2 \) graphs; moreover, the convergence \( V^j \to V \) happens locally in \( N \setminus B_{2\mathcal{R}}(p_\infty) \setminus \text{sing} V \) in the sense of \( C^2 \) graphs. The arbitrariness of \( R \) allows to extend these conclusions to \( N \setminus (\{p_\infty\} \cup \text{sing} V) \). At this stage, we do not know any regularity for \( \text{spt} \|V\| \) at the point \( p_\infty \). We note that smooth graphical convergence might fail at \( p_\infty \) even if \( p_\infty \) was known to be a point around which \( \text{spt} \|V\| \) is \( C^2 \) embedded (as in the example of catenoidal necks degenerating to a double plane).

If case (a) occurs, then the above argument can be followed without the need to remove \( B_{2\mathcal{R}}(p_\infty) \); in this case we obtain the same regularity conclusion for \( V \) in \( N \setminus \text{sing} V \), with \( \text{dim}_{\mathcal{R}}(\text{sing} V) \leq n - 7 \), and the convergence \( V^j \to V \) happens locally in \( N \setminus \text{sing} V \) in the sense of \( C^2 \) graphs.

The \( C^2_{\text{loc}} \) convergence \( V^j \to V \), away from a possible set of dimension \( \leq n - 7 \) and possibly away from an additional point, implies that the convergence \( V^j \to V \) (the regularity of \( V^j \) is already known to us) is also in \( C^2_{\text{loc}} \), away from a possible singular set \( \Sigma = \text{sing} V \cap \text{spt} \|V\| \) with \( \text{dim}_{\mathcal{R}}(\Sigma) \leq n - 7 \) and possibly away from an additional point \( p_{\text{sing}} \). (Clearly \( \text{spt} \|V\| \subset \text{spt} \|V\| \).) Note that \( V_j \neq 0 \) because \( V^j \neq 0 \) for all \( j \) and, in view of the locally uniform \( L^\infty \)-bound on \( (g_j) \), there is a positive lower bound for \( \|V^j\|(N) \) by the monotonicity formula applied in a ball of radius \( \text{inj}(N) \) around an arbitrary point of \( \text{spt} \|V^j\| \). By said convergence of \( V^j \) to \( V_{\text{g}} \), and since the \( C^2 \) graphs locally describing \( V^j \) on \( N \setminus \Sigma \) have scalar mean curvature \( g_j \), we conclude that \( V^j \) is, away from \( \Sigma \) and possibly from \( p_{\text{sing}} \), locally given by a union of \( C^2 \) graphs each having mean curvature \( g_{\text{g}} \) for one of the two choices of unit normal \( \hat{n} \) on each graph.

It now follows from Theorem 3.1 that \( V_{\text{g}} \) has multiplicity 1 on \( \text{reg} V_{\text{g}} \cap \{g > 0\} \). For if not, then there is a ball \( B = B_{\rho}(y) \) with \( g > 0 \) on \( B \) such that \( V_{\text{g}} \subset B \) has integer multiplicity \( k \geq 2 \) and \( \|V_{\text{g}}\| \) has mean curvature \( g_{\text{g}} \hat{n} \) for some choice of continuous unit normal \( \hat{n} \). By the local \( C^2 \) convergence \( V_{\text{g}} \to V_{\text{g}} \), it follows that there is an open set \( \Omega \subset N \) with \( \text{spt} \|V_{\text{g}}\| \cap B_{\rho}(y) \subset \Omega \) and \( V_{\text{g}} \subset \Omega \) consists of \( k \) graphs \( G_j^\ell, \ell = 1, 2, \ldots, k, \) of class \( C^2 \), over \( \text{spt} \|V_{\text{g}}\| \cap B_{\rho}(y) \), with each \( G_j^\ell \) converging in \( C^2 \) to \( \text{spt} \|V_{\text{g}}\| \cap B_{\rho}(y) \) as \( j \to \infty \). Furthermore, since \( g_j > 0 \) in \( B \),
it follows by Theorem 3.1 (iii) that for each $j$, the graphs $G_j^i$, $i = 1, 2, \ldots, k$, are distinct, and have mean curvature given by $g_j \nu_j$ with the unit normal vectors $\nu_j$ to $G_j^i$ close to $\hat{n}$ (by the $C^2$ convergence). Since $\bigcup_{i=1}^k G_j^i$ is contained in the phase boundary, this contradicts Theorem 3.1 (iii) which says that the mean curvature vector must point into the +1 phase everywhere on the phase boundary.

