Accepted: 27 January 2022

DOI: 10.1002/cpa.22144

RESEARCH ARTICLE

Multiplicity-1 minmax minimal hypersurfaces in manifolds with positive Ricci curvature

Costante Bellettini

Department of Mathematics, University College London, London, UK

Correspondence
Costante Bellettini, Department of Mathematics, University College London, 25 Gordon Street, WC1H 0AY, London, UK.
Email: c.bellettini@ucl.ac.uk

Funding information
EPSRC, Grant/Award Number: EP/S005641/1; National Science Foundation, Grant/Award Number: DMS-1638352

Abstract

We address the one-parameter minmax construction for the Allen–Cahn energy that has recently lead to a new proof of the existence of a closed minimal hypersurface in an arbitrary compact Riemannian manifold N^{n+1} with $n \geq 2$ (Guaraco's work, relying on works by Hutchinson, Tonegawa, and Wickramasekera when sending the Allen–Cahn parameter to 0). We obtain the following result: if the Ricci curvature of N is positive then the minmax Allen–Cahn solutions concentrate around a *multiplicity-1* minimal hypersurface (possibly having a singular set of dimension $\leq n-7$). This multiplicity result is new for $n \geq 3$ (for $n = 2$ it is also implied by the recent work by Chodosh–Mantoulidis). We exploit directly the minmax characterization of the solutions and the analytic simplicity of semilinear (elliptic and parabolic) theory in $W^{1,2}(N)$. While geometric in flavour, our argument takes advantage of the flexibility afforded by the analytic Allen–Cahn framework, where hypersurfaces are replaced by diffused interfaces; more precisely, they are replaced by sufficiently regular functions (from N to $\mathbb{R}$), whose weighted level sets give rise to diffused interfaces. We capitalise on the fact that (unlike a hypersurface) a function can be deformed both in the domain N (deforming the level sets) and in the target $\mathbb{R}$ (varying the values). We induce different geometric effects on the diffused interface by using these two types of deformations; this enables us to implement in a continuous way certain operations, whose analogues on a hypersurface would be discontinuous. An immediate corollary of the multiplicity-1 conclusion is that every compact Riemannian manifold N^{n+1} with $n \geq 2$ and positive Ricci curvature admits a two-sided closed minimal hypersurface, possibly with a singular set of dimension at most $n-7$. (This geometric corollary also follows from results obtained by different ideas in an Almgren–Pitts minmax framework.)

CONTENTS

1 | INTRODUCTION

The close link between Allen–Cahn energy and minimal hypersurfaces has its roots in the ideas pioneered by De Giorgi in the development of Γ-convergence. The works of Modica–Mortola [25] and Kohn–Sternberg [20], among many others, testify to the fine suitability of the Allen–Cahn approximation method for the study of area-minimisers. Moving away from the minimising case, the Allen-Cahn approximation has seen further success in recent years, starting with the combined works of Guaraco, Hutchinson, Tonegawa, and Wickramasekera [14, 16, 36, 37, 40]: their outcome was a new proof (that uses classical PDE minmax techniques) of the existence of a closed minimal hypersurface in an arbitrary compact Riemannian manifold of dimension 3 or higher. Moving to higher codimension problems, very recently the work of Pigati–Stern [27] made another fundamental contribution: after identifying a correct energy (and geometric framework), it carries out an approximation EPSRC procedure reminiscent in many ways of the Allen–Cahn one; this leads to a new proof (again via classical PDE minmax) of the existence of stationary integral varifolds of codimension 2 (the natural candidates for closed minimal submanifolds of codimension 2) in an arbitrary compact Riemannian manifold.

The original proof of the existence of stationary integral varifolds in a compact Riemannian manifold was obtained by Almgren [2] (in arbitrary codimension). In codimension-1, the work of Pitts [28], together with the regularity and compactness theory of Schoen–Simon–Yau [32] and Schoen–Simon [33], provided the information that the varifold obtained is in fact a closed minimal hypersurface (i.e., smooth except for an expected singular set of codimension 7 or higher). In answering (positively) the above existence question, Almgren and Pitts developed a considerable amount of machinery, which has been extended and further developed in the past decade, leading to impressive progress in the field, particularly for codimension-1 questions (starting with the resolution, by Marques–Neves [22], of the long-standing Willmore conjecture). The power so far deployed by the Almgren–Pitts minmax method is counterpoised by certain intrinsic difficulties that make it rather involved: the space of integral varifolds on N (or variants of it), on which the minmax is carried out lacks a linear structure and, moreover, no Palais–Smale condition is available for the area functional on this space. The Allen–Cahn minmax method looks, on the other hand, for saddle-type solutions to a semilinear elliptic problem on the Hilbert space $W^{1,2}$, and the validity of a Palais–Smale condition permits the use of classical PDE minmax tools, leading to convenient Morse index bounds.

A common feature in all variational minmax constructions is the fact that the geometric objects produced are integral varifolds, and as such carry an a.e. integer-valued multiplicity. Proving that this multiplicity is (a.e.) equal to 1 can lead to further geometric consequences. The work of Chodosh–Mantoulidis [6] (valid more generally for solutions with bounded Morse index, not necessarily minmax solutions) implies that the minimal surface obtained by a one- or multi-parameter Allen-Cahn minmax is two-sided and has multiplicity 1 when the Riemannian manifold is three-dimensional and the metric is bumpy or has positive Ricci curvature. Combining this with the result of [10] on the Weyl's law, [6] obtained the validity of the generic version of Yau's conjecture for three-manifolds (on the existence of infinitely many hypersurfaces).[1]

The multiplicity question is ubiquitous in the field. The work of Zhou [43] proves the multiplicity-1 conclusion for one- or multi-parameter Almgren–Pitts minmax, when the metric is bumpy or has positive Ricci and the dimension of the manifold is between 3 and 7, as conjectured by Marques–Neves [23, 1.2] (see also [23, Addendum]). For the recent viscosity approach to the minmax for surfaces in arbitrary codimension, proposed by Rivière in [30], the work of Pigati–Rivière [26] established a multiplicity-1 conclusion for the critical points constructed.

Heuristically, and regardless of the specific framework used for the construction, the sought submanifold (more precisely, integral varifold) is always obtained as a limiting object from a certain sequence; the relevance of the multiplicity-1 conclusion lies in the fact that it very much constrains the fashion in which this limit arises. Thanks to it, finer pieces of information that are available on the sequence can pass to the limit in a straightforward way. Higher multiplicity hypersurfaces, on the other hand, could arise in many different ways, possibly with degeneration of certain features and preventing the passage to the limit of certain properties. In the case of the Allen–Cahn equation, Wang [38] provides a $C^{1,\alpha}$-sheeting result (with Allard-type estimates) under multiplicity-1 convergence.

Our main theorem is the following multiplicity-1 result (new for $n \geq 3$), which applies to the Allen–Cahn (one-parameter) minmax construction in Guaraco's work [14]. The case $n = 2$ also follows from Chodosh–Mantoulidis's result [6].

[1] Yau's conjecture was then established in full, for manifolds of dimension between 3 and 7, by Song [35] in combination with the work of Marques–Neves [24].

Theorem 1.1. *Let N be a compact Riemannian manifold of dimension $n+1$ with $n \geq 2$ and with positive Ricci curvature. Then the Allen–Cahn minmax ([14], see Section 2.1 below) yields on N a multiplicity-1 smooth minimal hypersurface M with* $\dim(\overline{M} \setminus M) \leq n-7$.

Remark 1.2 (Additional consequences). The multiplicity-1 conclusion immediately implies that M is two-sided; in fact, $N \setminus \overline{M}$ is given by two disjoint open sets whose common boundary is $\overline{M}$. It also follows easily that M is connected and has Morse index 1.

To obtain the multiplicity-1 result of Theorem 1.1 we exploit directly the minmax characterization (rather than finite index properties). Recall that the Allen–Cahn energy $\mathcal{E}_\varepsilon$ involves a small parameter $\varepsilon > 0$ and the desired minimal hypersurface appears by taking a suitable (subsequential) limit, as $\varepsilon \to 0^+$, of varifolds naturally associated to the minmax critical points $u_\varepsilon \in W^{1,2}(N)$ constructed in [14]. (Heuristically, a diffused interface is constructed from weighted level sets of u_ε, following [16].) The minimal hypersurface is obtained in the $\varepsilon \to 0^+$ limit as a stationary integral varifold. Exploiting the fact that u_ε has Morse index at most 1, the varifold turns out to be smooth away from a singular set of codimension 7 or higher, ultimately thanks to Tonegawa and Wickramasekera's works [36, 37, 40]. We will not directly analyse the Allen–Cahn solutions u_ε constructed in [14] that concentrate on the minimal hypersurface. We will only retain the following information on these solutions: the minmax characterisation of u_ε, the fact that the minmax values $c_\varepsilon = \mathcal{E}_\varepsilon(u_\varepsilon)$ converge to the mass of the varifold as $\varepsilon \to 0^+$, and the smoothness properties of the varifold. We then prove the following result (see Section 1.1 for a sketch of the argument), from which Theorem 1.1 is easily deduced.

Theorem 1.3. *Let N be a compact Riemannian manifold of dimension $n+1$, $n \geq 2$, and with positive Ricci curvature. Let $M \subset N$ be any smooth minimal hypersurface such that* $\dim(\overline{M} \setminus M) \leq$ *$n-7$, M is stationary in N, and for every $x \in \overline{M}$ there exists a geodesic ball in N centred at x in which M is stable. Then the minmax value c_ε obtained by [14] (for $\varepsilon < 1$) satisfies*

$$\limsup_{\varepsilon \to 0^+} c_\varepsilon < 2\,\mathcal{H}^n(M).$$

Remark 1.4. The assumptions on M in Theorem 1.3 are valid for any minimal hypersurface whose closure is the support of a varifold produced by the minmax in [14, 16, 36, 37, 40]). Then it is readily checked that Theorem 1.1 follows from Theorem 1.3.

Remark 1.5. It is not hard to check that, under the assumptions of Theorem 1.1, the area of the minmax hypersurface is less than or equal to that of an arbitrary two-sided minimal hypersurface in N that has the properties listed for M in Theorem 1.3.

Remark 1.6. While $\mathrm{Ric}_N \geq 0$ on N would not suffice for our multiplicity-1 conclusion, the assumption $\mathrm{Ric}_N > 0$ in Theorems 1.1 and 1.3 can be weakened. Denoting by $\{\mathrm{Ric}_N = 0\}$ the set where the Ricci curvature is 0 in at least one direction, the curvature hypothesis can be relaxed by assuming $\mathrm{Ric}_N \geq 0$ on N and, additionally, one of the following: (i) $\mathcal{H}^n(\{\mathrm{Ric}_N = 0\}) = 0$ or (ii) $\{\mathrm{Ric}_N = 0\} \subset \bigcup_{i=1}^\infty A_i$ where A_i's are pairwise disjoint open sets, each having smooth mean-convex boundary, with mean curvature pointing towards the interior of A_i. (See Remark 8.2.)

Remark 1.7. As the assumption $\mathrm{Ric}_N > 0$ is only used at specific points in the proof (summarised in Remark 8.2), some ideas developed here could be employed more widely. For example, an argument in [5] is inspired by the present work, and analogues of Theorem 1.1 and of Theorem 1.8 below are obtained in [4] for $2 \leq n \leq 6$ when N is endowed with a bumpy metric.

While the Allen–Cahn and Almgren–Pitts frameworks are different in spirit (see also Remarks 1.10 and 1.11), Theorem 1.1 could be viewed as an Allen–Cahn counterpart of the combined results obtained in [19, 29, 41, 42] for the Almgren–Pitts minmax. In [19] Ketover–Marques–Neves show (relying also on [41]) that, when N^{n+1} is orientable with positive Ricci curvature and $2 \leq n \leq 6$, the minimal hypersurface is two-sided and has multiplicity 1. This result is extended to $n \geq 7$ by Ramírez-Luna in [29] (relying on [42]). Recalling Remark 1.2, Theorem 1.1 provides an alternative route to following the existence result for *two-sided* minimal hypersurfaces, also obtained in [19, 29].

Theorem 1.8. *In any compact Riemannian manifold of dimension $n+1$ with $n \geq 2$ and with positive Ricci curvature there exists a smooth* two-sided *minimal hypersurface M with* $\dim(\overline{M} \setminus M) \leq n-7$.

Remark 1.9. The curvature hypothesis in Theorem 1.8 can be weakened in one of the ways described in Remark 1.6.

1.1 | Strategy

We now outline the proof of Theorem 1.3. Given M as in Theorem 1.3, the idea is to produce, for all sufficiently small ε, a continuous path in $W^{1,2}(N)$ that joins the constant -1 to the constant $+1$ and such that the Allen–Cahn energy evaluated along the path stays below $2\mathcal{H}^n(M)$ by a fixed positive amount independent of ε (determined only by geometric properties of $M \subset N$). Since this is an admissible path for the minmax in [14] (see also Section 2.1), the inequality in Theorem 1.3 must hold.

The construction of the path is geometric in flavour and employs classical tools (coarea formula, semilinear PDE theory). For simplicity, in this introduction we illustrate it mainly in the case $2 \leq n \leq 6$, so that M is smooth and closed. We think of M with multiplicity 2 as an immersed two-sided hypersurface, namely, its double cover $\widetilde{M}$ with the standard projection. This immersion, that we denote by $\iota : \widetilde{M} \to N$, is minimal and unstable (by the positiveness of the Ricci curvature). It is possible to find a (sufficiently small) geodesic ball $B \subset M$ such that the lack of stability still holds for deformations that do not move B (this follows by a capacity argument). We then find a deformation of ι that is area-decreasing on some time interval $[0, t_0]$ and that does not move B. This deformation is depicted in the top row of Figure 1. (We can choose the initial speed of the deformation to be nonnegative on $\widetilde{M}$; therefore the deformation "pushes away from M".) We denote by $2\mathcal{H}^n(M) - \tau$ the area of the immersion at time t_0 for some $\tau > 0$. If we cut out B from M we are left with an immersion with boundary, namely $\iota|_{\widetilde{M} \setminus \iota^{-1}(B)}$. We can restrict the previous deformation to $\iota|_{\widetilde{M} \setminus \iota^{-1}(B)}$, obtaining an area-decreasing deformation (at fixed boundary) on the time interval $[0, t_0]$. This time the area changes from $2\mathcal{H}^n(M) - 2\mathcal{H}^n(B)$ to $2\mathcal{H}^n(M) - 2\mathcal{H}^n(B) - \tau$. This deformation is depicted in the middle row of Figure 1. Now we proceed to close the hole at B continuously (bottom row of Figure 1), reaching, say in time 1, the same immersion depicted in the top-right picture of Figure 1. It is helpful to think of closing the hole at B by inserting a weighted copy of B and letting the real-valued weight increase continuously from 0 to 2. (Abusing language, we will talk in this introduction of immersions also to indicate these "weighted immersions".) The

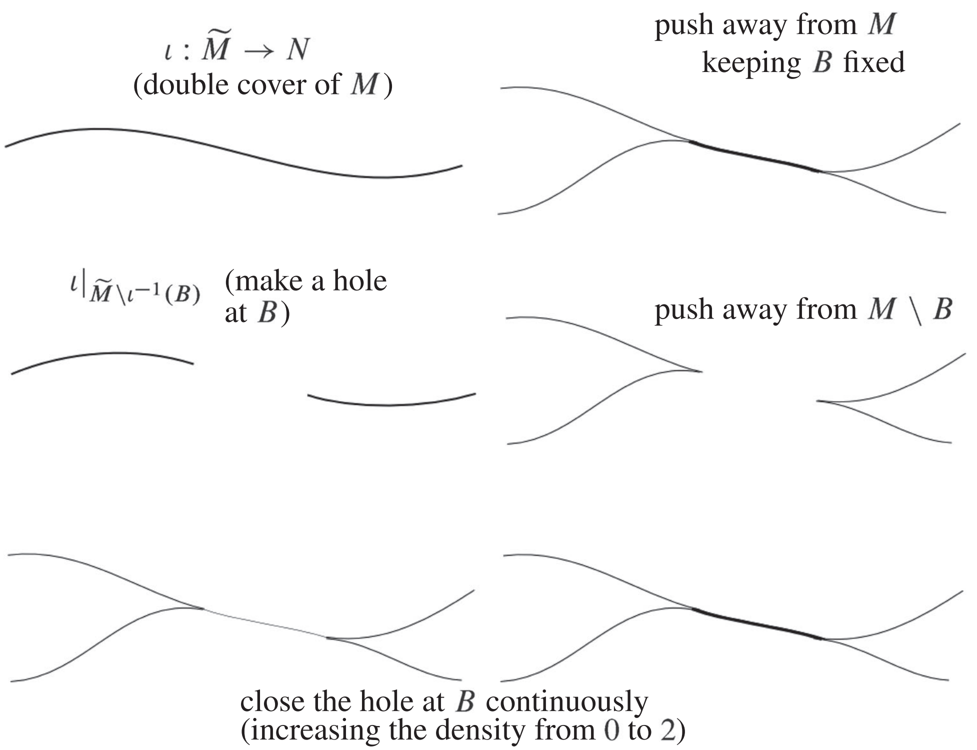

FIGURE 1 Cut, deform, fill in. The path of "immersions" in the second and third row reaches the same immersion depicted in the top-right picture.

area increases from $2\mathcal{H}^n(M) - 2\mathcal{H}^n(B) - \tau$ to $2\mathcal{H}^n(M) - \tau$. Therefore, in going from the middle-left picture to the bottom-right picture of Figure 1, we have produced a "path of immersions" along which the area stays strictly below $2\mathcal{H}^n(M)$, at least by $\min\{\tau, 2\mathcal{H}^n(B)\}$, a fixed positive amount that only depends on the geometry of $M \subset N$.

This path of immersions is then "reproduced at the Allen–Cahn level", that is, replaced by a continuous path $\gamma : [0, t_0 + 1] \to W^{1,2}(N)$. Each function in the image of this curve is a suitable "Allen–Cahn approximation" of the corresponding immersion. To construct this, one fits one-dimensional Allen–Cahn solutions in the normal bundle to the immersion, respecting multiplicities: at points with multiplicity 1 and 2 we will use, respectively, the top and bottom profiles in Figure 2. The image of the immersion corresponds to points where the function transitions between -1 and $+1$, with a double transition for points of multiplicity 2. The operation of closing the hole at B can be reproduced at the Allen–Cahn level thanks to the multiplicity-2 assumption on B: in the normal direction to B, the profile of the function goes from being constantly -1 to looking like the bottom picture in Figure 2, employing the continuous family of profiles depicted in Figure 3 (going from the last to the first picture). Moreover, this operation is continuous in $W^{1,2}(N)$. (Working in the Allen–Cahn framework, hypersurfaces are replaced by weighted level sets of functions and are thus naturally diffused, so continuous weights are allowed. This ultimately permits the geometric operation of closing the hole by increasing the weight of B continuously from 0 to 2. The analytic ingredient behind the implementation of such a geometric operation is the possibility to vary, as in Figure 3, the *values* of the function whose level sets give rise to the diffused hypersurface. This geometric effect cannot be obtained by composing

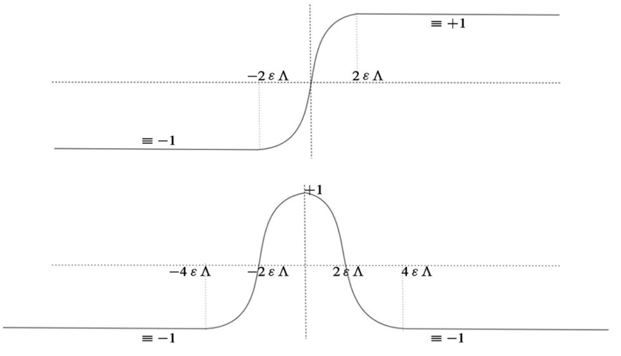

FIGURE 2 The (smooth) functions $\overline{\mathbb{H}}^{\varepsilon}$ (top) and $\Psi = \Psi_0$ (bottom), with $\Lambda = 3|\log \varepsilon|$.

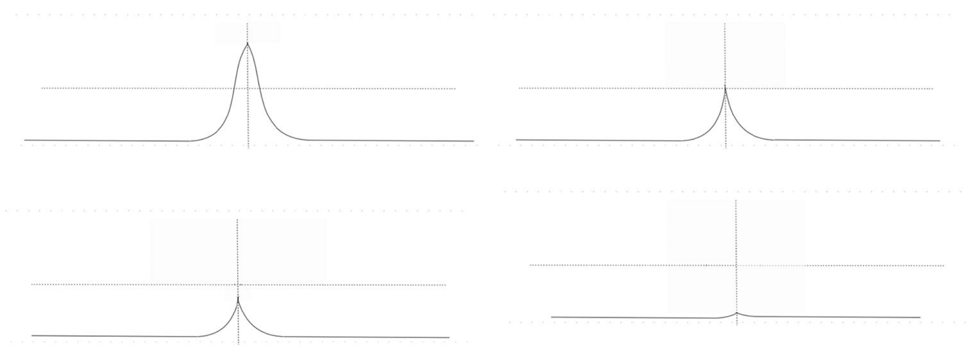

FIGURE 3 The profiles ψ_t, depicted for $t \in (0, 4\,\varepsilon\,\Lambda)$; see (3).

the function with a domain deformation.[2]) The construction of γ is done for all sufficiently small ε (the parameter of the Allen–Cahn energy) and, moreover, for all sufficiently small ε the Allen–Cahn energy all along γ is a close approximation of the area of the corresponding immersions; therefore, for all sufficiently small ε, the energies stay below $2\mathcal{H}^n(M)$ by a fixed "geometric" amount $\approx \min\{\tau, 2\mathcal{H}^n(B)\}$.