The smooth convergence $V_j^i \to V_j$ on $N \setminus (\Sigma \cup \{p_\infty\})$ further gives that condition (T) in [BelWic-2] is satisfied by $V_j$. We can then apply [BelWic-2, Theorem 4.1] and [BelWic-2, Remark 4.6] (whose validity extends without any modification to the case in which the prescribing function is $C^{0,\alpha}$) to conclude that the varifold $V_j \mathcal{L}(N \setminus (\Sigma \cup \{p_\infty\}))$ can be realized as the pushforward of an $n$-manifold via a two-sided immersion with mean curvature given by $g \nu$, where $\nu$ is a choice of unit normal to the immersion. In view of the low dimensionality of $\Sigma \cup \{p_\infty\}$, this is the same as affirming that $V_j$ can be realized by the same pushforward.

What is left to check is that when $n \leq 6$, if case (b) occurs, then $p_\infty \in \text{gen-reg } V_j$ (see Remark 4.1 and [BelWic-2, Definition 3.17]), so that $p_\infty$ can be absorbed in $\Sigma$ (for $n \geq 7$) without altering the condition $\dim_H(\Sigma) \leq n - 7$; then $\text{spt } V_j$, $j \in \mathbb{N}$, satisfies the optimal regularity conclusion on the whole of $N$, completing the proof of Theorem 1.1. In order to do this, we observe first of all (recalling Remark 7.2) that any tangent cone to $\Sigma$ is a choice of unit normal system of coordinates at $p$, in a finite collection of cylinders over balls $B$ contained in the phase boundary, this contradicts Theorem 3.1 (iii) which says that the mean curvature vector must point into the +1 phase everywhere on the phase boundary.

The smooth convergence $V_j^i \to V_j$ on $N \setminus (\Sigma \cup \{p_\infty\})$ further gives that condition (T) in [BelWic-2] is satisfied by $V_j$. We can then apply [BelWic-2, Theorem 4.1] and [BelWic-2, Remark 4.6] (whose validity extends without any modification to the case in which the prescribing function is $C^{0,\alpha}$) to conclude that the varifold $V_j \mathcal{L}(N \setminus (\Sigma \cup \{p_\infty\}))$ can be realized as the pushforward of an $n$-manifold via a two-sided immersion with mean curvature given by $g \nu$, where $\nu$ is a choice of unit normal to the immersion. In view of the low dimensionality of $\Sigma \cup \{p_\infty\}$, this is the same as affirming that $V_j$ can be realized by the same pushforward.

What is left to check is that when $n \leq 6$, if case (b) occurs, then $p_\infty \in \text{gen-reg } V_j$ (see Remark 4.1 and [BelWic-2, Definition 3.17]), so that $p_\infty$ can be absorbed in $\Sigma$ (for $n \geq 7$) without altering the condition $\dim_H(\Sigma) \leq n - 7$; then $\text{spt } V_j$, $j \in \mathbb{N}$, satisfies the optimal regularity conclusion on the whole of $N$, completing the proof of Theorem 1.1. In order to do this, we observe first of all (recalling Remark 7.2) that any tangent cone to $\Sigma$ is a choice of unit normal system of coordinates at $p$, in a finite collection of cylinders over balls $B$ contained in the phase boundary, this contradicts Theorem 3.1 (iii) which says that the mean curvature vector must point into the +1 phase everywhere on the phase boundary.

References

[Bel] C. Bellettini, Multiplicity-1 minmax minimal hypersurfaces in manifolds with positive Ricci curvature, arXiv 2020.

[BelWic-1] C. Bellettini, N. Wickramasekera, Stable CMC integral varifolds of codimension 1: regularity and compactness, arXiv 2018.

[BelWic-2] C. Bellettini, N. Wickramasekera, Stable prescribed-mean-curvature integral varifolds of codimension 1: regularity and compactness, arXiv 2019.

[ChoMan20] O. Chodosh, C. Mantoulidis, Minimal surfaces and the Allen–Cahn equation on 3-manifolds: index, multiplicity, and curvature estimates, Ann. of Math. vol. 191, no. 1, pp. 213-328 (2020).

[Dey] A. Dey, Existence of multiple closed CMC hypersurfaces with small mean curvature arXiv 2019 arXiv:1910.00089.

[EvaGar] L. C. Evans, R. F. Gariepy, Measure Theory and Fine Properties of Functions, CRC press, revised edition (2015).
[GilTru] D. Gilbarg, N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 3rd Edition, Springer-Verlag.