We now consider $\gamma(0)$ and $\gamma(t_0 + 1)$ (respectively, the Allen–Cahn approximations of the immersions in the middle-left and top-right picture of Figure 1). For the latter, we use a (negative) Allen–Cahn gradient flow (to which we add a small forcing term, infinitesimal in ε). We build a mean-convex barrier (by writing a suitable Allen– Cahn approximation of ι), that sits below $\gamma(t_0 + 1)$. Thanks to this, we show that the flow deforms $\gamma(t_0 + 1)$ continuously into a stable Allen– Cahn solution, which has to be the constant $+1$ by the Ricci-positive assumption. Along this flow, the Allen–Cahn energy is controlled by the initial bound $\approx 2\mathcal{H}^n(M) - \tau$. The function $\gamma(0)$ is $\approx +1$ close to $M \setminus B$ and ≈ -1 away from a tube around $M \setminus B$: we connect this function explicitly to the constant -1, continuously in $W^{1,2}$, with approximately decreasing Allen–Cahn energy. This is again possible thanks to the profiles in Figure 3. (A close geometric operation is to give weight 2 to $M \setminus B$ and let the real-valued weight decrease continuously to 0.) Reversing

[2] In a similar spirit, when we will write an Allen–Cahn approximations of an immersion with boundary, there will be no sharp transition of multiplicity at the boundary: the weight will instead continuously decrease to 0 in a neighbourhood of the boundary of the hypersurface.

the latter path, composing it with γ and then with the path obtained via the flow, we produce the promised continuous path in $W^{1,2}(N)$ that joins -1 to $+1$ and has the desired energy control.

We stress that the functions $\gamma(t)$, $t \in [0, t_0 + 1]$, that we call "Allen–Cahn approximations" of the corresponding immersions, are not solutions of an Allen–Cahn equation, even when they are built from minimal hypersurfaces; they only realize the "correct" energy value. In fact, we do not even analyse the Allen–Cahn first variation of $\gamma(t)$. The loss of information on the first variation is compensated by the ad hoc structure of the Allen–Cahn approximation: its level sets are by construction graphical over the given immersed hypersurface, so that the Allen–Cahn energy is an effective approximation of area (by the coarea formula) and the geometric information can be translated to the Allen–Cahn level.

We digress to comment briefly on the operation of connecting $\gamma(0)$ to the constant -1. We could in fact use an Allen–Cahn flow for this step, by first slightly deforming $\gamma(0)$ into another function (with a similar profile, so that it still approximates $2|M \setminus B|$, but with a more effective first variation) and then running the Allen-Cahn flow, which deforms this function to the constant -1. We do not argue in this way, since we are able to produce an explicit deformation of $\gamma(0)$ to -1, which is elementary and straightforward, thanks to the profiles in Figure 3. We stress, however, that the deformation that we exhibit mimics the Allen–Cahn flow, and can be viewed as a regularized version of the Brakke flow that starts at $2|M \setminus B|$ and vanishes instantaneously. While the Brakke flow creates a discontinuity in space-time, at the Allen–Cahn level we gain continuity (and the flow reaches -1 in time $O(\varepsilon \,|\log \varepsilon\,|)$). As we mentioned above, an intuitive geometric counterpart of the deformation connecting $\gamma(0)$ to -1, is the one that continuously decreases the weight of $M \setminus B$ from 2 to 0 in time $O(\varepsilon \,|\log \varepsilon\,|)$.

A remark in similar spirit can be made for the portion of path that "closes the hole at B". At the Allen–Cahn level we gain continuity for this operation, because the framework allows (heuristically) to increase the weight of B from 0 to 2 continuously. More precisely, with the parametrization that we employ (that takes time 1), if we were to take the $\varepsilon \to 0$ limit for this portion of path, we would see indeed a continuous increase of the density on B from 0 to 2 (going from the bottom-left picture to the bottom-right picture of Figure 1). We could have alternatively parametrized this portion of path by employing the same one-dimensional profiles in the normal direction to B, however parametrized at faster speed (as in (3)), in order to mimic a reversed Allen–Cahn flow on $\mathbb{R}$: in this case this portion of path would take time $O(\varepsilon \,|\log \varepsilon\,|)$, and if we were to take the $\varepsilon \to 0$ limit we would see the sudden appearance of $2|B|$.

We emphasise the following point of view on the construction of the path (connecting -1 to $+1$) that was sketched above. Consider $\iota : \widetilde{M} \to N$: we exhibit two one-sided deformations that decrease area and that can be reproduced for the Allen–Cahn approximations. One (from the top-left to the middle-left picture of Figure 1) has the geometric effect of removing $2|B|$. The other (from the top-left to the top-right picture of Figure 1) is a deformation of ι as an immersion, induced by an initial velocity compactly supported away from B. We will denote by G_0^ε in Section 4 the Allen–Cahn approximation of ι. Then, with reference to Figure 4, and using the notation $\gamma(0)$, $\gamma(t_0 + 1)$, respectively, for the Allen–Cahn approximations of the immersions in the middle-left and top-right picture of Figure 1, the two deformations just described, implemented at the Allen–Cahn level, correspond, respectively, to "going from G_0^ε to $\gamma(0)$" and "going from G_0^ε to $\gamma(t_0 + 1)$". The two deformations are linearly independent, as the former acts on a compact set containing B while the latter acts in the complement of this compact set. Note that it may well be that ι is an immersion with Morse index 1 (e.g., the double cover of an equator of $\mathbb{RP}^3$). The area-decreasing deformation that removes B is clearly not a deformation of ι as an immersion; it can be reproduced as a continuous deformation at the Allen–Cahn level thanks to the fact that multiplicity is 2 on

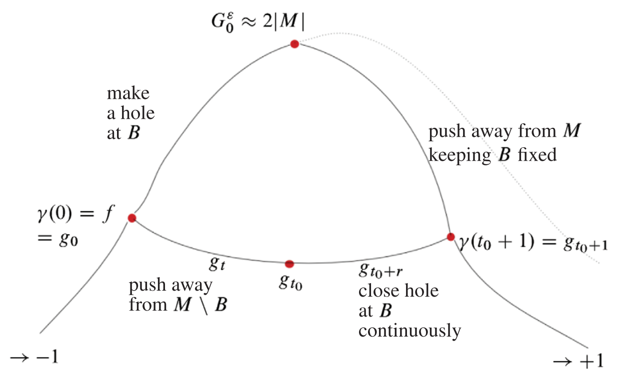

FIGURE 4 Lowering the peak (landscape for the Allen–Cahn energy). The same labels as in Figure 1 are used, to denote deformations that reproduce those in Figure 1.

B, so that the profile of G_0^ε in the normal bundle to B looks like the bottom one in Figure 2; this profile can be connected continuously to the constant -1 with controlled energy, employing the deformation depicted in Figure 3.

The function $\gamma(0)$ (that will be denoted by $f = g_0$ in Section 7) can be connected to the constant -1, and the function $\gamma(t_0 + 1)$ can be connected to the constant $+1$, as described in the sketch given earlier. We thus have a "recovery path" for the value $2\mathcal{H}^n(M)$: this path connects -1 to $+1$ (passing through G_0^ε) and the maximum of the Allen–Cahn energy along this path is $\approx 2\mathcal{H}^n(M)$. What we achieve is to deform this path in the surroundings of G_0^ε, exploiting the information that we have gained on the landscape; specifically, we deform the portion between $\gamma(0)$ and $\gamma(t_0 + 1)$. From $\gamma(0)$ we use a deformation that reproduces the one in the middle row of Figure 1. By doing this we reach a function g_{t_0} (notation as in Section 7). Now we close the hole continuously, replicating the deformation in the bottom row of Figure 1, reaching the function $\gamma(t_0 + 1)$ (which will be denoted by g_{t_0+1} in Section 7). We have thus found a path $\gamma : [0, t_0 + 1] \to W^{1,2}(N)$, from $\gamma(0)$ to $\gamma(t_0 + 1)$, that lowers the peak, compared to the initial "recovery path". This follows thanks to the fact that the Allen–Cahn energy is a close approximation of the area of the corresponding immersion, so we inherit the estimates that we had for the path of immersions that joins the middle-left picture to the bottom-right picture of Figure 1.

This shows that the landscape around G_0^ε is reminiscent of one where the Morse index is ≥ 2. *However*, G_0^ε is not a stationary point for the Allen–Cahn energy. In fact, we never need to compute the Allen–Cahn first or second variation along these deformations; it suffices to know that the Allen–Cahn energy at $\gamma(0)$ and $\gamma(t_0 + 1)$ is strictly less than its value at G_0^ε (by a fixed amount independent of ε). Knowledge of the first variation is only needed for G_0^ε in order to prove that a negative gradient flow connects $\gamma(t_0 + 1)$ to $+1$, for which we employ G_0^ε as a barrier.

Remark 1.10. It is natural to ask whether the path from -1 to $+1$ produced in the earlier sketch can be imitated (e.g., in an Almgren–Pitts framework) by a one-parameter family of boundaries in N. For the portion $\gamma : [0, t_0 + 1] \to W^{1,2}(N)$, rather than increasing the weight of B from 0 to 2 (which cannot be done in the class of boundaries) one can argue by doubling $M \setminus B$ and inserting a small cylindrical neck at B, then pushing this hypersurface away from M without moving the neck (and decreasing the area), then closing the neck. (An operation of this type is analysed in [19]. To avoid confusion, we point out that for our path γ, the nodal sets $\{\gamma_t = 0\}_{t \in [t_0, t_0+1]}$ are *not* cylindrical necks.) It is conceivable that one could then use mean-curvature flow to drift away from M

until extinction time and thus imitate, by using boundaries, the portion of path from $\gamma(t_0+1)$ to $+1$. The use of a flow for this purpose does not appear to have been investigated in the literature. (Gradient flows may be easier to use in the Allen–Cahn framework, since the parabolic problem is semilinear, has long-time existence, and singularities do not appear. This may be particularly true when $n \geq 7$ with singularities present in the geometric initial condition; see Remarks 1.11 and 1.12.) For the portion of path that goes from $\gamma(0)$ to the constant -1, the spirit of the Allen–Cahn deformation is again very different than a deformation of boundaries (compare with [19, 41]), since its geometric analogues are either a continuous weight decrease from 2 to 0 or a Brakke flow that instantaneously makes $M \setminus B$ disappear. The Allen–Cahn framework allows a very straightforward way to produce this portion of path. (Some extra challenges have to be overcome in [19], for example, the catenoid estimate.)

For $n \geq 7$, we still employ the idea illustrated in low dimensions. Its implementation, however, is rendered somewhat harder by the presence of the singular set: standard tubular neighbourhoods and Fermi coordinates for M (that are essential to fit one-dimensional Allen–Cahn profiles in the normal bundle to M) are not available. While the geometric ideas remain the same as in the low-dimensional case, we need to additionally study certain analytic properties. Denote by $d_{\overline{M}} : N \to [0,\infty)$ the distance function to $\overline{M}$. The value $d_{\overline{M}}(x)$ is always realized by a geodesic (possibly more than one) from x to a *smooth* point of $\overline{M}$. This allows us to analyse the cut-locus of $d_{\overline{M}}$ (restricting to $\{0 < d_{\overline{M}} < \mathrm{inj}(N)\}$), following [21], and obtain n-rectifiability properties for it. This leads (for the moment) to the existence of a suitable replacement for Fermi coordinates, which becomes the usual one on any compact subset of $\widetilde{M}$. Denote by $\iota : \widetilde{M} \to N$ the immersion given by the standard projection from the double cover of M. We choose $K \subset \widetilde{M}$ compact (sufficiently large) and a geodesic ball $B \subset \iota(K)$ (sufficiently small) so that $\iota : \widetilde{M} \to N$ admits a deformation as an immersion that decreases area and only moves $K \setminus \iota^{-1}(B)$. (This is analogous to what we did in the lower-dimensional case, except that this time we additionally need a deformation that does not move M close to the singular set.) The set K will play the role that was of $\widetilde{M}$ in the low-dimensional case. Around $\iota(K)$ we define Allen–Cahn approximations of suitable immersions by fitting one-dimensional Allen–Cahn profiles in the normal bundle. Away from $\iota(K)$, we use the level sets of $d_{\overline{M}}$ to complete the definition of the desired Allen–Cahn approximations and create (as in the low-dimensional case) a continuous path $\gamma : [0, t_0+1] \to W^{1,2}(N)$ with controlled energy. Exploiting further the n-rectifiability of the cut-locus, we analyse the singular part of $\Delta d_{\overline{M}}$ and (using also the Ricci-positive condition) we obtain that, restricting to $\{0 < d_{\overline{M}} < \mathrm{inj}(N)\}$, the distributional Laplacian of $d_{\overline{M}}$ is a negative Radon measure. This translates into a mean-convexity property for the Allen–Cahn approximation G_0^ε of $\iota : \widetilde{M} \to N$. With a (slightly nonstandard) smoothing operation, we obtain from G_0^ε a smooth barrier m that is still mean-convex for the negative Allen–Cahn gradient flow (as for $2 \leq n \leq 6$, we add an infinitesimal forcing term). By employing m we produce the part of the path that connects $\gamma(t_0+1)$ to the constant $+1$.

Remark 1.11. The continuity of the path from $\gamma(0)$ to $\gamma(t_0+1)$, its energy bounds, and the mean-convexity of G_0^ε ultimately rest on the fact that almost every level set of the distance function $d_{\overline{M}}$ is almost everywhere smooth, with mean curvature pointing away from M. These properties only require classical arguments. The almost everywhere information is sufficient for our purposes, because in the Allen–Cahn framework hypersurfaces are "diffused". For contrast, in the case of boundaries of Caccioppoli sets, all level sets of $d_{\overline{M}}$ have to be analysed; compare [29, prop. 2.2].

Remark 1.12. The almost everywhere properties at the previous remark are sufficient to set up a mean-convex Allen–Cahn flow starting at G_0^ε. For $n \geq 7$ this initial condition is built from a singular hypersurface. We expect that the $\varepsilon \to 0$ limit of these Allen-Cahn flows gives rise to an ancient (mean-convex) mean curvature flow with initial condition (at time $-\infty$) given by the singular minimal immersion $\iota : \widetilde{M} \to N$.

1.2 | Structure of the paper (and remarks for $n \leq 6$)

Except for properties of the distance function borrowed from [21] (in Section 3 we point out the relevant modifications needed to handle the singular set), the proof is self-contained.

After the preliminary Section 2, we begin the proof of Theorem 1.3, which we write for $n \geq 7$, assuming the existence of a (nonempty) singular set $\overline{M} \setminus M$ of dimension $\leq n-7$. While the underlying ideas are the same for all dimensions, the proof becomes considerably shorter and more straightforward in the absence of a singular set, in particular when $n \leq 6$. In detail, Sections 3 and 4, in which we study the distance function to $\overline{M}$ and its level sets, can be omitted when $\overline{M} = M$, and one can use standard facts about tubular neighbourhoods of smooth closed hypersurfaces. In Section 5 we identify a large unstable region $2|K \setminus B|$ and in Section 6 the immersions that will be relevant for the construction of the path. The compact set K that we need to work with in Sections 5 and 6 can be replaced simply by $\widetilde{M}$ when $\overline{M} = M$, and in this case the definitions of the Allen–Cahn approximations of the relevant immersions given in Section 7 become simpler. In Section 7.5 we construct a barrier m by suitably mollifying a Lipschitz function G_0^ε, which is defined from the level sets of $d_{\overline{M}}$ and is an Allen–Cahn approximation of $\iota : \widetilde{M} \to N$. This convolution procedure (described in Appendix A) ensures smoothness and mean-convexity of m, which is important for our arguments. If $\overline{M} = M$, G_0^ε is already smooth and mean-convex and no smoothing is needed, so Appendix A and part of Section 7.5 can be omitted. In Section 8 we complete the proof of Theorem 1.3, and subsequently of Theorems 1.1 and 1.8.

2 | PRELIMINARIES

We give a brief summmary of [14], then introduce the one-dimensional Allen–Cahn profiles that will be needed for our proof.

2.1 | Reminders: Allen–Cahn minmax approximation scheme

We recall the minmax construction in [14]. For $\varepsilon \in (0,1)$ consider the functional

$$\mathcal{E}_\varepsilon(u) = \frac{1}{2\sigma}\int_N \varepsilon\,\frac{|\nabla u|^2}{2} + \frac{W(u)}{\varepsilon}$$

on the Hilbert space $W^{1,2}(N)$. Here W is a C^3 “double well” potential, with exactly three critical points, two nondegenerate minima at ± 1 and a local maximum at 0, with (exactly) two zeroes of W'' (one between -1 and 0, one between 0 and 1) and with quadratic growth to ∞ at $\pm\infty$; the normalisation constant σ is $\sigma = \int_{-1}^{1}\sqrt{W(t)/2}\,dt$. A standard choice of potential is $W(x) = \frac{(1-x^2)^2}{2}$, suitably modified (to have quadratic growth) outside $[-2,2]$. Consider continuous paths in $W^{1,2}(N)$ that start at the constant -1 and end at the constant $+1$: this is the class of admissible paths. A "wall" (or "mountain pass") condition is ensured and yields the existence of a minmax solution u_ε to ${\mathcal{E}_\varepsilon}'(u_\varepsilon) = 0$. Moreover, upper and lower energy bounds are established, uniformly in ε. (We recall that ${\mathcal{E}_\varepsilon}'(u) = -\varepsilon\,\Delta u + \frac{W'(u)}{\varepsilon}$, where Δ is the Laplace-Beltrami operator, so the Euler-Lagrange equation ${\mathcal{E}_\varepsilon}'(u) = 0$ is elliptic semilinear.)

In order to produce a stationary varifold, one considers $w_\varepsilon = \Phi(u_\varepsilon)$ as in [16], with $\Phi(s) = \int_0^s \sqrt{W(t)/2}\,dt$, and defines the n-varifolds

$$V^\varepsilon(A) = \frac{1}{\sigma}\int_{-\infty}^{\infty} V_{\{w_\varepsilon = t\}}(A)dt.$$

The analysis in [16] (which only requires the stationarity of u_ε and no assumption on their second variation), together with the upper and lower bounds for $\mathcal{E}_\varepsilon(u_\varepsilon)$, gives that V^ε converges subsequentially, as $\varepsilon \to 0$, to an integral n-varifold $V \neq 0$ with vanishing first variation.

Thanks to the fact that the Morse index of u_ε is ≤ 1 for all ε, [14] reduces the problem locally in N to one that concerns stable Allen– Cahn solutions, as in [36]. For these, the regularity theory of [37, 40] applies and gives that spt $\|V\|$ is smoothly embedded away from a possible singular set of dimension $\leq n-7$; that is, V is the varifold of integration over a finite set of minimal hypersurfaces, each counted with integer multiplicity: $V = \sum_{j=1}^K q_j |M_j|$, with $q_j \in \mathbb{N}$ and M_j minimal and smooth with $\dim(\overline{M_j}\setminus M_j) \leq n-7$ ($|M_j|$ denotes the multiplicity-1 varifold of integration on M_j). In the case $n \leq 6$ all the M_j's are closed (and smooth). (In the case $\mathrm{Ric}_N > 0$ there is only one connected component, $K = 1$; see Remark 8.1.)

We point out that, denoting by ε_i the sequence extracted to guarantee the varifold convergence, $\mathcal{E}_{\varepsilon_i}(u_{\varepsilon_i}) \to \|V\|(N)$ in this construction, in other words the Allen–Cahn energy of u_{ε_i} converges to the mass $\sum_{j=1}^K q_j \mathcal{H}^n(M_j)$ of V.

2.2 | One-dimensional profiles

Let $\mathbb{H}(r)$ denote the monotonically increasing solution to $u'' - W'(u) = 0$ such that $\lim_{r\to\pm\infty}\mathbb{H}(r) = \pm 1$, with $\mathbb{H}(0) = 0$. (For the standard potential $\frac{(1-x^2)^2}{2}$ we have $\mathbb{H}(r) = \tanh(r)$.) Then also $\mathbb{H}(-r)$ and $\mathbb{H}(\pm r + z)$ solve $u'' - W'(u) = 0$ (for any $z \in \mathbb{R}$). The rescaled function $\mathbb{H}_\varepsilon(r) = \mathbb{H}(\frac{r}{\varepsilon})$ solves $\varepsilon u'' - \frac{W'(u)}{\varepsilon} = 0$.

Truncations. The arguments developed here will involve the construction of suitable Allen–Cahn approximations of certain immersions. For that purpose, we will make use of approximate versions of $\mathbb{H}_\varepsilon$. While this introduces small errors in the corresponding ODEs, it has the advantage that the approximate solutions are constant (± 1) away from an interval of the form $[-6\,\varepsilon\,|\log\varepsilon|, 6\,\varepsilon\,|\log\varepsilon|]$. An Allen–Cahn approximation of a hypersurface in N requires fitting the 1-dimensional profiles in the normal direction to the hypersurface, and we need to stay inside a tubular neighbourhood, so it is effective to have one-dimensional profiles that become constant before we reach the boundary of the tubular neighbourhood.

The cutoff for the heteroclininc $\mathbb{H}$ is done as follows (this truncation is also used in [6, 39]): for $\Lambda = 3|\log\varepsilon|$ define

$$\overline{\mathbb{H}}(r) = \chi(\Lambda^{-1}r)\mathbb{H}(r) \pm (1 - \chi(\Lambda^{-1}r)),$$

where $\pm$ is chosen, respectively, on $r > 0$ and $r < 0$, and χ is a smooth bump function that is $+1$ on $(-1, 1)$ and has support equal to $[-2, 2]$. With this definition, $\overline{\mathbb{H}} = \mathbb{H}$ on $(-\Lambda, \Lambda)$, $\overline{\mathbb{H}} = -1$ on $(-\infty, -2\Lambda]$, and $\overline{\mathbb{H}} = +1$ on $[2\Lambda, \infty)$. Moreover, $\overline{\mathbb{H}}$ satisfies (as we check below) $\|\overline{\mathbb{H}}'' - W'(\overline{\mathbb{H}})\|_{C^2(\mathbb{R})} \leq C\,\varepsilon^3$ for $C > 0$ independent of ε. Note that $\overline{\mathbb{H}}'' - W'(\overline{\mathbb{H}}) = 0$ away from $(-2\Lambda, -\Lambda) \cup (\Lambda, 2\Lambda)$, so it suffices to compute on $(-2\Lambda, -\Lambda) \cup (\Lambda, 2\Lambda)$:

$$\overline{\mathbb{H}}''(r) = \Lambda^{-2}\chi''(\Lambda^{-1}r)(\mathbb{H}(r) \mp 1) + 2\Lambda^{-1}\chi'(\Lambda^{-1}r)\mathbb{H}'(r) + \chi(\Lambda^{-1}r)\mathbb{H}''(r).$$

We have $\|\mathbb{H}(r) \mp 1\|_{C^0} \leq c\,\varepsilon^\alpha$ and $\|\mathbb{H}'(r)\|_{C^3} \leq c\,\varepsilon^\alpha$ for some $\alpha \geq 6$, and $c > 0$ depending only on W. This can be done by an explicit check for the standard potential (e.g., when $r > 0$ we must estimate $1 - \tanh(r) = \frac{2e^{-2r}}{1+e^{-2r}}$ for $r > -3\log\varepsilon$) and is true whenever W is quadratic around the minima by comparison. Therefore on $(-2\Lambda, -\Lambda) \cup (\Lambda, 2\Lambda)$ we get $\|\overline{\mathbb{H}}''\|_{C^2} \leq \tilde{c}\,\varepsilon^3$ for $\varepsilon < 1/2$ and $\tilde{c} > 0$ depending only on W. Similarly, on $(-2\Lambda, -\Lambda) \cup (\Lambda, 2\Lambda)$ one checks that $\|\overline{\mathbb{H}}'\|_{C^2} \leq \tilde{c}\,\varepsilon^3$ for $\varepsilon < 1/2$ and $\tilde{c} > 0$ depending only on W. Moreover, since $W'(\mathbb{H}) = \mathbb{H}''$ and $\mathbb{H} - \overline{\mathbb{H}} = (1 - \chi(\Lambda^{-1}t))(\mathbb{H} \mp 1)$, we find on $(-2\Lambda, -\Lambda) \cup (\Lambda, 2\Lambda)$

$$\begin{aligned}\|W'(\overline{\mathbb{H}})\|_{C^2} \leq \|W'(\mathbb{H})\|_{C^0} + \|W''\|_{C^0}\|\mathbb{H} - \overline{\mathbb{H}}\|_{C^0} \\ + 3\|W''\|_{C^1}(\|\overline{\mathbb{H}}'\|_{C^1} + \|\overline{\mathbb{H}}'\|^2_{C^0})0 \leq c\varepsilon^3.\end{aligned}$$

In conclusion, $\|\overline{\mathbb{H}}'' - W'(\overline{\mathbb{H}})\|_{C^2(\mathbb{R})} \leq C\varepsilon^3$ for some $C > 0$ (depending on W).

Notation. For $\varepsilon < 1$ we rescale these truncated solutions and let $\overline{\mathbb{H}}^\varepsilon(\cdot) = \overline{\mathbb{H}}(\frac{\cdot}{\varepsilon})$.

Computation of the Allen-Cahn energy of $\overline{\mathbb{H}}^\varepsilon$. To compute the energy of $\bar{H}^\varepsilon$, following [17], we have, for any $q : \mathbb{R} \to \mathbb{R}$,

$$\int_{(a,b)} \frac{\varepsilon}{2}|q'|^2 + \frac{W(q)}{\varepsilon} = \int_a^b \frac{1}{2}\left(\sqrt{\varepsilon}q' - \frac{1}{\sqrt{\varepsilon}}\sqrt{2W(q)}\right)^2 + q'\sqrt{2W(q)}.$$

The first term vanishes when $q = \mathbb{H}_\varepsilon$. Let G denote a primitive of $\sqrt{2W(t)}$. For the second term, noting that the integrand is $G(q)'$, we get $G(q(b)) - G(q(a))$. In particular, $\int_{\mathbb{R}} \frac{\varepsilon}{2}|\mathbb{H}_\varepsilon'|^2 + \frac{W(\mathbb{H}_\varepsilon)}{\varepsilon} = G(1) - G(-1) = 2\sigma$. Using the fact that $\mathbb{H}_\varepsilon(-2\,\varepsilon\,\Lambda) = -1 + O(\varepsilon^2)$, we get for $q = \mathbb{H}_\varepsilon$

$$\int_{-\infty}^{-2\,\varepsilon\,\Lambda} \frac{\varepsilon}{2}|q'|^2 + \frac{W(q)}{\varepsilon} = G(-1 + O(\varepsilon^2)) - G(-1) = O(\varepsilon^4) > 0,$$

and similarly $\int_{2\,\varepsilon\,\Lambda}^{\infty} \frac{\varepsilon}{2}|q'|^2 + \frac{W(q)}{\varepsilon} = O(\varepsilon^4) > 0$. Therefore

$$\int_{-2\,\varepsilon\,\Lambda}^{2\,\varepsilon\,\Lambda} \frac{\varepsilon}{2}|\mathbb{H}_\varepsilon'|^2 + \frac{W(\mathbb{H}_\varepsilon)}{\varepsilon} = 2\sigma - O(\varepsilon^4).$$

Recalling the definition of $\overline{\mathbb{H}}^\varepsilon$, we have that

$$\bar{H}^\varepsilon - \overline{\mathbb{H}}^\varepsilon = (1 - \chi(\varepsilon\,\Lambda^{-1}t))(\mathbb{H}^\varepsilon \pm 1)$$

which is controlled by $O(\varepsilon^2)$ in C^2-norm. Therefore

$$2\sigma - O(\varepsilon^2) \le \int_{-2\varepsilon\Lambda}^{2\varepsilon\Lambda} \frac{\varepsilon}{2}|(\overline{\mathbb{H}}^\varepsilon)'|^2 + \frac{W(\overline{\mathbb{H}}^\varepsilon)}{\varepsilon} \le 2\sigma + O(\varepsilon^2). \tag{1}$$

Families of profiles. Define the function $\Psi : \mathbb{R} \to \mathbb{R}$

$$\Psi(r) = \begin{cases} \overline{\mathbb{H}}^\varepsilon(r + 2\varepsilon\Lambda), & r \le 0, \\ \overline{\mathbb{H}}^\varepsilon(-r + 2\varepsilon\Lambda), & r > 0. \end{cases} \tag{2}$$

This function is smooth thanks to the fact that all derivatives of $\overline{\mathbb{H}}^\varepsilon$ vanish at $\pm 2\varepsilon\Lambda$. Define the following evolution for $t \in [0, \infty)$:

$$\Psi_t(r) := \begin{cases} 4\overline{\mathbb{H}}^\varepsilon(r + 2\varepsilon\Lambda - t), & r \le 0, \\ \overline{\mathbb{H}}^\varepsilon(-r + 2\varepsilon\Lambda - t), & r > 0. \end{cases} \tag{3}$$

Note that $\Psi_0 = \Psi$ and $\Psi_t \equiv -1$ for $t \ge 4\varepsilon\Lambda$. For $t \in (0, 4\varepsilon\Lambda)$ the function Ψ_t is equal to -1 for $|r| \ge 4\varepsilon\Lambda - t$. The functions Ψ_t form a family of even, Lipschitz functions, and $\mathcal{E}_\varepsilon(\Psi_t)$ is decreasing in t. Indeed, the energy contribution of the "tails" (± 1) is zero (so the energy is finite), and we have

$$\mathcal{E}_\varepsilon(\Psi_t) = \mathcal{E}_\varepsilon(\Psi) - \frac{1}{2\sigma}\int_{-t}^{t} \varepsilon\frac{|\Psi'|^2}{2} + \frac{W(\Psi)}{\varepsilon}.$$

Note that Ψ and Ψ_t depend on ε; however, we do not make explicit this dependence for notational convenience. The profiles Ψ_t and profiles of the type $\overline{\mathbb{H}}^\varepsilon(\cdot - t)$ will be used within our construction to produce Allen–Cahn approximations of relevant immersions (possibly with boundary), having multiplicity 1 or 2 on their image.

3 | DISTANCE FUNCTION TO $\overline{M}$

Let N be a Riemannian manifold of dimension $n+1$ with $n \ge 2$ and with positive Ricci curvature $\mathrm{Ric}_N > 0$. Let $M \subset N$ be a smooth minimal hypersurface such that $\dim(\overline{M} \setminus M) \le n - 7$, M is stationary in N, and M is locally stable in N, that is, for every point in $\overline{M}$ there exists a geodesic ball centred at the point in which M is stable. These properties are true for the $\varepsilon \to 0$ varifold limit of finite-index Allen–Cahn solutions on N, thanks to the analysis in [14, 16, 36, 37, 40]. The stationarity condition implies the existence of tangent cones at every point in $\overline{M}$. A consequence of the deep sheeting theorem in [33, 40] is that any point of $\overline{M}$ at which one tangent cone is supported on a hyperplane has to be a smooth point.

Let dist_N denote the (unsigned) Riemannian distance on N; we will be interested in the function $d_{\overline{M}} : N \to [0, \infty)$, $d_{\overline{M}}(\cdot) = \mathrm{dist}_N(\cdot, \overline{M})$. Since N is complete, for every x the value $d_{\overline{M}}(x)$ is realized by at least one geodesic from x to $\overline{M}$ (Hopf–Rinow). We recall a few facts that are true of the distance to an arbitrary closed set; see [21, sec. 3].[3] The function $d_{\overline{M}}$ is Lipschitz on N (with

[3] If the closed set is known to be a $C^{1,1}$ submanifold, then the existence of a tubular neighbourhood is guaranteed, in which the nearest point projection is a well-defined map; moreover, if C^2 regularity on the submanifold is assumed, Fermi coor-

constant 1) and locally semiconcave on $N \setminus \overline{M}$, so that its gradient is BV_{loc} on $N \setminus \overline{M}$ (equivalently, the distributional Hessian of $d_{\overline{M}}$ on $N \setminus \overline{M}$ is a Radon measure). We denote by $S_{d_{\overline{M}}}$ the subset of $N \setminus \overline{M}$ where $d_{\overline{M}}$ fails to be differentiable; $S_{d_{\overline{M}}}$ coincides with the set of points in $p \in N \setminus \overline{M}$ for which there exist at least two geodesics from p to $\overline{M}$ whose length realizes $d_{\overline{M}}(p)$. The function $d_{\overline{M}}$ is C^1 on $N \setminus (\overline{M} \cup \overline{S_{d_{\overline{M}}}})$ and $S_{d_{\overline{M}}}$ is countably n-rectifiable (this uses [1]). However, rectifiability is not necessarily true of its closure unless extra hypotheses on the closed set are available; for example, $\overline{S_{d_{\overline{M}}}}$ would be countably n-rectifiable (even in the C^{k-2} sense) if $\overline{M}$ were a C^k submanifold with $k \geq 3$, thanks to [21, sec. 4]. While this statement does not apply immediately in our case due to the presence of the singular set $\overline{M} \setminus M$, the proof in [21, sec. 4] can still be carried out without any change by virtue of the following observation, which can also be found in [13, 42].

Lemma 3.1. *Let $x \in N \setminus \overline{M}$. For any geodesic from x to $\overline{M}$ (whose length realizes $d_{\overline{M}}(x)$), we have that the endpoint y on $\overline{M}$ actually belongs to M.*

Proof. Let γ be any geodesic from x to $\overline{M}$, let y be its endpoint on $\overline{M}$, and fix a point $z \in N$ that lies in the image of γ and such that $\text{dist}_N(z, y) < \text{inj}(N)$. Consider the (open) geodesic ball $B(z) \subset N$ centred at z with radius $\text{dist}_N(z, y)$. Then $\overline{M} \cap B(z) = \emptyset$ (otherwise there would be a shorter curve than γ joining x to $\overline{M}$) and $y \in \overline{M} \cap \partial B(z)$. Since the monotonicity formula holds at all points of $\overline{M}$ (M is stationary in N), we can blow up at y to obtain tangent cones. Then every tangent cone to M at y has to be supported in a (closed) half-space (the complement of the open half-space obtained by blowing up $B(z)$ at y). By [34, chap. 7, theorem 4.5, remark 4.6] every tangent cone to M at y is the hyperplane tangent to $B(z)$ at y, possibly with multiplicity. As pointed out above, the sheeting theorem in [33, 40] implies that y is a smooth point. □

In other words, any geodesic that realizes the distance to $\overline{M}$ has to end at a smooth point, that is, on M (and it meets M orthogonally). This is the key fact that allows us to repeat the arguments in [21, sec. 4] (as we briefly sketch below) and obtain Proposition 3.2. In the rest of this work we will be interested in the set $T_\omega = \{x \in N : \text{dist}_N(x, \overline{M}) < \omega\}$, where ω is chosen in $(0, \text{inj}(N))$, therefore we restrict to this open set for our analysis (even though not strictly necessary for this section).

Proposition 3.2 (as in [21]). *The set $\overline{S_{d_{\overline{M}}}} \cap T_\omega$ is countably n-rectifiable.*[4] *Moreover, $\nabla d_{\overline{M}} \in SBV_{\text{loc}}(T_\omega \setminus \overline{M})$ and the singular part (with respect to $\mathcal{H}^{n+1} \llcorner (T_\omega \setminus \overline{M})$) of the Radon measure $D^2 d_{\overline{M}} \llcorner (T_\omega \setminus \overline{M})$ is supported on $\overline{S_{d_{\overline{M}}}} \cap (T_\omega \setminus \overline{M})$.*

Remark 3.3. Additionally, we have, since M is smooth, that the absolutely continuous part of $D^2 d_{\overline{M}}$ has a smooth density with respect to $\mathcal{H}^{n+1} \llcorner \left(T_\omega \setminus (\overline{M} \cup \overline{S_{d_{\overline{M}}}})\right)$. This density coincides with the pointwise Hessian of $d_{\overline{M}}$.

Sketch: relevant arguments in [21]. Consider the map $F(y, v, t) = \exp_y(tv)$ for $y \in M$ and v a unit vector orthogonal to M at y. For fixed (y, v) the curve $F(y, v, t)$ is a geodesic leaving M

dinates can be used. In our case, due to the presence of the singular set $\overline{M} \setminus M$, one cannot have a tubular neighbourhood of $\overline{M}$.

[4] Even C^k countably n-rectifiable for all k, however we will not need this stronger property.

orthogonally. We will limit ourselves to $t \leq \omega$, since we are only interested in T_ω. If t_0 is sufficiently small (depending on (y, v)) the geodesic $t \in [0, t_0] \to F(y, v, t)$ is the minimizing curve between its endpoint $F(y, v, t_0)$ and $\overline{M}$; equivalently, its length is $d_{\overline{M}}(F(y, v, t_0))$. However, for large enough t_0 the geodesic may fail to be minimizing; therefore one can consider $\sigma = \sigma_{y,v} \in (0, \omega]$ defined as follows: $\sigma = \omega$ if $F(y, v, t)$ is minimizing (between its endpoint and $\overline{M}$) for all $t \leq \omega$; otherwise, σ is chosen in $(0, \omega)$ so that $F(y, v, t)$ is minimizing (between its endpoint and $\overline{M}$) for $t \leq \sigma$, and $F(y, v, t)$ is not minimizing if $t > \sigma$. The set of points

$$\mathrm{Cut}(M) = \{F(y, v, \sigma_{(y,v)}) : y \in M, v \in (T_y M)^\perp, |v| = 1, \sigma_{(y,v)} < \omega\}$$

is the restriction to T_ω of the so-called cut-locus of M, and it is a subset of $T_\omega \setminus \overline{M}$ whose closure in T_ω does not intersect M. Recall that the unit sphere bundle of M is just $\widetilde{M}$, the oriented double cover of M, so we will also write $(y, v) \in \widetilde{M}$.

Standard theory of geodesics (e.g., [31, chap. 2, lemma 4.8, and chap. 3, lemma 2.11], which give the analogue of [21, prop. 4.7] for M), gives that if $x \in \mathrm{Cut}(M)$, then at least one of the following two conditions holds: (a) there exist (at least) two distinct geodesics from x to M that realize $d_{\overline{M}}(x)$; (b) the map $F : \widetilde{M} \times (0, \omega) \to T_\omega$ has noninvertible differential at $(y, v, \sigma_{(y,v)})$, where $x = F(y, v, \sigma_{(y,v)})$. Conversely, if (a) or (b) holds, then the geodesic $t \to F(y, v, t)$ cannot be minimal on $t \in [0, t_0]$ when $t_0 \in (\sigma_{(y,v)}, \omega)$. Option (a) is equivalent to $x \in S_{d_{\overline{M}}} \cap (T_\omega \setminus \overline{M})$ (see [21, prop. 3.7]).

Using these facts, the arguments of [21, prop. 4.8] adapt to give that

$$\overline{S_{d_{\overline{M}}}} \cap (T_\omega \setminus \overline{M}) = \mathrm{Cut}(M);$$

therefore in order to prove the rectifiability in Proposition 3.2 it suffices (since $S_{d_{\overline{M}}}$ is countably n-rectifiable, see above) to show that $\mathrm{Cut}(M) \setminus S_{d_{\overline{M}}}$ is a countably n-rectifiable set in $T_\omega \setminus \overline{M}$, that is, the analogue of [21, theorem 4.11]. Note that $\overline{S_{d_{\overline{M}}}} \cap \overline{M} \subset \overline{M} \setminus M$ is $\mathcal{H}^n$-negligible, so it does not affect rectifiability. The points in $\mathrm{Cut}(M) \setminus S_{d_{\overline{M}}}$ are characterised by the validity of option (b) above, and the arguments in [21] are local around the points $(y, v, \sigma_{y,v}) \in \widetilde{M} \times (0, \omega)$, so they apply verbatim to our case.

Once the countable n-rectifiability of $\overline{S_{d_{\overline{M}}}}$ has been obtained, it follows that $\nabla d_{\overline{M}}$ is $SBV_{\mathrm{loc}}(T_\omega \setminus \overline{M})$. Indeed, we know to begin with (see above for these statements about the distance to a closed set) that $\nabla d_{\overline{M}}$ is in $BV_{\mathrm{loc}}(T_\omega \setminus \overline{M})$ and notice that $d_{\overline{M}}$ is C^2 (even C^k for all k) on $T_\omega \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$ thanks to the smoothness of M. The "Cantor part" of the Radon measure $D^2 d_{\overline{M}}$ gives 0 measure to countably n-rectifiable sets (see [3, prop. 3.92 or prop. 4.2]); in particular, it gives 0 to $\overline{S_{d_{\overline{M}}}}$. The smoothness of $D^2 d_{\overline{M}}$ in $T_\omega \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$ then implies that there is no "Cantor part", that is, $\nabla d_{\overline{M}}$ is $SBV_{\mathrm{loc}}(T_\omega \setminus \overline{M})$. This concludes the sketch of proof of Proposition 3.2.

Remark 3.4 (on the diffeomorphism F). We point out a couple of further facts, mainly adapted from [21], for future reference. The level sets of $d_{\overline{M}}$ are smooth in the open set $T_\omega \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$, thanks to the implicit function theorem, the smoothness of $d_{\overline{M}}$, and the invertibility of F on this open set.

The map $F(y, v, t) = \exp_y(tv)$ for $y \in M$ and v a unit vector orthogonal to M at y, $t \in (0, \omega)$ is a map from $\widetilde{M} \times (0, \omega)$ into T_ω (since the oriented double cover $\widetilde{M}$ of M is defined as the set of (y, v) with $y \in M$, v unit vector normal to M at y). Arguing as in [21, Prop. 4.8] we see that the following restriction of F (still denoted by F)

$$F : \{((y,v),s) : (y,v) \in \widetilde{M}, s \in (0,\sigma_{(y,v)})\} \to T_\omega \setminus \mathrm{Cut}(M) \setminus \overline{M} \tag{4}$$

is a (smooth) diffeomorphism.[5] This diffeomorphism extends as a continuous map to $\widetilde{M} \times \{0\}$ by sending $((y,v),0)$ to $y \in M$ (note that it is a $2-1$ map here, the standard projection from $\widetilde{M}$ to M). The image of this continuous map is then $T_\omega \setminus \mathrm{Cut}(M) \setminus (\overline{M} \setminus M)$. Again following verbatim [21, prop. 4.8], we also have that the function $\sigma_{(y,v)}$ is continuous on $\widetilde{M}$. The diffeomorpshism F in (4), continuously extended to $\widetilde{M}$, provides the natural replacements for Fermi coordinates around M in our situation, where the singular set $\overline{M} \setminus M$ is present. We will write

$$V_{\widetilde{M}} := \{((y,v),s) : (y,v) \in \widetilde{M}, s \in [0,\sigma_{(y,v)})\},$$

for the domain of (the extension of) F.

Let us take a closer look at the level sets $\Gamma_t = \{x \in T_\omega : d_{\overline{M}}(x) = t\}$ for $t \in (0,\omega)$. By the previous discussion, the smooth hypersurface $\Gamma_t \setminus \overline{S_{d_{\overline{M}}}}$ can be retracted smoothly, staying in $T_\omega \setminus \overline{S_{d_{\overline{M}}}}$ onto a subset of M and at each time the image of the retraction is contained in (a smooth portion of) a level set of $d_{\overline{M}}$. In fact, we have a retraction

$$\left(T_\omega \setminus \mathrm{Cut}(M) \setminus (\overline{M} \setminus M)\right) \times [0,1] \to T_\omega \setminus \mathrm{Cut}(M) \setminus (\overline{M} \setminus M)$$

explicitly given, using the identification (4), by (here $q = (y,v) \in \widetilde{M}$ and $\sigma_q = \sigma_{(y,v)}$)

$$R : \{(q,s) : q \in \widetilde{M}, s \in [0,\sigma_q)\} \times [0,1] \to \{(q,s) : q \in \widetilde{M}, s \in [0,\sigma_q)\},$$
$$R(q,s,\alpha) = (q,(1-\alpha)s).$$

Under the identification (4), the function s is just $d_{\overline{M}}$, so it follows that the retraction preserves level sets of $d_{\overline{M}}$.