[GMT80] E. Gonzalez, U. Massari, I. Tamanini *Existence and regularity for the problem of a pendent liquid drop*, Pacific J. Math. 88 (1980) n. 2, 399-420.

[GMT83] E. Gonzalez, U. Massari, I. Tamanini *On the regularity of boundaries of sets minimizing perimeter with a volume constraint* Indiana Univ. Math. J. 32 (1983), no. 1, 25-37.

[Gra] A. Gray, *Tubes*, second edition, Progress in Mathematics, Birkhäuser (2004).

[Gro98] K. Große-Brauckmann, *Complete embedded constant mean curvature surfaces*, Habilitationsschrift, Universität Bonn, 1998.

[Gua18] M. Guaraco, *Min-max for phase transitions and the existence of embedded minimal hypersurfaces*, J. Diff. Geom. 108, 1 (2018), 91-133.

[Hic18] F. Hiesmayr. *Spectrum and index of two-sided Allen–Cahn minimal hypersurfaces*. Comm. in PDE. 43, 11, (2018), 1541-1565.

[HutTon00] J. E. Hutchinson, Y. Tonegawa, *Convergence of phase interfaces in the van der Waals–Cahn–Hilliard theory*, Calc. Var. 10 (2000), no. 1, 49-84.

[Ilm93] T. Ilmanen, *Convergence of the Allen–Cahn equation to Brakke’s motion by mean curvature*, J. Diff. Geom. 38, 2 (1993), 417-461.

[ManMen02] C. Mantegazza, A. Mennucci, *Hamilton–Jacobi Equations and Distance Functions on Riemannian Manifolds*, Appl. Math. Optim. vol. 47, 1-25 (2002).

[ModMor77] L. Modica, S. Mortola, *Un esempio di Γ-convergenza*, Boll. U.M.I. 14-B (1977) 285-299.

[RogTon08] M. Röger, Y. Tonegawa, *Convergence of phase-field approximations to the Gibbs–Thomson law*, Calc. Var. 32 (2008) 111-136.

[Sch01] R. Schätzle, *Hypersurfaces with mean curvature given by an ambient Sobolev function*. J. Diff. Geom. 58, 3, 371-420 (2001).

[Sch04] R. Schätzle, *Quadratic tilt excess decay and strong maximum principle for varifolds*. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 3(1), 171-231 (2004).

[SchSim81] R. Schoen, L. Simon, *Regularity of stable minimal hypersurfaces*, Comm. Pure Appl. Math. 34 (1981), 741-797.

[SSY75] R. Schoen, L. Simon, S. -T. Yau. *Curvature estimates for minimal hypersurfaces*. Acta Math. 134 (1975), 276–288.

[Sim68] J. Simons. *Minimal varieties in Riemannian manifolds*. Ann. of Math. 88 (1968), 62–105.

[Str] M. Struwe, *Variational methods*, 4th edition, Springer.

[Ton05-1] Y. Tonegawa, *On stable critical points for a singular perturbation problem*, Comm. An. Geom. (2005), 13, 2, 439-459.

[TonWic12] Y. Tonegawa, N. Wickramasekera, *Stable phase interfaces in the van der Waals–Cahn–Hilliard theory*, J. Reine Angew. Math., 668:191-210, 2012.

[Ton05-2] Y. Tonegawa, *A diffused interface whose chemical potential lies in a Sobolev space*, Ann Sc. Norm. Super. Pisa Cl. Sci. (5) 4(3), 487-510 (2005).
[Whi91] B. White, *The space of minimal submanifolds for varying Riemannian metrics* Indiana Univ. Math. J. 40 (1) (1991), 161-200.

[Wic14] N. Wickramasekera, *A general regularity theory for stable codimension 1 integral varifolds*, Ann. of Math. (2), 179(3):843-1007, 2014.

[ZhoZhu19] X. Zhou, J. Zhu *Min-max theory for constant mean curvature hypersurfaces* Invent. Math. (2019) 218: 441-490.

[ZhoZhu20] X. Zhou, J. Zhu *Existence of hypersurfaces with prescribed mean curvature I - Generic min-max* Camb. Journ. of Math. (2020) 8, 2, 311-362.

Costante Bellettini  
Department of Mathematics  
University College London  
London WC1E 6BT, United Kingdom  
C.Bellettini@ucl.ac.uk

Neshan Wickramasekera  
DPMMS  
University of Cambridge  
Cambridge CB3 0WB, United Kingdom  
N.Wickramasekera@dpmms.cam.ac.uk