We will now analyse the jump part of the Hessian of $d_{\overline{M}} : T_\omega \setminus \overline{M} \to (0,\infty)$; this will lead to Lemma 3.5 below. To this end, we perform, for $\mathcal{H}^n$-a.e. point $x \in \left(\overline{S_{d_{\overline{M}}}} \setminus \overline{M}\right) \cap T_\omega$, a blowup of $d_{\overline{M}}$ as follows. Using normal coordinates around x, for all sufficiently small $\rho > 0$ consider the function $d_\rho : B_1^{n+1}(0) \to (0,\infty)$ defined by

$$d_\rho(y) = \frac{d_{\overline{M}}(x+\rho y) - d_{\overline{M}}(x)}{\rho}.$$

Then $(\nabla d_\rho)(y) = (\nabla d_{\overline{M}})(x+\rho y)$. Note that d_ρ have Lipschitz-constant 1 and $d_\rho(0) = 0$, therefore we can extract a sequence $\rho_j \to 0$ such that d_{ρ_j} converge in $C^{0,\alpha}$ (for all $\alpha < 1$) to a 1-Lipschitz function $d_x : B_1^{n+1}(0) \to \mathbb{R}$ with $d_x(0) = 0$. Recall Proposition 3.2: the rectifiability of $\overline{S_{d_{\overline{M}}}}$ implies that at $\mathcal{H}^n$-a.e. point $x \in \overline{S_{d_{\overline{M}}}} \setminus \overline{M}$ there exists a measure-theoretic unit normal $\hat{n}_x$ to $\overline{S_{d_{\overline{M}}}}$ (rather, two choices of it); moreover, the left and right limits in the Lebesgue sense of the SBV_{loc} function

[5] This diffeomorphism shows, in particular, the following. If $x \in T_\omega \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$, then we know that there exists a unique geodesic γ from x to $\overline{M}$ and its endpoint y is on M by Lemma 3.1. Then, by the properties (a), (b) discussed above, no point of γ is in $\mathrm{Cut}(M)$ and therefore all points on γ except y have the property that they lie in $T_\omega \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$.

$\nabla d_{\overline{M}}$ are well-defined in the two half-spaces identified by the normal (see [3, theorem 3.77]). This means that there exists two constant vectors $a \neq b$ in $\mathbb{R}^{n+1}$ such that

$$\frac{1}{\rho^{n+1}} \int_{\{z \in B_\rho(x) : z \cdot \hat{n}_x < 0\}} |\nabla d_{\overline{M}} - a| \to 0$$

$$\text{and} \quad \frac{1}{\rho^{n+1}} \int_{\{z \in B_\rho(x) : z \cdot \hat{n}_x > 0\}} |\nabla d_{\overline{M}} - b| \to 0$$

as $\rho \to 0$. This is equivalent, by a change of variables, to

$$\int_{\{z \in B_1(0) : z \cdot \hat{n}_x < 0\}} |\nabla d_\rho - a| \to 0, \quad \int_{\{z \in B_1(0) : z \cdot \hat{n}_x > 0\}} |\nabla d_\rho - b| \to 0$$

as $\rho \to 0$. Therefore ∇d_{ρ_j} converge in $L^1(B_1^{n+1}(0))$ to the function F_{ab} defined to be constant on each of the two half-balls $\{z \in B_1^{n+1}(0) : z \cdot \hat{n}_x < 0\}$ and $\{z \in B_1^{n+1}(0) : z \cdot \hat{n}_x > 0\}$, with respective values a and b. This function must be the (distributional) gradient of d_x. Indeed, for every $v \in C_c^1(B_1(0))$ we have

$$\int_{B_1(0)} d_x \nabla v = \lim_{j \to \infty} \int_{B_1(0)} d_{\rho_j} \nabla v$$
$$= - \lim_{j \to \infty} \int_{B_1(0)} \nabla d_{\rho_j} v = - \int_{B_1(0)} F_{ab} v,$$

where we used, in the two limits, respectively, the uniform convergence $d_{\rho_j} \to d_x$ and the L^1-convergence $\nabla d_{\rho_j} \to F_{ab}$. The equality obtained expresses the fact that $F_{ab} = \nabla d_x$ and proves that

$$d_{\rho_j} \to d_x \text{ in } W^{1,1}(B_1(0)) \text{ and in } C^{0,\alpha}(B_1(0)).$$

Recall now that $d_{\overline{M}}$ is locally semiconcave, so it has at least an element in the superdifferential; that is, there exists a C^1 function ϖ in a neighbourhood of x that is $\geq d_{\overline{M}}$ and such that $\varpi(x) = d_{\overline{M}}(x)$. Performing the same blowup on ϖ, we consider the rescalings $\frac{\varpi(x+\rho y) - \varpi(x)}{\rho}$. These functions converge in $C^1(B_1(0))$ to an affine function ϖ_x. By uniform convergence, $\varpi_x \geq d_x$ on $B_1(0)$ and $\varpi_x(0) = d_x(0) = 0$. Recalling that $\nabla d_x = F_{ab}$, we obtain

$$(a - b) \cdot \hat{n}_x \geq 0. \tag{5}$$

The jump part of $D(\nabla d_{\overline{M}})$ is characterized as the measure that is absolutely continuous with respect to $\mathcal{H}^n \llcorner \overline{S_{d_{\overline{M}}}}$ and with density that is given for $\mathcal{H}^n$-a.e. $x \in (\overline{S_{d_{\overline{M}}}} \setminus M) \cap T_\omega$ by $(b - a) \otimes \hat{n}_x$ (see, e.g., [3, (3.90)]). Taking the trace and using (5) this implies the following:

Lemma 3.5. *Let Δ denote the Laplace-Beltrami operator on $T_\omega \setminus \overline{M}$. The singular (jump) part of $\Delta d_{\overline{M}}$ in $T_\omega \setminus \overline{M}$ is a negative measure (supported on $\overline{S_{d_{\overline{M}}}}$).*

Next we analyse the absolutely continuous part (with respect to $\mathcal{H}^{n+1}$) of $\Delta d_{\overline{M}}$ for $d_{\overline{M}} : T_\omega \setminus \overline{M} \to (0, \infty)$. By Proposition 3.2 it suffices to analyse the smooth function $\Delta d_{\overline{M}}$ on $T_\omega \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$. For this, we will need the Ricci curvature assumption (which has not been used so far).

Lemma 3.6. *The function $d_{\overline{M}}$ satisfies $\Delta d_{\overline{M}} \leq 0$ on $T_\omega \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$.*

Proof. Recall Remark 3.4. For $x \in T_\omega \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$, $d_{\overline{M}}(x)$ is realized by the length of a unique geodesic from x to a point in M that we denote by $\pi(x)$, and the level set $\{y \in N \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M}) : d_{\overline{M}}(y) = d_{\overline{M}}(x)\}$ passing through x is C^2, and its scalar mean curvature at x (with respect to the normal that points away from M) is $-\Delta d_{\overline{M}}(x)$. We are thus in the classical situation in which we look at level sets of the distance function to a smooth submanifold, in this case a geodesic ball $B_r(\pi(x))$ in M. This gives the information on the Laplacian in a neighbourhood of x. By Riccati's equation [11, cor. 3.6], using the non-negativity of the Ricci curvature, we get that the mean curvature of the level sets $\{y \in N \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M}) : y = \exp_z(t\nu), z \in B_r(\pi(x))\}$ (this is a disk at distance t from $B_r(\pi(x))$), for either of the choices of unit normal ν on $B_r(\pi(x))$ increases in t, hence $\Delta d_{\overline{M}} \leq 0$ on $N \setminus (\overline{S_{d_{\overline{M}}}} \cup \overline{M})$. □

From Lemmas 3.5 and 3.6, we have $\Delta d_{\overline{M}} \llcorner (T_\omega \setminus \overline{M}) \leq 0$ in the sense of distributions.[6] We now analyse $\Delta d_{\overline{M}}$ at M. For $p \in M$ take a sufficiently small open ball U containing p that is disjoint from $(\overline{M} \setminus M)$ and from $\mathrm{Cut}(M)$ and such that $U \setminus M$ is the union of two disjoint connected open sets U^+ and U^-. We compute the action of the distribution $\Delta d_{\overline{M}}$ on an arbitrary test function $u \in C_c^\infty(U)$ and obtain $(\Delta d_{\overline{M}})(u) = -\int \nabla d_{\overline{M}} \nabla u = -\int_{U^+} \nabla d_{\overline{M}} \nabla u - \int_{U^-} \nabla d_{\overline{M}} \nabla u$. Note that $\nabla d_{\overline{M}}$ extends to a smooth vector field in a neighbourhood of U^+, so $\int_{U^+} \nabla d_{\overline{M}} \nabla u = \int_{U^+} \mathrm{div}(u \nabla d_{\overline{M}}) - \int_{U^+} u \,\mathrm{div}(\nabla d_{\overline{M}})$, where in the last term $\mathrm{div}(\nabla d_{\overline{M}}) = \Delta d_{\overline{M}}$ in the classical sense. The unit outer normal to ∂U^+ agrees, on $\mathrm{supp}(u)$, with $-\nabla d_{\overline{M}}$ (this relevant portion of ∂U^+ is contained in M). The divergence theorem then gives $\int_{U^+} \mathrm{div}(u \nabla d_{\overline{M}}) = -\int_{\partial U^+} u$. Arguing similarly for U^-, we find $(\Delta d_{\overline{M}})(u) = \int_{U^+ \cup U^-} u \Delta d_{\overline{M}} + 2 \int_{M \cap U} u$.

In conclusion, $\Delta d_{\overline{M}} \llcorner (T_\omega \setminus (\overline{M} \setminus M)) = \Delta d_{\overline{M}} \llcorner (T_\omega \setminus \overline{M}) + 2\mathcal{H}^n \llcorner M$. In particular, $\Delta d_{\overline{M}} \llcorner (T_\omega \setminus (\overline{M} \setminus M))$ is a Radon measure (we have given its Hahn decomposition into negative and positive parts). We will now extend across $\overline{M} \setminus M$ by a capacity argument.

Proposition 3.7. *Let N be a closed $(n+1)$-dimensional Riemannian manifold with positive Ricci curvature and M a smooth minimal hypersurface as in Theorem 1.3. Denote by $d_{\overline{M}}$ the distance function to $\overline{M}$ and by $T_\omega = \{x \in N : d_{\overline{M}}(x) < \omega\}$, where $\omega < \mathrm{inj}(N)$. Then $\Delta d_{\overline{M}}$ is a Radon measure on T_ω, with positive part $2\mathcal{H}^n \llcorner M$.*

Proof. Let $\delta > 0$ be arbitrary and choose $\chi \in C_c^\infty(T_\omega)$ to be a function that takes values in [0,1], is identically 1 in an open neighbourhood of $\overline{M} \setminus M$, identically 0 away from a (larger) neighbourhood of $\overline{M} \setminus M$, and such that $\int_{T_\omega} |\nabla \chi| < \delta$ (see [8, 4.7]). Then we have, for $v \in C_c^\infty(T_\omega)$,

$$
\begin{aligned}
&(\Delta d_{\overline{M}} - 2\mathcal{H}^n \llcorner M)(v) \\
&\quad = (\Delta d_{\overline{M}} - 2\mathcal{H}^n \llcorner M)((1-\chi)v) + (\Delta d_{\overline{M}})(\chi v) - 2\int_M \chi v \\
&\quad = (\Delta d_{\overline{M}} - 2\mathcal{H}^n \llcorner M)((1-\chi)v) - \int_{T_\omega} \nabla d_{\overline{M}} \nabla \chi \, v \\
&\qquad - \int_{T_\omega} \nabla d_{\overline{M}} \nabla v \, \chi - 2\int_M \chi v.
\end{aligned}
\tag{6}
$$

[6] A distribution is said to be ≤ 0 if for every nonnegative test function the result is ≤ 0. A distribution that is ≥ 0 or ≤ 0 is necessarily a Radon measure; see, for example [8, theorem 1.39].

For the second term recall that the distribution $\nabla d_{\overline{M}}$ is an L^∞ function with $|\nabla d_{\overline{M}}| = 1$ a.e. and so

$$\left|\int_{T_\omega} \nabla d_{\overline{M}} \nabla \chi v\right| \le \|v\|_{L^\infty}\left(\int_{T_\omega} |\nabla \chi|\right) < \delta \|v\|_{L^\infty}.$$

This tends to 0 as $\delta \to 0$. As $\delta \to 0$, the corresponding χ will go to 0 in $L^1(T_\omega)$ so the third term will also tend to 0. For the fourth term, we notice that (by the construction of χ) $\mathrm{supp}(\chi)$ is contained in $\{\mathrm{dist}_N(\cdot, \overline{M} \setminus M) < d\}$ with $d \to 0$ for $\delta \to 0$; as $\mathcal{H}^n \llcorner M$ is a finite measure, we have that

$$(\mathcal{H}^n \llcorner M)\Big(\{\mathrm{dist}_N(\,\cdot\,, \overline{M} \setminus M) < d\}\Big) \to 0;$$

hence the fourth term also tends to 0 for $\delta \to 0$.

The distribution $\Delta d_{\overline{M}}$ is a priori of order ≤ 1:

$$\left|\int_{T_\omega} (\Delta d_{\overline{M}}) v\right| = \left|\int_{T_\omega} \nabla d_{\overline{M}} \nabla v\right| \le \mathcal{H}^{n+1}(N)\|v\|_{C^1}.$$

For the first term in the rightmost side of (6), observe that

$$(1-\chi)v \in C^\infty_c(T_\omega \setminus (\overline{M} - M))$$

and $\Delta d_{\overline{M}} - 2\mathcal{H}^n \llcorner M$ is a negative Radon measure on this open set (by Lemma 3.6 and by the observation preceding Proposition 3.7), so that

$$(\Delta d_{\overline{M}} - 2\mathcal{H}^n \llcorner M)((1-\chi)v) \le 0$$

if $v \ge 0$ (because $(1-\chi)v \ge 0$ by the choice of χ). As (6) holds for all δ, and its last three terms tend to 0 as $\delta \to 0$, for every $v \in C^\infty_c(T_\omega)$ and $v \ge 0$ we have

$$(\Delta d_{\overline{M}} - 2\mathcal{H}^n \llcorner M)(v) = \lim_{\delta\to 0}(\Delta d_{\overline{M}} - 2\mathcal{H}^n \llcorner M)((1-\chi)v) \le 0.$$

The distribution $\Delta d_{\overline{M}} - 2\mathcal{H}^n \llcorner M$ is therefore a negative Radon measure on T_ω. □

4 | LEVEL SETS OF $d_{\overline{M}}$

We consider the level sets $\Gamma_t = \{x : d_{\overline{M}}(x) = t\}$, for $t \in [0, \omega/2]$ (we fixed an arbitrary $\omega \in (0, \mathrm{inj}(N))$); we will obtain that the areas of Γ_t are "essentially" decreasing in t. Further, we will consider an "Allen–Cahn approximation" $G^\varepsilon_0 : N \to \mathbb{R}$ of $\Gamma_{6\varepsilon|\log\varepsilon|} = \Gamma_{2\varepsilon\Lambda}$ defined, for ε sufficiently small (to ensure $4\varepsilon\Lambda < \omega/2$), as follows:

$$G^\varepsilon_0(x) = \begin{cases} -1 & \text{for } x \in N \setminus T_\omega \\ \overline{\mathbb{H}}^\varepsilon(-d_{\overline{M}}(x) + 2\varepsilon\Lambda) & \text{for } x \in T_\omega \end{cases}. \tag{7}$$

Since $\overline{\mathbb{H}}^\varepsilon$ is constantly -1 on $(-\infty, -2\varepsilon\Lambda]$, the function G^ε_0 is constantly -1 on $\{x : d_{\overline{M}}(x) > 4\varepsilon\Lambda\}$. Since $\overline{\mathbb{H}}^\varepsilon$ is smooth, G^ε_0 has the same regularity of $d_{\overline{M}}$, that is, it is locally Lipschitz, $G^\varepsilon_0 \in W^{1,\infty}(N)$. Moreover, its gradient (which equals $-(\overline{\mathbb{H}}^\varepsilon)'(-d_{\overline{M}}(x) + 2\varepsilon\Lambda)\nabla d_{\overline{M}}(x)$ in T_ω and 0 otherwise) is in

$BV(N)$ and its distributional Laplacian ΔG_0^ε is a Radon measure (as computed within (4) below). Note that the profile of G_0^ε in the normal direction at any point of M is given by the function $\Psi = \Psi_0$ in (3), therefore G_0^ε can also be thought of as an Allen–Cahn approximation of $2|M|$, or equivalently of the immersion $\iota : \widetilde{M} \to N$ that covers M twice. The fact that $\mathcal{E}_\varepsilon(G_0^\varepsilon)$ is approximately $2|M|$ will be established later.

The Allen–Cahn first variation of G_0^ε (which is clearly 0 outside T_ω) can be computed in T_ω as follows:

$$\begin{aligned}
-(2\sigma)\mathcal{E}'_\varepsilon(G_0^\varepsilon) &= \varepsilon\,\Delta G_0^\varepsilon - \frac{W'(G_0^\varepsilon)}{\varepsilon} \\
&= \varepsilon\,\overline{\mathbb{H}}^{\varepsilon\prime\prime}(-d_{\overline{M}} + 2\,\varepsilon\,\Lambda)|\nabla d_{\overline{M}}|^2 - \varepsilon\,\overline{\mathbb{H}}^{\varepsilon\prime}(-d_{\overline{M}} + 2\,\varepsilon\,\Lambda)\Delta d_{\overline{M}} \\
&\quad - \frac{W'(\overline{\mathbb{H}}^{\varepsilon}(-d_{\overline{M}} + 2\,\varepsilon\,\Lambda))}{\varepsilon} \\
&= \underbrace{\varepsilon\,\overline{\mathbb{H}}^{\varepsilon\prime\prime}(-d_{\overline{M}} + 2\,\varepsilon\,\Lambda) - \frac{W'(\overline{\mathbb{H}}^{\varepsilon}(-d_{\overline{M}} + 2\,\varepsilon\,\Lambda))}{\varepsilon}}_{O(\varepsilon^2)} \\
&\quad - \underbrace{\varepsilon\,\overline{\mathbb{H}}^{\varepsilon\prime}(-d_{\overline{M}} + 2\,\varepsilon\,\Lambda)}_{0\leq\cdot\leq 3}\,\underbrace{\Delta d_{\overline{M}}}_{\leq 0},
\end{aligned}$$

in the distributional sense. Since $\Delta d_{\overline{M}}$ a Radon measure thanks to Proposition 3.7, we will think of $-\mathcal{E}'_\varepsilon(G_0^\varepsilon)$ as a Radon measure. The term $O(\varepsilon^2)$ (first brace) is a Lipschitz function that we interpret as a density with respect to $\mathcal{H}^{n+1}$. The last term is the measure $\Delta d_{\overline{M}}$ multiplied by a bounded Lipschitz function; recall that $\overline{\mathbb{H}}^{\varepsilon\prime}(2\,\varepsilon\,\lambda) = 0$. With abuse of notation, we neglect the positive part of $\Delta d_{\overline{M}}$ in the third brace, as it is supported on $\{d_{\overline{M}} = 0\}$.

Denote by $\mathcal{F}_{\varepsilon,\mu}$, for a constant $\mu > 0$, the functional on $W^{1,2}(N)$ given by

$$\mathcal{F}_{\varepsilon,\mu}(u) = \mathcal{E}_\varepsilon(u) - \frac{\mu}{2\sigma}\int_N u.$$

The computation in (4) shows that for every ε there exists $\mu_\varepsilon > 0$, $\mu_\varepsilon \to 0$ as $\varepsilon \to 0$, such that[7] (we need $\mu_\varepsilon > 4\sigma\|O(\varepsilon^2)\|_{L^\infty}$ where $O(\varepsilon^2)$ is the first term in the last line of (4))

$$-\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(G_0^\varepsilon) = -\mathcal{E}'_\varepsilon(G_0^\varepsilon) + \frac{\mu_\varepsilon\mathcal{H}^{n+1}}{2\sigma} \geq \frac{1}{2}\frac{\mu_\varepsilon}{2\sigma}\mathcal{H}^{n+1}.$$

(The inequality means that the Radon measure on the left minus the Radon measure on the right is a nonnegative measure.) The function G_0^ε will form the starting point for the construction of a barrier for the negative $\mathcal{F}_{\varepsilon,\mu_\varepsilon}$-gradient flow in Section 7.5.

Areas of Γ_t. Since $\overline{S_{d_{\overline{M}}}}$ is countably n-rectifiable (and thus has Hausdorff dimension $\leq n$ and vanishing $\mathcal{H}^{n+1}$ measure) we get that, for a.e. $t > 0$, $\mathcal{H}^n(\overline{S_{d_{\overline{M}}}} \cap \Gamma_t) = 0$. We will denote by $\Omega \subset (0,\omega)$ the set with $\mathcal{H}^1(\Omega) = 0$ such that

$$t \in (0,\omega) \setminus \Omega \Rightarrow \mathcal{H}^n(\overline{S_{d_{\overline{M}}}} \cap \Gamma_t) = 0$$

[7] A precise choice of μ_ε will be made in (35).

(and therefore, for $t \notin \Omega$, Γ_t is a smooth hypersurface away from a $\mathscr{H}^n$-negligible set). Therefore for $t \in (0,\omega) \setminus \Omega$ we have $\mathcal{H}^n(\Gamma_t) = \mathcal{H}^n(\Gamma_t \setminus \overline{S_{d_{\overline{M}}}})$; that is, we only need to compute the area of the smooth part of Γ_t. Thanks to this, we will compare the area of Γ_t to that of M for $t \in (0,\omega) \setminus \Omega$.

Lemma 4.1. *Let* $\Gamma_t = \{x \in N : d_{\overline{M}}(x) = t\}$ *and* $\Omega \subset (0,\omega)$ *as above* ($\mathcal{H}^1(\Omega) = 0$)*. Then*

1. *for* $t \in (0,\omega) \setminus \Omega$ *the set* Γ_t *is a smooth hypersurface away from a set of vanishing* $\mathcal{H}^n$*-measure and* $\mathcal{H}^n(\Gamma_t) < 2\mathcal{H}^n(M)$*;*
2. *the function* $t \in (0,\omega) \to \mathcal{H}^n(\Gamma_t)$ *satisfies for* $t_1 < t_2$, $t_2 \notin \Omega$ ($t_1 \in \Omega$ *is allowed*)*, the inequality* $\mathcal{H}^n(\Gamma_{t_2}) < \mathcal{H}^n(\Gamma_{t_1})$.

Proof. The first part of (1) has already been discussed above. Recall the diffeomorphism induced by F in Remark 3.4. Endow $\{(q,s) : q \in \widetilde{M}, s \in [0,\sigma_q)\}$ with the pull-back metric (via F) from $T_\omega \setminus \mathrm{Cut}(M) \setminus \overline{M}$. The metric extends continuously to $\widetilde{M} \times \{0\}$ to give the natural metric on $\widetilde{M}$. We will thus work in $V_{\widetilde{M}} = \{(q,s) : q \in \widetilde{M}, s \in [0,\sigma_q)\}$; note that $F^{-1}(\Gamma_{t_0} \setminus \overline{S_{d_{\overline{M}}}}) = \{(x,s) \in V_{\widetilde{M}} : s = t_0\}$. Denoting by Π the map $\Pi(q,s) = (q,0)$, recall that from the structure of $V_{\widetilde{M}}$ we obtain the following. For every $t < t_0$ the set $\{(x,s) \in V_{\widetilde{M}} : x \in \Pi(F^{-1}(\Gamma_{t_0} \setminus \overline{S_{d_{\overline{M}}}})), s = t\}$ is contained in $F^{-1}(\Gamma_t \setminus \overline{S_{d_{\overline{M}}}})$. It is then enough, for (1) and (2), to prove that, if $t_0 \notin \Omega$ and $t < t_0$, then $\{(x,s) \in V_{\widetilde{M}} : s = t_0\}$ has area bounded by $\{(x,s) \in V_{\widetilde{M}} : x \in \Pi(\{(x,s) \in V_{\widetilde{M}} : s = t_0\}), s = t\}$.

Let $(x_1, \dots, x_n, s)$ be local coordinates on $V_{\widetilde{M}}$ chosen so that $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ form a local frame around a point $x_0 \in \widetilde{M}$, that is orthonormal at $x_0 \in \widetilde{M}$, and $\frac{\partial}{\partial s}$ is the unit speed of the geodesics $\{x = \text{const}\}$. Then the Riemannian metric on $V_{\widetilde{M}}$ induces an area element θ_{s_0} for the level set $\{s = s_0\}$ at the point (x_0, s_0). By [11, theorem 3.11] it satisfies the ODE $\frac{\partial}{\partial s} \log \theta_s = -\vec{H}(x_0,s) \cdot \frac{\partial}{\partial s}$, where $\vec{H}_{(x_0,s)}$ is the mean curvature of the level set at distance s evaluated at the point (x_0,s). (Note that in [11] θ_s denotes the volume element, but since $\frac{\partial}{\partial s}$ is a unit vector, the area and volume elements are the same.) By Riccati's equation [11, cor. 3.6] we find that $H(x_0,s) = \vec{H}_{(x_0,s)} \cdot \frac{\partial}{\partial s}$ is strictly increasing in s, at least at linear rate, thanks to the positiveness of the Ricci curvature, $H_{(x_0,s)} \geq s(\min_N \mathrm{Ric}_N)$. Therefore $\frac{\partial}{\partial s} \log \theta_s \leq -s(\min_N \mathrm{Ric}_N)$ and we find for $s_0 \geq 0$, $t \geq 0$

$$\log\left(\frac{\theta(s_0+t)}{\theta(s_0)}\right) \leq -(\min_N \mathrm{Ric}_N) \int_{s_0}^{s_0+t} s\, ds$$

and therefore

$$\theta_{s_0+t} \leq \theta_{s_0} e^{-\frac{\min_N \mathrm{Ric}_N}{2}(2s_0 t + t^2)} \quad \text{for } (x_0, s_0+t) \in V_{\widetilde{M}}.$$

In particular, $\theta(t)$ is decreasing in t. From this (1) and (2) follow by integrating the area element. (Recall that $\int_{\widetilde{M}} \theta_0 dx^1 \dots dx^n = 2\mathcal{H}^n(M)$.) □

Allen–Cahn energy of G_0^ε. Thanks to Lemma 4.1 we can control the Allen–Cahn energy of G_0^ε by twice the area of M. Indeed, recalling that the energy is 0 in the complement of $T_{\omega/2}$ and that ∇G_0^ε is parallel to $\nabla d_{\overline{M}}$, we use the coarea formula for the slicing function $d_{\overline{M}}$ (for which $|\nabla d_{\overline{M}}| = 1$) and we get

$$\int_{T_{\omega/2}} \varepsilon \frac{|\nabla G_0^\varepsilon|^2}{2} + \frac{W(G_0^\varepsilon)}{\varepsilon} = \int_0^\omega \left(\int_{\Gamma_s} \frac{|\nabla G_0^\varepsilon|^2}{2} + \frac{W(G_0^\varepsilon)}{\varepsilon} \right) ds \underbrace{=}_{\text{by}(7)}$$

$$= \int_{-\omega/2+2\varepsilon\Lambda}^{2\varepsilon\Lambda} \left(\int_{\Gamma_{2\varepsilon\Lambda-s}} \varepsilon \frac{(\overline{\mathbb{H}}^{\varepsilon\prime}(s))^2}{2} + \frac{W(\overline{\mathbb{H}}^\varepsilon(s))}{\varepsilon} \right) ds \underbrace{\leq}_{\text{Lemma 4.1}}$$

$$\leq 2\mathcal{H}^n(M) \left(\int_{\mathbb{R}} \varepsilon \frac{(\overline{\mathbb{H}}^{\varepsilon\prime})^2}{2} + \frac{W(\overline{\mathbb{H}}^\varepsilon)}{\varepsilon} \right),$$

where we used Lemma 4.1(a) for a.e. s, namely $s \notin \Omega$. By the estimates in (1) we get

$$\mathcal{E}_\varepsilon(G_0^\varepsilon) \leq 2\mathcal{H}^n(M)\,(1 + O(\varepsilon^2)). \tag{8}$$

5 | INSTABILITY PROPERTIES OF *M* (CHOICE OF *B*)

Let $\iota : \widetilde{M} \to N$ be the (smooth) minimal immersion induced by the standard projection (2–1 map) from the oriented double cover of M onto M. Let ν be a choice (on $\widetilde{M}$) of unit normal to the immersion ι. Recall (Remark 3.4) the coordinates $((y, v), s) = (q, s)$ on $V_{\widetilde{M}}$, which is diffeomorphic to $T_\omega \setminus \overline{S_{d_{\overline{M}}}} \setminus (\overline{M} \setminus M)$; here $y \in M$ and v a unit vector orthogonal to M at y, or, equivalently, $q = (y, v) \in \widetilde{M}$. For every compact set $K \subset \widetilde{M}$ there exists $c_K > 0$ such that $c_K < \sigma_{(y,v)}$ for all $(y, v) \in K$. This follows from the continuity of σ_q on $\widetilde{M}$ (Remark 3.4). Choosing K even (i.e., such that K is the double cover of a compact set $\iota(K)$ in M) this means that $\iota(K)$ admits a two-sided tubular neighbourhood of semiwidth c_K.

We will now consider deformations of ι with initial velocity dictated by a function $\varphi \in C_c^2(\widetilde{M})$. For $\varphi \in C_c^2(\widetilde{M})$, choose $c_{\text{supp}\,\varphi}$ as above and consider the following one-parameter family of immersions $\iota_t : \widetilde{M} \to N$ defined for $t \in (-\delta_0, \delta_0)$, where $\delta_0 \in \left(0, \frac{c_{\text{supp}\,\varphi}}{\max \varphi}\right)$:

$$(y, v) \to \exp_{\iota(y)}(t\varphi((y, v))\nu((y, v))),$$

for $(y, v) \in \widetilde{M}$. The first variation of area at $t = 0$ is 0 because M is minimal. The second variation of area at $t = 0$ is given by

$$\int_{\widetilde{M}} |\nabla\varphi|^2 d\mathcal{H}^n - \int_{\widetilde{M}} \varphi^2(|A|^2 + \text{Ric}_N(\nu, \nu)) d\mathcal{H}^n, \tag{9}$$

where A denotes the second fundamental form of ι, ∇ the gradient on $\widetilde{M}$ (with respect to g_0, the Riemannian metric induced by the pull-back from M), Ric_N the Ricci tensor of N and $\mathcal{H}^n$ is induced on $\widetilde{M}$ by g_0 (equivalently, integrate with respect to $d\,\text{vol}_{g_0}$).

Lemma 5.1 (Unstable region). *There exist a geodesic ball $D \Subset M$ and $\widetilde{\phi} \in C_c^2(\widetilde{M})$ with $\widetilde{\phi} \geq 0$, such that, writing $\widetilde{D} = \iota^{-1}(D)$, the support of $\widetilde{\phi}$ is contained in $\widetilde{M} \setminus \overline{\widetilde{D}}$ and*

$$\int_{\widetilde{M}} |\nabla\widetilde{\phi}|^2\, d\mathcal{H}^n - \int_{\widetilde{M}} \widetilde{\phi}^2(|A|^2 + \text{Ric}_N(\nu, \nu)) d\mathcal{H}^n < 0. \tag{10}$$

Proof. The second variation of M is only defined for initial velocities induced by a function with compact support in M. Fix an arbitrary point $b \in M$. Let $\delta > 0$ be arbitrary and choose $\rho = \rho_\delta \in C_c^\infty(N)$ such that $0 \le \rho \le 1$, $\rho = 1$ in an open neighbourhood of $\{b\} \cup (\overline{M} \setminus M)$, $\rho = 0$ in the complement of a (larger) open neighbourhood of $\{b\} \cup (\overline{M} \setminus M)$, and $\int_N |\nabla \rho|^2 < \delta$. This is possible because $\{b\} \cup (\overline{M} \setminus M)$ has finite (actually 0 when $n > 2$) $\mathcal{H}^{n-2}$-measure, and the mass growth is Euclidean around every point of $\overline{M}$ (since $2|M|$ is a stationary integral varifold, which gives the validity of the monotonicity formula): the previous two facts allow us to conclude that the two-capacity of $\{b\} \cup (\overline{M} \setminus M)$ is 0 (see [8, sec. 4.7]), establishing the existence of ρ with the desired properties.

Then the function $\varphi(q) = 1 - \rho(\iota(q))$ is admissible in (9) and the expression becomes (integrating on M)

$$2 \int_M |\nabla \rho|^2 \, d\mathcal{H}^n - 2 \int_M (1-\rho)^2 (|A_M|^2 + \mathrm{Ric}_N(\nu, \nu)) d\mathcal{H}^n.$$

(Note that on M the choice of ν is in general only permitted up to sign; this suffices for the term $\mathrm{Ric}_N(\nu, \nu)$ to make sense.) Sending $\delta \to 0$ the second term tends to $-2 \int_M (|A_M|^2 + \mathrm{Ric}_N(\nu, \nu)) d\mathcal{H}^n$ and the first term tends to 0, so the above expression converges to a negative number (recall that $\mathrm{Ric}_N > 0$). Therefore there exists δ sufficiently small such that

$$2 \int_M |\nabla \rho|^2 \, d\mathcal{H}^n - 2 \int_M (1-\rho)^2 (|A_M|^2 + \mathrm{Ric}_N(\nu, \nu)) d\mathcal{H}^n < 0.$$

We let, for this δ, $\tilde{\phi}(q) = 1 - \rho(\iota(q))$. Since $1 - \rho$ vanishes in a neighbourhood of b, there exists a geodesic ball D whose closure is disjoint from $\mathrm{supp}(1-\rho)$, and therefore its double cover $\tilde{D}$ is a positive distance away from $\mathrm{supp}\, \tilde{\phi}$. □

Remark 5.2. This lemma uses $n \ge 2$ to argue that $\{b\}$ has codimension ≥ 2 (for $n = 1$ the lemma fails, e.g., for $\mathbb{RP}^1 \subset \mathbb{RP}^2$).

Remark 5.3. By the construction of ρ in [8], $\rho(x) = 0$ when $\mathrm{dist}_N(x, \{b\} \cup (\overline{M} \setminus M)) > d_\delta$ for some $d_\delta \to 0$ as $\delta \to 0$. This means that for δ sufficiently small the support of ρ has at least two (compact) connected components one of which contains b (and thus $\overline{\tilde{D}}$) while the union of the others contains an open neighbhourhood O_1 of $\overline{M} \setminus M$. Let $O \Subset O_1$ be an open set containing $\overline{M} \setminus M$ (to avoid technical difficulties, we ensure also that $\partial O \cap \overline{M}$ is $(n-1)$-dimensional, thanks to the coarea formula for dist_N). For $\tilde{\phi} = 1 - \rho \circ \iota$, we have that the complement of $\mathrm{supp}\, \tilde{\phi}$ has at least two (open) connected components in $\tilde{M}$, one containing $\overline{\tilde{D}}$ while the other contains $\overline{\iota^{-1}(O)}$. Note that $K = \tilde{M} \setminus \iota^{-1}(O)$ is compact. These facts guarantee that $\tilde{\phi}$ vanishes in a neighbourhood of $\partial \tilde{D}$ and of $\partial(\iota^{-1}(O)) = \partial K$, a condition that will be technically useful in Section 6.

Remark 5.4 (Choice of B). Choose the ball B in M to be concentric with D and with half the radius. Denote by $R > 0$ the radius of B. Let $\tilde{B} = \iota^{-1}(B)$: this is the union of two geodesic balls in $\tilde{M}$. The choices of B and $\tilde{\phi}$ will be kept until the end.

Remark 5.5. The geometric counterpart of Lemma 5.1 is that the minimal immersion ι is unstable with respect to the area functional also if we restrict to deformations that leave $\tilde{D}$ (and D) fixed and that do not move M close to its singular set $\overline{M} \setminus M$. We will be more specific in Section 6 below.

6 | RELEVANT IMMERSIONS (CHOICE OF τ)

Recall Remark 5.3. We will fix the compact subset $K = \widetilde{M} \setminus \iota^{-1}(O)$ and will denote by K_B the compact set $K \setminus \widetilde{B}$, where $\widetilde{B}$ is as in Remark 5.4. Note that both K and K_B are even in $\widetilde{M}$; that is, they are double covers (via ι) of compact subsets of M. We have $\operatorname{supp} \widetilde{\phi} \subset K_B \subset K$ for $\widetilde{\phi}$ chosen in Lemma 5.1. Recall that $\widetilde{\phi}$ vanishes in a neighbourhood of ∂K_B (and of ∂K). We will define on K and K_B suitable two-sided immersions into N, smooth up to the boundaries ∂K and ∂K_B (this means that there exist open neighbourhoods of K and K_B to which the immersions can be smoothly extended).

Choose $c_K > 0$ such that $c_K < \min_{(y,v)\in K} \sigma_{(y,v)}$ (by the continuity of $\sigma > 0$ on $\widetilde{M}$ the minimum exists and is positive). We therefore have a well-defined one-sided tubular neighbourhood of K in $V_{\widetilde{M}}$, namely $K \times [0, c_K)$, with closure contained in $V_{\widetilde{M}}$. Note that there exists an open neighbourhood of K on which $\sigma_{(y,v)} > c_K$, by continuity of σ on $\widetilde{M}$.

Recall that $V_{\widetilde{M}}$ is endowed with the Riemannian metric induced by the pull-back from N. Let Π_K denote the nearest-point projection onto K (in coordinates, $\Pi_K(q,s) = (q,0)$). For future purposes, we ensure that c_K above is also suitably small to ensure that, for $x = (q,s) \in K \times [0, c_K)$, then

$$\big||J\Pi_K|(x) - 1\big| \le 2C_K s \quad \text{and} \quad \left|\frac{1}{|J\Pi_K|(x)} - 1\right| \le 2C_K s, \tag{11}$$

where $|J\Pi_K| = \sqrt{(D\Pi_K)(D\Pi_K)^T}$, and the constant $C_K > 0$ is the maximum of the norm of the second fundamental form of $\iota : \widetilde{M} \to N$ restricted to $K \subset \widetilde{M}$. Note that s is just the Riemannian distance of (q,s) to K (and to $\overline{M}$).

Choosing $\tilde{c}_0 > 0$ and $\tilde{t}_0 > 0$ sufficiently small, we can ensure that

$$\big(q, c + t\widetilde{\phi}(q)\big) \in K \times \left[0, \frac{c_K}{2}\right)$$

for all $t \in [0, \tilde{t}_0]$ and for all $c \in [0, \tilde{c}_0]$. For any such c, t we thus have a smooth two-sided immersion $q = (y,v) \in \operatorname{Int}(K) \to \exp_y\big((c + t\widetilde{\phi}(q))v\big)$ from the interior of K into N.

Remark 6.1. Note that, since $\widetilde{\phi} = 0$ in a neighbourhood of ∂K, the immersion

$$q = (y,v) \in \operatorname{Int}(K) \to \exp_y\big((c + t\widetilde{\phi}(q))v\big)$$

agrees with $q = (y,v) \in \operatorname{Int}(K) \to \exp_y(cv)$ in a neighbourhood of ∂K; therefore it extends smoothly to ∂K. Similarly,

$$q = (y,v) \in \operatorname{Int}(K_B) \to \exp_y\big((c + t\widetilde{\phi}(q))v\big)$$

extend smoothly to ∂K_B because $\widetilde{\phi} = 0$ vanishes in a neighbourhood of ∂K.

Remark 6.2.

(a) Again thanks to the fact that $\widetilde{\phi} = 0$ in a neighbourhood of ∂K, we have the following technically useful fact. For the two-sided immersion

$$q = (y,v) \in K \to \exp_y\big((c + t\widetilde{\phi}(q))v\big),$$

with $c > 0$, denote by ν a choice of unit normal (which extends continuously up to ∂K) and by $K_{c,t,\widetilde{\phi}}$ its image. We can find $\underline{c} > 0$ such that, for any $t \in [0, \tilde{t}_0]$ and $c \in [0, \tilde{c}_0]$, the set $\{\exp_x(s\nu) : s \in (-\underline{c}, \underline{c}), x \in K_{c,t,\widetilde{\phi}}\}$ is contained in $K \times [0, c_K)$. By making $\underline{c}$ smaller if necessary, we can also ensure that the set

$$\left\{\exp_x(s\nu) : s \in (-\min\{c, \underline{c}\}, \min\{c, \underline{c}\}), x \in K_{c,t,\widetilde{\phi}}\right\}$$

is a tubular neighbourhood of $K_{c,t,\widetilde{\phi}}$, in the sense that it admits a well-defined nearest-point projection $\Pi_{c,t}$ onto $K_{c,t,\widetilde{\phi}}$. This projection extends smoothly up to the boundary portion $\{\exp_x(s\nu) : s \in (-\min\{c, \underline{c}\}, \min\{c, \underline{c}\}), x \in \partial K_{c,t,\widetilde{\phi}}\}$. In fact, close to $\{\exp_x(s\nu) : s \in (-\min\{c, \underline{c}\}, \min\{c, \underline{c}\}), x \in \partial K_{c,t,\widetilde{\phi}}\}$ we have that $\Pi_{c,t}$ agrees with the nearest-point projection onto Γ_c.

These properties essentially say that we can work with tubular neighbourhoods of $K_{c,t,\widetilde{\phi}}$ without interfering with the complement of $F(K \times [0, c_K))$, and it will be useful when writing Allen–Cahn approximations of the immersions $q = (y, v) \in \mathrm{Int}(K) \to \exp_y\left((c + t\widetilde{\phi}(q))v\right)$.[8]

(b) For notational convenience we redefine $\widetilde{c_0}$, by choosing the minimum of $\widetilde{c_0}$ specified above and $\underline{c}$ specified in (a). Then we have a well-defined nearest point projection

$$\Pi_{c,t} : \{\exp_x(s\nu) : s \in (-c, c), x \in K_{c,t,\widetilde{\phi}}\} \to K_{c,t,\widetilde{\phi}}$$

for all $c \in (0, \tilde{c_0}]$ and all $t \in [0, \tilde{t_0}]$.

Remark 6.3. Choosing a suitably small $t_0 \leq \tilde{t}_0$, $t_0 > 0$, we can further ensure that the area of the immersion $q = (y, v) \in \mathrm{Int}(K) \to \exp_y\left((t\widetilde{\phi}(q))v\right)$ is strictly decreasing in t on the interval $[0, t_0]$. This follows upon noticing that the first variation (with respect to area) at $t = 0$ is 0 (by minimality of M), and the second variation at $t = 0$ is negative by Lemma 5.1 (see Remark 5.5). Note that the immersions $q = (y, v) \in \mathrm{Int}(K_B) \to \exp_y\left((t\widetilde{\phi}(q))v\right)$ (the previous family of immersions restricted to $\mathrm{Int}(K_B)$) have the same area-decreasing property, since $\widetilde{\phi} = 0$ on $\widetilde{D}$. For the latter family of immersions, the area at $t = 0$ is

$$\mathcal{H}^n(K) - \mathcal{H}^n(\widetilde{B}) \leq 2\mathcal{H}^n(M) - 2\mathcal{H}^n(B).$$

Lemma 6.4. *Let t_0 be as in Remark 6.3 and $\widetilde{c}_0$ as in Remark 6.2(b). There exist $c_0 \in (0, \widetilde{c}_0]$ and $\tau > 0$ such that*

(i) for all $c \in [0, c_0]$ and for all $t \in [0, t_0]$ the area of the immersion

$$q = (y, v) \in \mathrm{Int}(K_B) \to \exp_y\left((c + t\widetilde{\phi}(q))v\right)$$

is $\leq \mathcal{H}^n(K) - \frac{3}{4}\mathcal{H}^n(\widetilde{B}) = \mathcal{H}^n(K) - \frac{3}{2}\mathcal{H}^n(B)$;

(ii) for all $c \in [0, c_0]$ the area of the immersion

$$q = (y, v) \in \mathrm{Int}(K) \to \exp_y\left((c + t_0\widetilde{\phi}(q))v\right)$$

is $\leq \mathcal{H}^n(K) - \tau$.

[8] More precisely, we can patch the definition of Allen–Cahn approximation given in the tubular neighbourhood of $K_{c,t,\widetilde{\phi}}$ (for $c = 2\varepsilon\Lambda$ to be chosen) with the function G_0^ε defined in (7).

Proof. Let us prove that (i) holds for some $c_0' \in (0, \widetilde{c}_0]$ (in place of c_0). Argue by contradiction: if not, then there exists $c_i \to 0$ and $t_i \in [0, t_0]$ such that the area of $q \in \mathrm{Int}(K_B) \to \exp_y\big((c_i + t_i\widetilde{\phi}(q))v\big)$ is $\geq 2(\mathcal{H}^n(M) - \frac{3}{4}\mathcal{H}^n(B))$ for all i. Upon extracting a subsequence we may assume $t_i \to t \in [0, t_0]$, and by continuity of the area we get that the area of $q \in \mathrm{Int}(K_B) \to \exp_y\big((t\widetilde{\phi}(q))v\big)$ is $\geq (\mathcal{H}^n(K) - \frac{3}{2}\mathcal{H}^n(B))$. This is however in contradiction with Remark 6.3, which says that this area is $\leq \mathcal{H}^n(K) - 2\mathcal{H}^n(B)$.

Let us prove that (ii) holds for some $c_0'' \in (0, \widetilde{c}_0]$ (in place of c_0) and for some $\tau > 0$. By Remark 6.3 the area of $q = (y, v) \in \mathrm{Int}(K) \to \exp_y\big((t_0\widetilde{\phi}(q))v\big)$ is strictly smaller than $\mathcal{H}^n(K)$. Denote by 2τ the positive difference of the two areas. By continuity, there exists $c_0'' > 0$ such that for all $c \in [0, c_0'']$ the area of the immersion $q = (y, v) \in \mathrm{Int}(K) \to \exp_y\big((c + t_0\widetilde{\phi}(q))v\big)$ is smaller than $\mathcal{H}^n(K) - \tau$.

Choosing $c_0 = \min\{c_0', c_0''\}$ concludes. □

We will write, in Section 7, Allen–Cahn approximations of the immersions in Lemma 6.4. To that end, we will work in the tubular neighbourhoods specified in Remark 6.2, restricting the range of c and t to $[0, c_0]$ and $[0, t_0]$, respectively (in order to exploit the area bounds obtained in the lemma). We will also make use of the following bounds.

Remark 6.5. There exists a constant $C_{K,c_0,t_0} > 0$, depending only on c_0, t_0, on the Riemannian metric and on the C^3 norms of $\widetilde{\phi}$ on K and of F, such that for $c \in [0, c_0]$ and $t \in [0, t_0]$

$$\big||J\Pi_{c,t}|(x) - 1\big| \leq C_{K,c_0,t_0} s \quad \text{and} \quad \left|\frac{1}{|J\Pi_{c,t}|(x)} - 1\right| \leq C_{K,c_0,t_0} s, \tag{12}$$

where $|J\Pi_{c,t}| = \sqrt{(D\Pi_{c,t})(D\Pi_{c,t})^T}$, and s is the distance of x to $K_{c,t,\widetilde{\phi}}$.

Signed distance $\mathrm{dist}_{K_{c,t,\widetilde{\phi}}}$. To write the Allen–Cahn approximation of the immersions in Lemma 6.4 we will need to use the following notion of signed distance to $K_{c,t,\widetilde{\phi}}$. Recall that $\widetilde{\phi} \geq 0$ is smooth and $\widetilde{\phi} = 0$ in a neighbourhood of ∂K. In the coordinates of $V_{\widetilde{M}}$, $K_{c,t,\widetilde{\phi}}$ is identified with a graph, namely (for $c \in [0, c_0]$ and $t \in [0, t_0]$)

$$F^{-1}\big(K_{c,t,\widetilde{\phi}}\big) = \big\{(q, s) \in K \times [0, c_K) : s = c + t\widetilde{\phi}(q)\big\}.$$

We define, on $K \times (0, c_K)$, the following "signed distance to $F^{-1}\big(K_{c,t,\widetilde{\phi}}\big)$" for $c > 0$. First, we decide the sign of the distance: we say that $(q, s) \in K \times (0, c_K)$ has negative distance to $F^{-1}\big(K_{c,t,\widetilde{\phi}}\big)$ if $s < c + t\widetilde{\phi}(q)$ and positive distance to $F^{-1}\big(K_{c,t,\widetilde{\phi}}\big)$ if $s > c + t\widetilde{\phi}(q)$. Next we define its modulus. The modulus of the signed distance is the unsigned distance of (q, s) to $F^{-1}\big(K_{c,t,\widetilde{\phi}}\big)$ in $K \times (0, c_K)$ (recall that $K \times (0, c_K)$ is endowed with the Riemannian metric pulled back from N). Note that if $(q, s) \in F^{-1}\big(K_{c,t,\widetilde{\phi}}\big)0$, then the distance extends smoothly at (q, s) with value 0. Also, note that we do not define the signed distance on $K \times \{0\}$. The signed distance just defined descends to a smooth function on $F(K \times (0, c_K)) \subset N$ that we will denote by $\mathrm{dist}_{K_{c,t,\widetilde{\phi}}}$. The set $F(K \times (0, c_K))$ is an open tubular neighbourhood of $\iota(K)$ of semiwidth c_K, with M removed.

7 | ALLEN–CAHN APPROXIMATIONS AND PATHS IN $W^{1,2}(N)$

The overall aim in the sections that follow is to produce, for all sufficiently small ε, a continuous path in $W^{1,2}(N)$ that starts at the constant -1, ends at the constant $+1$ and such that $\mathcal{E}_\varepsilon$ is bounded by $\approx 2\mathcal{H}^n(M) - \min\{\frac{\mathcal{H}^n(B)}{2}, \frac{\tau}{2}\}$, where B and τ were chosen respectively in Remark 5.4 and Lemma 6.4 and depend only on geometric data (not on ε). Theorem 1.3 (and Theorems 1.1, 1.8) will follow immediately once this is achieved.

7.1 | Choice of ε

Let B be as in Remark 5.4 and c_0, t_0, τ be as in Lemma 6.4. The geometric quantities $\mathcal{H}^n(B)$ and τ are relevant in the forthcoming construction.

In the following sections we are going to exhibit, for every sufficiently small ε, a continuous path in $W^{1,2}(N)$ with $\mathcal{E}_\varepsilon$ suitably bounded along the whole path. We will specify now an initial choice $\varepsilon < \varepsilon_1$ that permits the construction of the $W^{1,2}$-functions describing the path. When we will estimate $\mathcal{E}_\varepsilon$ along the path, we will do so in terms of geometric quantities (typically, areas of certain hypersurfaces, hence independent of ε) plus errors that will depend on ε. For sufficiently small ε, that is, $\varepsilon < \varepsilon_2$ for a choice of $\varepsilon_2 \leq \varepsilon_1$ to be specified, these errors will be $\leq C(\varepsilon\,|\log\varepsilon|)$ for some $C > 0$ independent of ε; we will not keep track of the constants and will instead write $O(\varepsilon\,|\log\varepsilon|)$. At the very end (Section 8), in order to make these errors much smaller than τ and $\mathcal{H}^n(B)$, and thus have an effective estimate on $\mathcal{E}_\varepsilon$ along the path, we may need to revisit the smallness choice: for some ε_3, (possibly $\varepsilon_3 \leq \varepsilon_2$) we will get that for $\varepsilon < \varepsilon_3$ the errors can be absorbed in the geometric quantities. Therefore for $\varepsilon < \varepsilon_3$, we will have an upper bound for $\mathcal{E}_\varepsilon$ along the path that is independent of ε.

Now we choose ε_1. The choices of ε_2, ε_3 will be made as we proceed into the forthcoming arguments. We restrict to $\varepsilon_1 < 1$, so that the $O(\varepsilon^2)$ controls that we have on the approximated one-dimensional solutions in Section 2.2 are valid for all $\varepsilon < \varepsilon_1$. We then require $\varepsilon_1 < \frac{1}{e}$ so to have $\varepsilon\,|\log\varepsilon|$ is decreasing as ε decreases so that the conditions specified on ε_1 hold also for each $\varepsilon < \varepsilon_1$ and, moreover,

$$6\varepsilon_1|\log\varepsilon_1| < \frac{c_0}{20}$$

(and implicitly $< \frac{1}{2}\omega$). Since the quantity $6\varepsilon|\log\varepsilon|$ will appear frequently (due to the choice of truncation in Section 2.2), we will use the shorthand notation $\Lambda = 3|\log\varepsilon|$ when working at fixed ε.

7.2 | Allen–Cahn approximation of $2(|M| - |B|)$

Recall the function $G_0^\varepsilon : N \to \mathbb{R}$ defined in (7), which is an Allen–Cahn approximation of $\iota : \widetilde{M} \to N$, that is, a $W^{1,2}$ function with nodal set close to the image of ι and such that its Allen–Cahn energy $\mathcal{E}_\varepsilon(G_0^\varepsilon)$ is approximately[9] the area of ι (i.e., $\approx 2\mathcal{H}^n(M)$). Due to the fact that we replace

[9] In Section 4 we only established an upper bound for $\mathcal{E}_\varepsilon(G_0^\varepsilon)$, and most of the time an upper bound is all that will matter for our Allen–Cahn approximations (although a lower bound in terms of the area of the corresponding immersion is also going to be always valid). In the case of G_0^ε, such a lower bound for $\mathcal{E}_\varepsilon(G_0^\varepsilon)$ will be established later.

hypersurfaces by nonsharp transitions, the function G_0^ε can also be thought of as an Allen–Cahn approximation of $\Gamma_{2\varepsilon\Lambda}$ (that is exactly the nodal set of G_0^ε).

Definition of f. We will now "remove the ball B" from $G_0^\varepsilon : N \to \mathbb{R}$. In other words, we will write an Allen–Cahn approximation f of $2(|M| - |B|)$, or, equivalently, of $\iota|_{\widetilde{M}\setminus\widetilde{B}}$. Always because we have nonsharp transitions, we can think of f also as an Allen–Cahn approximation of $\Gamma_{2\varepsilon\Lambda}$ with two balls removed. Although $f = f^\varepsilon$ does depend on ε, we drop the ε for notational convenience. What is important to keep in mind is that we can perform the construction of f given below for any $\varepsilon < \varepsilon_1$ and that we will obtain estimates on $\mathcal{E}_\varepsilon(f^\varepsilon)$ that are uniform in ε.

To this end, we let $\chi \in C_c^\infty(\widetilde{M})$ be smooth and even (i.e., $\chi(p) = \chi(q)$ if $\iota(p) = \iota(q)$), with $\chi = 1$ on $\widetilde{B}$, $|\nabla\chi| \le \frac{2}{R}$, where R is the radius of B, and $\operatorname{supp}\chi \Subset \widetilde{D}$. Then we define, using coordinates $(q, s) \in K \times [0, c_K) \subset V_{\widetilde{M}}$,

$$G_{0,B}^\varepsilon(q, s) = \Psi_{4\varepsilon\Lambda\chi(q)}(s), \tag{13}$$

where Ψ_t is as in (3). Since χ is even, the function $G_{0,B}^\varepsilon$ descends to a well-defined function f on $F(K \times [0, c_K))$ (this is a tubular neighbourhood of semiwidth c_K around $\iota(K)$). Note that f agrees with G_0^ε on $F((K \setminus \widetilde{D}) \times [0, c_K))$ and on $F((K \times (c_K/2, c_K))$ (on the latter both are equal to -1); therefore we extend f to N by setting it equal to G_0^ε on $N \setminus F(K \times [0, c_K))$,

$$f(x) = \begin{cases} G_0^\varepsilon & \text{for } x \in N \setminus F(K \times [0, c_K)), \\ G_{0,B}^\varepsilon(F^{-1}(x)) & \text{for } x \in F(K \times [0, c_K)); \end{cases} \tag{14}$$

then f is $W^{1,\infty}$ on the complement of $F(\widetilde{D} \times [0, c_K/2])$. Since $\Psi_t(x)$ is even and Lipschitz on $\mathbb{R}$ (see (3)), we will in fact conclude that f is $W^{1,\infty}$ on N. We only need to check it around points $x \in D$. Let $\chi_0 : M \to \mathbb{R}$ be defined by $\chi_0(y) = \chi(\iota^{-1}(y))$; this is a smooth function compactly supported in D. In a neighbourhood of $x \in D$ we can choose a small geodesic ball $B_r(x) \subset M$ and use Fermi coordinates $(y, a) \in B_r(x) \times (-c_K, c_K)$. Then in this neighbourhood $f(y, a) = \Psi_{4\varepsilon\Lambda\chi_0(y)}(a)$. Since $\Psi_t(z)$ is Lipschitz in $(t, z) \in [0, \infty) \times \mathbb{R}$, we conclude that f is Lipschitz on $B_r(x) \times (-c_K, c_K)$. (The Jacobian factor that measures the distortion of the Riemannian metric from the product metric on $B_r(x) \times (-c_K, c_K)$ is bounded by a constant that only depends on the geometric data $F(K) \subset M \subset N$; therefore it suffices to observe that $\Psi_{4\varepsilon\Lambda\chi_0(y)}(a)$ is Lipschitz with respect to the product metric.) Therefore $f \in W^{1,\infty}(N)$.

Allen–Cahn energy of f. To estimate from above the Allen–Cahn energy of f, since $f = G_0^\varepsilon$ in the complement of $F(\widetilde{D} \times [0, c_K))$ and we estimated $\mathcal{E}_\varepsilon(G_0^\varepsilon)$ in (8), we only need to compute the energy of f on $F\big(\widetilde{D} \times [0, c_K)\big)$ (and, similarly, the energy of G_0^ε on $F\big(\widetilde{D} \times [0, c_K)\big)$. We can therefore use coordinates (q, s) on $\widetilde{D} \times [0, c_K) \subset V_{\widetilde{M}}$ as in (13) and apply the coarea formula (for the function $\Pi_K(q, s) = (q, 0)$, whose Jacobian determinant $|J\Pi_K|$ is computed with respect to the Riemannian metric induced from N):

$$\begin{aligned} &\int_{F(\widetilde{D}\times[0,c_K))} \varepsilon\frac{|\nabla f|^2}{2} + \frac{W(f)}{\varepsilon} \\ &= \int_{\widetilde{B}\times(0,c_K)} \left(\frac{\varepsilon}{2}|\nabla G_{0,B}^\varepsilon|^2 + \frac{W(G_{0,B}^\varepsilon)}{\varepsilon}\right) \\ &\quad + \int_{\widetilde{D}\setminus\widetilde{B}} \int_{(0,c_K)} \frac{1}{|J\Pi_K|}\left(\frac{\varepsilon}{2}\left|\frac{\partial}{\partial s}G_{0,B}^\varepsilon\right|^2 + \frac{W(G_{0,B}^\varepsilon)}{\varepsilon}\right) ds\, dq \\ &\quad + \int_{(\widetilde{D}\setminus\widetilde{B})\times(0,c_K)} \frac{\varepsilon}{2}|\nabla_q G_{0,B}^\varepsilon|^2. \end{aligned} \tag{15}$$

The notation ∇_q stands for the gradient projected onto the level sets of s (recall that $\frac{\partial}{\partial s}$ is orthonormal to the level sets of s). By definition of $G^\varepsilon_{0,B}$ we have, at $(q,z) \in \widetilde{D} \times (0, c_K)$:

$$\frac{\partial G^\varepsilon_{0,B}}{\partial q_i} = \frac{d}{da}(\Psi_a)(z)\Big|_{a=4\,\varepsilon\,\Lambda\chi(q)} 4\,\varepsilon\,\Lambda \frac{\partial \chi}{\partial q_i},$$

where, with a slight abuse of notation, $\chi(q,z) = \chi(q)$. As a function on $\widetilde{M}$, χ satisfies $|\nabla\chi| \le \frac{2}{R}$ (where R denotes the radius of B). Moreover, $\left|\frac{d}{da}(\Psi_a)(z)\right| = |\Psi'(|z| + a)| \le \frac{3}{\varepsilon}$. These bounds imply ($\Lambda = 3|\log\varepsilon|$)

$$\varepsilon\,|\nabla_q G^\varepsilon_{0,B}|^2 \le \varepsilon\,\frac{C}{\varepsilon^2}\frac{\varepsilon^2\,|\log\varepsilon|^2}{R^2} = \frac{C\,\varepsilon\,|\log\varepsilon|^2}{R^2}. \tag{16}$$

(Here $C = (8\cdot 6)^2 C'$, where $C' > 0$ depends on the distortion factor between the Riemannian metric and the product metric.) Since $\widetilde{B}, \widetilde{D}, R$, and C are independent of ε, (16) implies that the third term on the right-hand side of (15) can be made arbitrarily small by choosing ε sufficiently small; this term is $O(\varepsilon^2\,|\log\varepsilon|^3)$, since the integrand is zero on $(\widetilde{D}\setminus\widetilde{B}) \times (4\,\varepsilon\,\Lambda, c_K)$. The first term on the right-hand side of (15) vanishes because $G^\varepsilon_{0,B} = -1$ on that domain. For the second term on the right-hand side of (15), note that the inner integral only gives a contribution in $[0, 4\,\varepsilon\,\Lambda]$ ($G^\varepsilon_{0,B} = -1$ on $s \in [4\,\varepsilon\,\Lambda, c_K]$). Recalling the bounds on the Jacobian factor $|J\Pi_K|$ given in (11) and the energy estimates on the one-dimensional profiles (see (1) and (3), we find second term on right-hand side of (15)

$$\begin{aligned}
&\le (1 + 8\,\varepsilon\,\Lambda C_K)\int_{\widetilde{D}\setminus\widetilde{B}}\left(\int_0^{4\,\varepsilon\,\Lambda}\frac{1}{2}\,\varepsilon\big(\Psi'_{4\,\varepsilon\,\Lambda\chi(q)}\big)^2 + \frac{W(\Psi_{4\,\varepsilon\,\Lambda\chi(q)})}{\varepsilon}\right)dq\\
&\le \mathcal{H}^n(\widetilde{D}\setminus\widetilde{B})\,(1 + 8\,\varepsilon\,\Lambda C_K)\,\mathcal{E}_\varepsilon(\overline{\mathbb{H}}^\varepsilon)\\
&\le (\mathcal{H}^n(\widetilde{D}) - \mathcal{H}^n(\widetilde{B}))\,(1 + 8\,\varepsilon\,\Lambda C_K)(2\sigma + O(\varepsilon^2)).
\end{aligned}$$

We can thus rewrite (15) as a leading term $2\sigma(\mathcal{H}^n(\widetilde{D}) - \mathcal{H}^n(\widetilde{B}))$ plus errors; for a sufficiently small choice of $\varepsilon_2 \le \varepsilon_1$ for $\varepsilon < \varepsilon_2$ all errors are of the type $O(\varepsilon\,|\log\varepsilon|)$. We therefore conclude that the following estimate holds for all $\varepsilon < \varepsilon_2$:

$$\int_{F(\widetilde{D}\times[0,c_K))} \varepsilon\,\frac{|\nabla f|^2}{2} + \frac{W(f)}{\varepsilon} \le 4\sigma(\mathcal{H}^n(D) - \mathcal{H}^n(B)) + O(\varepsilon\,|\log\varepsilon|).$$

Going back to G^ε_0, we can give a lower bound to its energy on $F(\widetilde{D}\times[0,c_K))$ with a computation analogous to the one just carried out. With coordinates $(q,s) \in D \times [0, c_K)$ we have that G^ε_0 is the function $\Psi(s)$ and therefore $|\nabla G^\varepsilon_0|$ is given by $\left|\frac{\partial}{\partial s}\Psi(s)\right|$ (the gradient is parallel to the $\frac{\partial}{\partial s}$). Using the coarea formula (again[10] with Π_K) we get

$$\begin{aligned}
&\int_{F(\widetilde{D}\times[0,c_K))} \varepsilon\,\frac{|\nabla G_0^\varepsilon|^2}{2}+\frac{W(G_0^\varepsilon)}{\varepsilon}\\
&=\int_{\widetilde{D}}\left(\int_0^{4\,\varepsilon\,\Lambda}\frac{1}{|J\Pi_K|}\left(\frac{\varepsilon}{2}\left|\frac{\partial}{\partial s}\Psi(s)\right|^2+\frac{W(\Psi(s))}{\varepsilon}\right)ds\right)dq\\
&\geq \mathcal{H}^n(\widetilde{D})(1-8\,\varepsilon\,\Lambda C_K)(2\sigma+O(\varepsilon^2)),
\end{aligned}\tag{17}$$

where we used (11), (1), and (3). The result in (17) is of the form $4\sigma\mathcal{H}^n(D)$ plus errors. The errors are of the form $O(\varepsilon\,|\log\varepsilon|)$ for all $\varepsilon<\varepsilon_2$ for some suitably small choice of $\varepsilon_2\leq\varepsilon_1$.

Remark 7.1 (On the choice of ε_2). We make the choice of ε_2 several times along the construction, always within the scope of making the errors controlled by $C\,\varepsilon\,|\log\varepsilon|$ with C independent of $\varepsilon\in(0,\varepsilon_2)$. The specific value ε_2 might change from one instance to the next, but since we make finitely many choices we implicitly assume that the correct ε_2 is the smallest of all. From now on, this remark will apply every time we say that the errors are of the form $O(\varepsilon\,|\log\varepsilon|)$ for all $\varepsilon<\varepsilon_2$ for some suitably small choice of ε_2.

In conclusion, for all $\varepsilon<\varepsilon_2$ we have that

$$\begin{aligned}
&\frac{1}{2\sigma}\int_{F(D\times[0,c_K))}\varepsilon\,\frac{|\nabla G_0^\varepsilon|^2}{2}+\frac{W(G_0^\varepsilon)}{\varepsilon}\\
&-\frac{1}{2\sigma}\int_{F(D\times[0,c_K))}\varepsilon\,\frac{|\nabla f|^2}{2}+\frac{W(f)}{\varepsilon}\geq 2\mathcal{H}^n(B)-|O(\varepsilon\,|\log\varepsilon|)|.
\end{aligned}\tag{18}$$

Recall that f does depend on ε, although we are not expliciting the dependence for notational convenience, and that we can produce f (as defined above) for every $\varepsilon<\varepsilon_1$. By (8) and (18), and the fact that $f=G_0^\varepsilon$ on $N\setminus F(\widetilde{D}\times[0,c_K))$, we conclude that for a sufficiently small choice of $\varepsilon_2\leq\varepsilon_1$, for all $\varepsilon<\varepsilon_2$, the following estimate holds:

$$\mathcal{E}_\varepsilon(f)\leq 2\big(\mathcal{H}^n(M)-\mathcal{H}^n(B)\big)+O(\varepsilon\,|\log\varepsilon|).\tag{19}$$

This says that f is a good[11] Allen–Cahn approximation of $2(|M|-|B|)$. In terms of the immersions of Lemma 6.4, f is also an Allen–Cahn approximation of $q=(y,v)\in\mathrm{Int}(K_B)\to\exp_y(2\,\varepsilon\,\Lambda v)$ (the nodal set of f contains the image of this immersion with boundary).

7.3 | From $\mathcal{E}_\varepsilon(-1)=0$ to $2(|M|-|B|)$

In this section we construct a continuous path in $W^{1,2}(N)$ that joins f to the constant -1, keeping $\mathcal{E}_\varepsilon$ along the path controlled by $\mathcal{E}_\varepsilon(f)$.

[10] It would also be possible to use the coarea formula slicing by the distance to M, as done in (8), making use of Lemma 4.1.

[11] We only need the upper bound (19), however a lower bound of the form $\mathcal{E}_\varepsilon(f)\geq 2(\mathcal{H}^n(\iota(K))-\mathcal{H}^n(D))-O(\varepsilon\,|\log\varepsilon|)$ is also easily seen to be valid.

We begin by introducing the following one-parameter family of functions: for $r \in [0, 4\varepsilon\Lambda]$ define

$$Y_r^\varepsilon(x) = \begin{cases} -1 & \text{for } x \in N \setminus T_\omega, \\ \Psi_r(d_{\overline{M}}(x)) & \text{for } x \in T_\omega, \end{cases} \tag{20}$$

where Ψ_r is as in (3). Since $\overline{\mathbb{H}}^\varepsilon$ is constantly -1 on $(-\infty, -2\varepsilon\Lambda]$, the function Y_t^ε is constantly -1 on $\{x : d_{\overline{M}}(x) > 4\varepsilon\Lambda - r\}$. Moreover, since $d_{\overline{M}}$ is Lipschitz on N and Ψ_r is Lipschitz on $\mathbb{R}$, denoting the Lipschitz constants of Ψ_r and $d_{\overline{M}}$, respectively, by $C_{\Psi_r}, C_{d_{\overline{M}}}$, we have

$$|\Psi_r(d_{\overline{M}}(x)) - \Psi_r(d_{\overline{M}}(y))| \le C_{\Psi_r}|d_{\overline{M}}(x) - d_{\overline{M}}(y)| \le C_{\Psi_r} C_{d_{\overline{M}}} \operatorname{dist}_N(x, y).$$

Therefore $Y_r^\varepsilon \in W^{1,\infty}(N)$.

Notice that $Y_0^\varepsilon = G_0^\varepsilon$. We compute $\mathcal{E}_\varepsilon(Y_r^\varepsilon)$ by using the coarea formula (slicing by the distance function $d_{\overline{M}}$, for which $|\nabla d_{\overline{M}}| = 1$) as we did for G_0^ε (see (8)). We obtain

$$\begin{aligned} \mathcal{E}_\varepsilon(Y_r^\varepsilon) &\le 2\mathcal{H}^n(M)\left(\frac{1}{2\sigma}\int_0^{4\varepsilon\Lambda - r} \varepsilon\frac{(\Psi_r')^2}{2} + \frac{W(\Psi_r)}{\varepsilon}\right) \\ &\le 2\mathcal{H}^n(M)(1 + O(\varepsilon^2)), \end{aligned} \tag{21}$$

using (1) and the fact that

$$\int_0^{4\varepsilon\Lambda - r} \varepsilon\frac{(\Psi_r')^2}{2} + \frac{W(\Psi_r)}{\varepsilon} \le \int_0^\infty \varepsilon\frac{(\Psi_r')^2}{2} + \frac{W(\Psi_r)}{\varepsilon} = 2\sigma + O(\varepsilon^2).$$

Note that $\mathcal{E}_\varepsilon(Y_r^\varepsilon) \to 0$ as $r \to 4\varepsilon\Lambda$.

Now we give a lower bound for the energy of Y_r^ε on the domain $F(\widetilde{D} \times [0, c_K))$ as we did for G_0^ε in (17), that is, using the coarea formula for the function Π_K. Note that on this domain we can use the coordinates (q, s) on $\widetilde{D} \times [0, c_K)$ and the fact that the gradient of Y_r^ε is parallel to $\frac{\partial}{\partial s}$. We have

$$\begin{aligned} &\int_{F(\widetilde{D}\times[0,c_K))} \varepsilon\frac{|\nabla Y_r^\varepsilon|^2}{2} + \frac{W(Y_r^\varepsilon)}{\varepsilon} \\ &= \int_{\widetilde{D}}\left(\int_0^{4\varepsilon\Lambda - r} \frac{1}{|J\Pi_K|}\left(\frac{\varepsilon}{2}\left|\frac{\partial}{\partial s}\Psi_r(s)\right|^2 + \frac{W(\Psi_r(s))}{\varepsilon}\right)ds\right)dq \\ &\ge 2\mathcal{H}^n(D)(1 - 8\varepsilon\Lambda C_K)\int_0^{4\varepsilon\Lambda - r} \varepsilon\frac{(\Psi_r')^2}{2} + \frac{W(\Psi_r)}{\varepsilon}, \end{aligned} \tag{22}$$

where we used (11) and the fact that $\frac{\varepsilon}{2}\left|\frac{\partial}{\partial s}\Psi_r(s)\right|^2 + \frac{W(\Psi_r(s))}{\varepsilon}$ is independent of q. We therefore conclude, from the first inequality in (21) and from (22), the following estimate for the Allen-Cahn energy of Y_r^ε in $N \setminus F(\widetilde{D} \times [0, c_K))$: there exists $\varepsilon_2 \le \varepsilon_1$ sufficiently small such that for all $\varepsilon < \varepsilon_2$

$$\begin{aligned} &\int_{N\setminus F(\widetilde{D}\times[0,c_K))} \varepsilon\frac{|\nabla Y_r^\varepsilon|^2}{2} + \frac{W(Y_r^\varepsilon)}{\varepsilon} \\ &\le 2(\mathcal{H}^n(M) - \mathcal{H}^n(D))\int_0^{4\varepsilon\Lambda - r} \varepsilon\frac{(\Psi_r')^2}{2} + \frac{W(\Psi_r)}{\varepsilon} + O(\varepsilon|\log\varepsilon|). \end{aligned} \tag{23}$$

Definition of the path f_r. We now define a continuous path $r \in [0, 4\varepsilon\Lambda] \to f_r \in W^{1,2}(N)$ as follows. Recalling the definition of $\chi \in C_c^\infty(\widetilde{M})$ and using coordinates $(q, s) \in \widetilde{D} \times [0, c_K)$ we set

$$Y_{r,B}(q, s) = \Psi_{4\varepsilon\Lambda\chi(q)+r}(s),$$

where Ψ_t is as in (3). The function $f_r : N \to \mathbb{R}$ is then defined by

$$f_r(x) = \begin{cases} Y_r^\varepsilon(x) & \text{if } x \in N \setminus F(\widetilde{D} \times [0, c_K)), \\ Y_{r,B}(F^{-1}(x)) & \text{if } x \in F(\widetilde{D} \times [0, c_K)). \end{cases} \tag{24}$$

Note that f_r is well-defined on D since χ is even. Note also that for $r = 0$ this function is f and for $r = 4\varepsilon\Lambda$ it is the constant -1. Moreover, $f_r \in W^{1,\infty}(N)$ for every r. To see this, notice that $Y_{r,B}$ is smooth on $\widetilde{D} \times (0, c_K)$, so f_r is smooth on $F(\widetilde{D} \times (0, c_K))$. Moreover,

$$f_r \in W^{1,\infty}(N \setminus F(\overline{\widetilde{D}} \times [0, c_K]))$$

because it agrees with Y_r^ε on this open set. The smoothness at $F(\overline{\widetilde{D}} \times \{c_K\})$ is immediate because $f_r = -1$ in a neighbourhood of $F(\overline{\widetilde{D}} \times \{c_K\})$. We thus only need to check that f_r is Lipschitz locally around any point $x \in \overline{D}$. Using Fermi coordinates $(y, a) \in B(x) \times (-\delta, \delta)$, where $B(x)$ is a small geodesic ball in M centred at x and $\delta > 0$, we have the following expression for f, thanks to the fact that $\Psi_r : \mathbb{R} \to \mathbb{R}$ is even for every r: $f_r(y, a) = \Psi_{4\varepsilon\Lambda\chi_0(y)+r}(a)$, where $\chi_0(p) = \chi(F^{-1}(p))$. Since $\Psi_r(z)$ is Lipschitz on $\{(r, z) : r \in [0, \infty), z \in \mathbb{R}\}$, and since χ_0 is smooth, we obtain that $f_r \in W^{1,\infty}$ on the chosen neighbourhood of x. (As we did in (14), we use the fact that being Lipschitz for the product metric on $B(x) \times (-\delta, \delta)$ implies Lipschitz with respect to the Riemannian metric induced from N.) In conclusion, we have $f_r \in W^{1,\infty}(N)$.

The path $r \in [0, 4\varepsilon\Lambda] \to f_r \in W^{1,2}(N)$ is moreover continuous. Let us check the continuity of ∇f_r in r (with respect to the L^2-topology on N). The partial derivatives of f_r on $F(\widetilde{D} \times [0, c_K))$ are given by, using (q, s)-coordinates on $\widetilde{D} \times (0, c_K)$:

$$\left(\dots, 4\varepsilon\Lambda \frac{\partial\chi(q)}{\partial q_i}\Psi_0'(s + 4\varepsilon\Lambda\chi(q) + r), \dots, \Psi_0'(4\varepsilon\Lambda\chi(q) + r + s)\right).$$

By continuity of translations in L^p, and smoothness of χ and of the Riemannian metric, we get that ∇f_r is continuous in r (with respect to the L^2-topology, or even L^p for any p). Similarly, we can argue for $T_\omega \setminus F(\widetilde{D} \times [0, c_K))$, where $f = Y_r^\varepsilon$ and the gradient is $\Psi_0'(r + d_{\overline{M}}(x))\nabla d_{\overline{M}}(x)$: this changes continuously with r (with respect to the L^2-topology, or even L^p for any p). Therefore we have that $r \in [0, 4\varepsilon\Lambda] \to \nabla f_r \in L^2(N)$ is continuous. The fact that f_r changes continuously in r with respect to the L^2-topology is even more straightforward.

Energy along the path. To estimate $\mathcal{E}_\varepsilon(f_r)$ we compute the energy on $F(\widetilde{D} \times [0, c_K))$ using the coarea formula for Π_K, similarly to (19), in the coordinates $(q, s) \in \widetilde{D} \times [0, c_K)$. Notice that $Y_{0,B}^\varepsilon(q, s) = -1$ for $q \in B$. Then we obtain

$$\int_{F(\widetilde{D}\times[0,c_K))} \varepsilon\frac{|\nabla f_r|^2}{2} + \frac{W(f_r)}{\varepsilon}$$
$$= \int_{\widetilde{D}\setminus\widetilde{B}} \int_0^{c_K} \frac{1}{|J\Pi_K|}\left(\varepsilon\frac{|\Psi_r'(s)|^2}{2} + \frac{W(\Psi_r(s))}{\varepsilon}\right) ds\, dq \underbrace{\leq}_{(11)}$$

$$\leq (1+8\,\varepsilon\,\Lambda C_K)\int_{\widetilde{D}\setminus\widetilde{B}}\int_0^{c_K}\left(\varepsilon\,\frac{|\Psi_r'(s)|^2}{2}+\frac{W(\Psi_r(s))}{\varepsilon}\right)ds\,dq$$

$$\leq 2(1+8\,\varepsilon\,\Lambda C_K)\mathcal{H}^n(D\setminus B)\left(\int_r^{4\,\varepsilon\,\Lambda}\varepsilon\,\frac{(\Psi')^2}{2}+\frac{W(\Psi)}{\varepsilon}\right).$$

Recalling that $f_r = Y_r^\varepsilon$ on $N\setminus F(\widetilde{D}\times[0,c_K))$ and by the estimate in (23) we conclude that there exists $\varepsilon_2\leq\varepsilon_1$ such that for all $\varepsilon\leq\varepsilon_2$ the following estimates hold for $r\in[0,4\,\varepsilon\,\Lambda]$:

$$\begin{aligned}\mathcal{E}_\varepsilon(f_r)&\leq 2(\mathcal{H}^n(M)-\mathcal{H}^n(B))\left(\frac{1}{2\sigma}\int_r^{4\,\varepsilon\,\Lambda}\varepsilon\,\frac{(\Psi')^2}{2}+\frac{W(\Psi)}{\varepsilon}\right)\\&\quad+O(\varepsilon\,|\log\varepsilon|),\\\mathcal{E}_\varepsilon(f_r)&\leq 2(\mathcal{H}^n(M)-\mathcal{H}^n(B))+O(\varepsilon\,|\log\varepsilon|).\end{aligned}\tag{25}$$

(The second follows from the first since the energy of Ψ in parentheses is $\leq 1+O(\varepsilon^2)$.) The second estimate shows the uniform energy control on $r\in[0,4\,\varepsilon\,\Lambda]$; the first shows that $\mathcal{E}_\varepsilon(f_r)\to 0$ as $r\to 4\,\varepsilon\,\Lambda$.

Remark 7.2. At least for $n\leq 6$ it is possible to produce a continuous path from f to -1, with a similar energy control as in (25), by employing alternatively a negative $\mathcal{E}_\varepsilon$-gradient flow starting at a suitably constructed function f_2 that is $W^{1,2}$-close to f and with $\mathcal{E}_\varepsilon(f_2)\approx\mathcal{E}_\varepsilon(f)$. One can choose f_2 such that the flow is mean convex and converges (decreasingly) to the constant -1, reaching it in time $O(\varepsilon\,|\log\varepsilon|)$. The $\varepsilon\to 0$ limit of such paths is then the Brakke flow that starts at $2(|M|-|B|)$ and vanishes instantaneously. The family (24) that we gave in this section mimics exactly this flow; however, it is more elementary, even for $n\leq 6$, as we can exhibit the path explicitly (and moreover present no additional difficulties for $n\geq 7$). Note that the path f_r that we produced also reaches -1 in time $O(\varepsilon\,|\log\varepsilon|)$.

7.4 | Lowering the peak

In this section we construct the next portion of our path, starting at f. The immersions in Lemma 6.4 are particularly relevant, as they provide the geometric counterpart of this portion of the $W^{1,2}$-path: first we use the immersions in (i) of Lemma 6.4, keeping $c=2\,\varepsilon\,\Lambda$ and increasing t from 0 to t_0; then we connect the final immersion just obtained to the one in (ii) of Lemma 6.4 with $t=t_0$ and $c=2\,\varepsilon\,\Lambda$ (in doing so, we "close the hole at B"). The portion of the path that we exhibit in this section is made of Allen–Cahn approximations of the immersions just described. It is this portion of the path that "lowers the peak" of $\mathcal{E}_\varepsilon$ (compare Figure 4), keeping it a fixed amount below $2\mathcal{H}^n(M)$ (thanks to the estimates in Lemma 6.4).

We will keep using the shorthand notation $\Lambda=3|\log\varepsilon|$. All the functions that we will construct in this section coincide with G_0^ε in the complement of $F(K\times[0,c_K))$. By construction they will in fact agree with G_0^ε in a neighbourhood of $\partial F(K\times[0,c_K))$ (guaranteeing a smooth patching), and thanks to Remark 6.2 and since $2\,\varepsilon\,\Lambda<c_0/20$ (Section 7.1) we can use tubular neighbourhoods of semiwidth $2\,\varepsilon\,\Lambda$ around $K_{c,t,\widetilde{\phi}}$ for every $c\geq 2\,\varepsilon\,\Lambda$ to define Allen–Cahn approximations of the immersions in Lemma 6.4.

Recall the notation $K_{c,t,\widetilde{\phi}}$ from Section 6: it denotes the image via $F:V_{\widetilde{M}}\to N$ of the graph $\{(q,s)\in V_{\widetilde{M}}\,:\,q\in K, s=c+t\widetilde{\phi}(q)\}$ for $t\in[0,t_0]$ and $c\in[0,c_0]$. In other words, $K_{c,t,\widetilde{\phi}}$ is the image of the immersion (smoothly extended up to ∂K; see Remark 6.1) $q = (y, v) \in K \to \exp_y((c + t\widetilde{\phi}(q))v)$. Recall the definition of the signed distance provided in Section 6 and denote by $\text{dist}_{K_{c,t,\widetilde{\phi}}}$ the signed distance to $K_{c,t,\widetilde{\phi}}$, well-defined on $F(\widetilde{D} \times (0, c_K))$. If $t = 0$, then $\text{dist}_{K_{c,0,\widetilde{\phi}}}$ extends continuously to $F(\widetilde{D} \times [0, c_K))$ with value $-c$ on $F(\widetilde{D} \times \{0\})$. With this in mind, the definition of f in (13)–(14), can equivalently be given as follows:

$$f(x) = \begin{cases} \overline{\mathbb{H}}^{\varepsilon}_{4\,\varepsilon\,\Lambda\chi_0(\Pi_K(x))}\left(-\text{dist}_{K_{2\,\varepsilon\,\Lambda,0,\widetilde{\phi}}}(x)\right) & \text{for } x \in F(K \times [0, c_K)), \\ G_0^{\varepsilon}(x) & \text{for } x \in N \setminus F(K \times [0, c_K)), \end{cases}$$

where

$$\overline{\mathbb{H}}^{\varepsilon}_s(\cdot) = \overline{\mathbb{H}}^{\varepsilon}(\cdot - s),$$

$\chi_0 = \chi \circ F^{-1}$, and, with a slight abuse of notation, $\Pi_K(x)$ is the nearest-point projection of x onto M. (In the coordinates of $V_{\widetilde{M}}$ we have $\Pi_K(q, s) = (q, 0)$, which is the notation used in Section 6; the map on $F(K \times [0, c_K))$ that we are using above should then be $F \circ \Pi_K \circ F^{-1}$, we however denote both the map in $K \times [0, c_K)$ and the map in $F(K \times [0, c_K))$ by the same symbol Π_K.)

Remark 7.3. The signed distance $\text{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}(x)$ is defined on $F(K \times (0, c_K))$. We point out the following facts. Let $x \in F(K \times \{0\})$ and $x_j \to x$, $x_j \in F(K \times (0, c_K))$ (so that the signed distance is negative on x_j for j sufficiently large). Then

$$\limsup_{j \to \infty} \text{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}(x_j) \leq -2\,\varepsilon\,\Lambda.$$

Moreover, $\text{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}$ extends continuously to $F((K \setminus \text{supp}(\widetilde{\phi})) \times \{0\})$ with value $-2\,\varepsilon\,\Lambda$. In particular, the continuous extension is valid on (a neighbourhood of) $\overline{D}$.

Definition of g_t. We construct now the portion of path $t \in [0, t_0] \to g_t \in W^{1,2}(N)$ whose geometric counterpart is given by the immersions in (i) of Lemma 6.4 with $c = 2\,\varepsilon\,\Lambda$ and $t \in [0, t_0]$. These immersions "have a hole at $\widetilde{B}$". We set, for $t \in [0, t_0]$:

$$g_t(x) = \begin{cases} G_0^{\varepsilon}(x) (\text{see} (7)) & \text{for } x \in N \setminus F(K \times [0, c_K)), \\ \overline{\mathbb{H}}^{\varepsilon}_{4\,\varepsilon\,\Lambda\chi_0(\Pi_K(x))}(-\text{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}(x)) & \text{for } x \in F(K \times (0, c_K)) \cup D, \\ 1 & \text{for } x \in F((K \setminus \overline{D}) \times \{0\}). \end{cases} \tag{26}$$

In the second line of (26) we are using the fact that $\text{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}$ is well-defined and continuous on $\overline{D}$, with value $-2\,\varepsilon\,\Lambda$ (see Remark 7.3). Also note that on $F(\partial K \times [0, c_K))$ the definition in the second line agrees with the definition of G_0^{ε} ($\widetilde{\phi}$ vanishes in a neighbourhood of ∂K, see Section 6) and the same is true on $F(K \times \{c_K\})$ ($g_t = -1$ in a neighbourhood). For $t = 0$ we have $g_0 = f$, by the expression of f given earlier in this section.

$g_t \in W^{1,\infty}(N)$ *for each* t. Let us check first that g_t is continuous on N for each t. In view of the comments just made, this needs to be checked only at an arbitrary x in $F((\text{Int}(K) \setminus \overline{D}) \times \{0\})$. Let $x_j \to x$, then for sufficiently large j we have $x_j \in F((\text{Int}(K) \setminus \overline{D}) \times [0, c_K))$. Then $x, x_j \notin \text{supp}\,\chi_0 \times [0, c_K)$. Therefore by (26) we get $g_t(x_j) = \overline{\mathbb{H}}^{\varepsilon}(-\text{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}(x_j))$. Recall Remark 7.3. By continuity of $\overline{\mathbb{H}}^{\varepsilon}$ and the fact that $\overline{\mathbb{H}}^{\varepsilon}(z) = 1$ for $z \geq 2\,\varepsilon\,\Lambda$, we conclude that $g_t(x_j) \to 1$; hence g_t is continuous at x.

To check that $g_t \in W^{1,\infty}(N)$, note first that the definition in the second line of (26) is equal to the one of G_0^{ε} in a neighbourhood of the boundary of $F(K \times [0, c_K))$. Moreover, g_t is smooth on $F(K \times (0, c_K))$ and G_0^{ε} is $W^{1,\infty}(N)$. These facts imply that $g_t \in W^{1,\infty}(N \setminus F(K \times \{0\}))$, and actually even in a neighbourhood of the boundary of $F(K \times [0, c_K))$. Moreover, for $x \in B$ we have $g_t = -1$ in a neighbourhood of x, because $\chi_0 = 1$ on B and $\overline{\mathbb{H}}^{\varepsilon}_{4\,\varepsilon\,\Lambda}(z) \equiv -1$ for $z \leq 2\,\varepsilon\,\Lambda$.

Therefore we only need to check that g_t is locally Lipschitz around points $x \in F((\mathrm{Int}(K) \setminus \widetilde{B}) \times \{0\})$. We distinguish two cases. If $x \notin \overline{D}$, that is, if $x \in F((\mathrm{Int}(K) \setminus \overline{\widetilde{D}}) \times \{0\})$, then g_t is actually C^1 in a neighbourhood of x. This is seen by repeating the argument used above (for the continuity of g_t at such point) to prove that $|\nabla g_t(x_j)| \to 0$ (using the fact that $\overline{\mathbb{H}}^{\varepsilon\prime}$ is smooth on $\mathbb{R}$ and equal to 0 on $[2\,\varepsilon\,\Lambda, \infty)$). We therefore have g_t is C^1 on $F((\mathrm{Int}(K) \setminus \overline{\widetilde{D}}) \times (0, c_K))$, g_t extends continuously to $F((\mathrm{Int}(K) \setminus \overline{\widetilde{D}}) \times \{0\})$ with constant value 1 and ∇g_t extends continuously to this set with value 0. From these facts it follows that the L^{∞} function equal to ∇g_t on $F((\mathrm{Int}(K) \setminus \overline{\widetilde{D}}) \times (0, c_K))$ is the distributional derivative of g_t in a neighbourhood of x, and therefore g_t is C^1 in a neighbourhood of x. In the second case, that is, if $x \in \overline{D} \setminus B$, then for a sufficiently small ball $B_{\rho}(x) \subset M$ we have $\widetilde{\phi} = 0$ on $F^{-1}(B_{\rho}(x))$ (because $\mathrm{supp}(\widetilde{\phi})$ and $\overline{\widetilde{D}}$ are disjoint), and we can use a well-defined system of Fermi coordinates $(y, a) \in B_{\rho}(x) \times (-c_K, c_K)$. In these coordinates we have $\mathrm{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}(F(y,a)) = |a| - 2\,\varepsilon\,\Lambda$, and $g_t(F(y,a)) = \overline{\mathbb{H}}^{\varepsilon}_{4\,\varepsilon\,\Lambda\chi(y)}(-|a| + 2\,\varepsilon\,\Lambda)$, which is Lipschitz in the neighbourhood.

The path $t \to g_t$ is continuous. It suffices to check that the second line in (26) is continuous in t. The proof can be carried out using the coordinates on $V_{\widetilde{M}}$ and the fact that the graph $\{(q, s) : q \in K, s = 2\,\varepsilon\,\Lambda + t\widetilde{\phi}(q)\}$ changes smoothly in t, hence so does the function $\mathrm{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}$. In fact, for our purposes it suffices to observe that if $t_i \to \bar{t}$, then $K_{2\,\varepsilon\,\Lambda,t_i,\widetilde{\phi}}$ converges to $K_{2\,\varepsilon\,\Lambda,\bar{t},\widetilde{\phi}}$ in the Hausdorff distance, from which it follows that $\mathrm{dist}_N(\cdot, K_{2\,\varepsilon\,\Lambda,t_i,\widetilde{\phi}})$ converges pointwise a.e. to $\mathrm{dist}_N(\cdot, K_{2\,\varepsilon\,\Lambda,\bar{t},\widetilde{\phi}})$. This implies that ∇g_{t_i} converges pointwise a.e. to $\nabla g_{\bar{t}}$ and, by dominated convergence (since N is compact and $|\nabla g_t|$ is uniformly bounded independently of t) $\nabla g_{t_i} \to \nabla g_{\bar{t}}$ in $L^2(N)$. The fact that $g_{t_i} \to g_{\bar{t}}$ in $L^2(N)$ follows easily by checking that $t \to g_t$ is a Lipschitz curve with respect to $L^{\infty}(N)$. Therefore the path $t \in [0, t_0] \to g_t \in W^{1,2}(N)$ is continuous.

Energy of g_t. To give an upper bound for $\mathcal{E}_{\varepsilon}(g_t)$ we first need a lower bound for the energy of G_0^{ε} on $F(K \times [0, c_K))$. This is analogous to the estimate in (17):

$$
\begin{aligned}
&\int_{F(K\times[0,c_K))} \varepsilon\,\frac{|\nabla G_0^{\varepsilon}|^2}{2} + \frac{W(G_0^{\varepsilon})}{\varepsilon} \\
&= \int_K \left(\int_0^{4\,\varepsilon\,\Lambda} \frac{1}{|J\Pi_K|} \left(\frac{\varepsilon}{2} \left| \frac{\partial}{\partial s} \Psi_0(s) \right|^2 + \frac{W(\Psi_0(s))}{\varepsilon} \right) ds \right) dq \underbrace{\geq}_{(11)} \\
&\geq (1 - 8\,\varepsilon\,\Lambda C_K) \int_K \left(\int_0^{4\,\varepsilon\,\Lambda} \left(\frac{\varepsilon}{2} \left| \frac{\partial}{\partial s} \Psi_0(s) \right|^2 + \frac{W(\Psi_0(s))}{\varepsilon} \right) ds \right) dq \\
&\geq \mathcal{H}^n(K)(1 - 8\,\varepsilon\,\Lambda C_K) \left(\int_{-2\,\varepsilon\,\Lambda}^{2\,\varepsilon\,\Lambda} \frac{\varepsilon}{2} \left| \overline{\mathbb{H}}^{\varepsilon\prime} \right|^2 + \frac{W(\overline{\mathbb{H}}^{\varepsilon})}{\varepsilon} \right) \\
&= \mathcal{H}^n(K)(1 - 8\,\varepsilon\,\Lambda C_K)(2\sigma + O(\varepsilon^2)),
\end{aligned} \tag{27}
$$

where we used (1).

From (8) and (27) we obtain that, for some suitably small choice of $\varepsilon_2 \leq \varepsilon_1$, for all $\varepsilon < \varepsilon_2$ the following holds for the energy of G_0^ε (and thus also of g_t) in $N \setminus F(K \times [0, c_K))$:

$$\frac{1}{2\sigma}\int_{N\setminus F(K\times[0,c_K))} \varepsilon\,\frac{|\nabla G_0^\varepsilon|^2}{2} + \frac{W(G_0^\varepsilon)}{\varepsilon} \leq \mathcal{H}^n(\widetilde{M}\setminus K) + O(\varepsilon\,|\log\varepsilon|). \tag{28}$$

We now pass to an estimate for the energy of g_t in $F(K \times [0, c_K))$. For this we will use Fermi coordinates for a tubular neighbourhood of $K_{2\varepsilon\Lambda,t,\tilde{\phi}}$ of semiwidth $2\varepsilon\Lambda$. Denote by $(y, a) \in K_{2\varepsilon\Lambda,t,\tilde{\phi}} \times (-2\varepsilon\Lambda, 2\varepsilon\Lambda)$ such coordinates and by $\Pi_{2\varepsilon\Lambda,t}$ the nearest-point projection from the chosen tubular neighbourhood onto $K_{2\varepsilon\Lambda,t,\tilde{\phi}}$ (see Remark 6.2). Notice that $g_t = -1$ on $F(\widetilde{B} \times [0, c_K))$, so there is no energy contribution in this open set. The coarea formula (for the function $\Pi_{2\varepsilon\Lambda,t}$) then gives[12]

$$\begin{aligned}
&\int_{F(K\times[0,c_K))} \varepsilon\,\frac{|\nabla g_{t_0+r}|^2}{2} + \frac{W(g_{t_0+r})}{\varepsilon} \\
&\leq \int_{K_{2\varepsilon\Lambda,t_0,\tilde{\phi}}} \left(\int_{-2\varepsilon\Lambda}^{2\varepsilon\Lambda} \frac{1}{|J\Pi_{2\varepsilon\Lambda,t_0}|}\left(\frac{\varepsilon}{2}\left|\overline{\mathbb{H}}^{\varepsilon\prime}(a)\right|^2 + \frac{W(\overline{\mathbb{H}}^\varepsilon(a))}{\varepsilon}\right)da\right)dy \underbrace{\leq}_{(12)} \\
&\leq (1 + 2\varepsilon\Lambda C_{K,t_0,c_0,\tilde{\phi}})\mathcal{H}^n(K_{2\varepsilon\Lambda,t_0,\tilde{\phi}})\left(\int_{-2\varepsilon\Lambda}^{2\varepsilon\Lambda} \frac{\varepsilon}{2}\left|\overline{\mathbb{H}}^{\varepsilon\prime}\right|^2 + \frac{W(\overline{\mathbb{H}}^\varepsilon)}{\varepsilon}\right) \\
&= (1 + 2\varepsilon\Lambda C_{K,t_0,c_0,\tilde{\phi}})\mathcal{H}^n(K_{2\varepsilon\Lambda,t_0,\tilde{\phi}})(2\sigma + O(\varepsilon^2)).
\end{aligned} \tag{29}$$

Therefore for some suitably small choice of $\varepsilon_2 \leq \varepsilon_1$, for all $\varepsilon < \varepsilon_2$ the following holds:

$$\begin{aligned}
&\frac{1}{2\sigma}\int_{F(K\times[0,c_K))} \varepsilon\,\frac{|\nabla g_t|^2}{2} + \frac{W(g_t)}{\varepsilon} \\
&\qquad\leq \mathcal{H}^n(K_{2\varepsilon\Lambda,t,\tilde{\phi}} \setminus F(\widetilde{B}\times[0,c_K))) + O(\varepsilon\,|\log\varepsilon|).
\end{aligned}$$

Note that $K_{2\varepsilon\Lambda,t,\tilde{\phi}} \setminus F(\widetilde{B} \times [0, c_K))$ is the image of K_B via the immersion in (i) of Lemma 6.4 when $c = 2\varepsilon\Lambda$. Using Lemma 6.4 in the last estimate and putting it together with (28) we finally obtain that, for some suitably small choice of $\varepsilon_2 \leq \varepsilon_1$, for all $\varepsilon < \varepsilon_2$ the following estimate holds for all $t \in [0, t_0]$:

$$\mathcal{E}_\varepsilon(g_t) \leq 2\left(\mathcal{H}^n(M) - \frac{3}{4}\mathcal{H}^n(B)\right) + O(\varepsilon\,|\log\varepsilon|). \tag{30}$$

Definition of g_{t_0+r}: "closing the hole at B". We have constructed a continuous path $t \in [0, t_0] \to g_t \in W^{1,2}(N)$ with $g_0 = f$ and with $\mathcal{E}_\varepsilon$ uniformly controlled by (30), reproducing the middle row of Figure 1. The next portion of the path will start from g_{t_0} and will "close the hole at B". On the geometric side, we are starting at the immersion in Lemma 6.4 (i) with $c = 2\varepsilon\Lambda$ and $t = t_0$, and ending at the immersion in Lemma 6.4 (ii) with $c = 2\varepsilon\Lambda$ and $t = t_0$, reproducing the bottom row

[12] In the first inequality that follows we use the fact that for $y \in K_{2\varepsilon\Lambda,t,\tilde{\phi}} \cap F((\widetilde{D} \setminus \widetilde{B}) \times [0, c_K))$, integration in a is in the domain $-2\varepsilon\Lambda \leq a \leq 2\varepsilon\Lambda(1 - 2\chi(\Pi_K((y, a))))$, and we can bound the top endpoint of this interval by $2\varepsilon\Lambda$.

of Figure 1. We define for $r \in [0,1]$

$$g_{t_0+r}(x) = \begin{cases} G_0^\varepsilon(x)(\text{see}(7)) & \text{for } x \in N \setminus F(K \times [0, c_K)), \\ \overline{\mathbb{H}}^\varepsilon_{4\,\varepsilon\,\Lambda(1-r)\chi_0(\Pi_K(x))}(-\operatorname{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}(x)) & \text{for } x \in F(K \times (0, c_K)) \cup D, \\ 1 & \text{for } x \in F((K \setminus \overline{D}) \times \{0\}). \end{cases} \tag{31}$$

Note that $g_{t_0+r} = g_{t_0}$ when $r = 0$ (justifying the notation). Moreover, $g_{t_0+r}(x) = g_{t_0}(x)$ for $r \in [0,1]$ and $x \in N \setminus F(\operatorname{supp}(\chi) \times [0, c_K))$. In other words, we are only making changes to the values of g_{t_0} in the set $F(\widetilde{D} \times [0, c_K))$ (equivalently, introducing Fermi coordinates centred at D, the set $D \times (-c_K, c_K)$).

The fact that $g_{t_0+r} \in W^{1,\infty}(N)$ for every $r \in [0,1]$ follows by repeating the arguments used for g_t, where the only part that has to be altered is the local expression of g_{t_0+r} around points of D. Using Fermi coordinates (y, a) with $y \in D$, $a \in (-c_K, c_K)$ we get $g_{t_0+r}(F(y,a)) = \overline{\mathbb{H}}^\varepsilon_{4\,\varepsilon\,\Lambda(1-r)\chi(y)}(-|a| + 2\,\varepsilon\,\Lambda)$, which is Lipschitz. Notice that this is the domain in N where we are "closing the hole": when $r = 1$ the expression just obtained becomes $g_{t_0+1}(F(y,a)) = \overline{\mathbb{H}}^\varepsilon_0(-|a| + 2\,\varepsilon\,\Lambda) = \Psi_0(a)$ and so

$$g_{t_0+1}(x) = \begin{cases} G_0^\varepsilon(x)(\text{see}(7)) & \text{for } x \in N \setminus F(K \times [0, c_K)), \\ \overline{\mathbb{H}}^\varepsilon(-\operatorname{dist}_{K_{2\,\varepsilon\,\Lambda,t,\widetilde{\phi}}}(x)) & \text{for } x \in F(K \times (0, c_K)), \\ 1 & \text{for } x \in F(K \times \{0\}). \end{cases} \tag{32}$$

Note also that $r \in [0,1] \to g_{t_0+r} \in W^{1,2}(N)$ is a continuous path (with a proof as the ones used earlier for g_t and f_r).

Energy of g_{t_0+r}. We use the coarea formula as we did to reach (30). We get

$$\begin{aligned} &\int_{F(K\times[0,c_K))} \varepsilon \frac{|\nabla g_{t_0+r}|^2}{2} + \frac{W(g_{t_0+r})}{\varepsilon} \\ &\leq \int_{K_{2\,\varepsilon\,\Lambda,t_0,\widetilde{\phi}}} \left(\int_{-2\,\varepsilon\,\Lambda}^{2\,\varepsilon\,\Lambda} \frac{1}{|J\Pi_{2\,\varepsilon\,\Lambda,t_0}|} \left(\frac{\varepsilon}{2} \left|\overline{\mathbb{H}}^{\varepsilon\prime}(a)\right|^2 + \frac{W(\overline{\mathbb{H}}^\varepsilon(a))}{\varepsilon} \right) da \right) dy \underbrace{\leq}_{(12)} \\ &\leq (1 + 2\,\varepsilon\,\Lambda C_{K,t_0,c_0,\widetilde{\phi}}) \mathcal{H}^n(K_{2\,\varepsilon\,\Lambda,t_0,\widetilde{\phi}}) \left(\int_{-2\,\varepsilon\,\Lambda}^{2\,\varepsilon\,\Lambda} \frac{\varepsilon}{2} \left|\overline{\mathbb{H}}^{\varepsilon\prime}\right|^2 + \frac{W(\overline{\mathbb{H}}^\varepsilon)}{\varepsilon} \right) \\ &= (1 + 2\,\varepsilon\,\Lambda C_{K,t_0,c_0,\widetilde{\phi}}) \mathcal{H}^n(K_{2\,\varepsilon\,\Lambda,t_0,\widetilde{\phi}})(2\sigma + O(\varepsilon^2)). \end{aligned} \tag{33}$$

Therefore for some suitably small choice of $\varepsilon_2 \leq \varepsilon_1$ for all $\varepsilon < \varepsilon_2$ the following holds:

$$\frac{1}{2\sigma} \int_{F(K\times[0,c_K))} \varepsilon \frac{|\nabla g_{t_0+r}|^2}{2} + \frac{W(g_{t_0+r})}{\varepsilon} \leq \mathcal{H}^n(K_{2\,\varepsilon\,\Lambda,t_0,\widetilde{\phi}}) + O(\varepsilon\,|\log \varepsilon|).$$

Note that $K_{2\,\varepsilon\,\Lambda,t_0,\widetilde{\phi}}$ is the image of K via the immersion in (ii) of Lemma 6.4 when $c = 2\,\varepsilon\,\Lambda$. Using Lemma 6.4 in the last estimate and putting it together with (28) we finally obtain that for some suitably small choice of $\varepsilon_2 \leq \varepsilon_1$ for all $\varepsilon < \varepsilon_2$ the following estimate holds for all $r \in [0,1]$:

$$\mathcal{E}_\varepsilon(g_{t_0+r}) \leq 2\mathcal{H}^n(M) - \tau + O(\varepsilon\,|\log\varepsilon|). \tag{34}$$

7.5 | Connect to +1

To conclude the construction of our path, we will connect g_{t_0+1} to the constant $+1$ by means of a (negative) gradient flow. To this end, we will produce a suitable barrier m, constructed from G_0^ε. First we check that

$$G_0^\varepsilon \leq g_{t_0+1} \quad \text{on } N.$$

To see this, recall that $G_0^\varepsilon = g_{t_0+1}$ on $N \setminus F(K\times[0,c_K))$, so we only need to compare the two functions on $F(K\times[0,c_K))$. On this domain we use coordinates $(q,s)\in K\times[0,c_K)$. Use the following temporary notation: $H(x) = \overline{\mathbb{H}}^\varepsilon(-x)$, $T = \{(q,s) : q\in K, s = 2\,\varepsilon\,\Lambda + t_0\widetilde{\phi}(q)\}$, and $d(q,s) = \text{dist}_{K_{2\,\varepsilon\,\Lambda, t_0,\widetilde{\phi}}}(F(q,s))$. Equivalently, the latter signed distance is $\text{sgn}_{(q,s)}\,\text{dist}((q,s),T)$, where dist is the Riemannian distance (induced from N) and $\text{sgn}_{(q,s)} = -1$ on

$$\{(q,s) : q\in K, 0 < s < 2\,\varepsilon\,\Lambda + t_0\widetilde{\phi}(q)\}$$

and $\text{sgn}_{(q,s)} = +1$ on $\{(q,s) : q\in K, 2\,\varepsilon\,\Lambda + t_0\widetilde{\phi}(q) \leq s < c_K\}$. Then $G_0^\varepsilon(q,s) = H(s - 2\,\varepsilon\,\Lambda)$ and $g_{t_0+1}(q,s) = H(d(q,s))$. If $\text{sgn}_{(q,s)} = -1$, then the Riemannian distance to T is $\geq 2\,\varepsilon\,\Lambda - s$, because T lies above $\{s = 2\,\varepsilon\,\Lambda\}$. Similarly, if $\text{sgn}_{(q,s)} = +1$ then the Riemannian distance to T is $\leq s - 2\,\varepsilon\,\Lambda$. Therefore in either case we have $d(q,s) \leq s - 2\,\varepsilon\,\Lambda$. This implies (since H is decreasing) that $G_0^\varepsilon(q,s) \leq g_{t_0+1}(q,s)$.

We are going to work with the "modified" Allen–Cahn energy

$$\mathcal{F}_{\varepsilon,\mu_\varepsilon}(u) = \mathcal{E}_\varepsilon(u) - \frac{\mu_\varepsilon}{2\sigma}\int_N u\; d\mathcal{H}^{n+1},$$

where $\mu_\varepsilon > 0$ tends to 0 as $\varepsilon \to 0$. The role of μ_ε is that of a forcing term to ensure that the flow "moves in the desired direction" and is moreover "mean-convex". There is flexibility on the choice of μ_ε; we fix the following (note that in Section 4 we only required $\mu_\varepsilon > |O(\varepsilon^2)|$ in order to obtain (4)):

$$\mu_\varepsilon = \varepsilon\,|\log\varepsilon|. \tag{35}$$

We are now ready to construct the barrier.

Lemma 7.4. *For all sufficiently small ε there exists a smooth function $m : N \to \mathbb{R}$ ($m = m^\varepsilon$) such that $m < g_{t_0+1}$ and $-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m) = \varepsilon\,\Delta m - \frac{W'(m)}{\varepsilon} + \mu_\varepsilon > 0$.*

Proof. In Section 4 we obtained that, for all sufficiently small ε,

$$-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(G_0^\varepsilon) \geq \frac{\mu_\varepsilon}{2}\mathcal{H}^{n+1}$$

for $\mu_\varepsilon > 0$ as in (35). Recall that this inequality means that (the positive Radon measure) $-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(G_0^\varepsilon)$ minus $\frac{\mu_\varepsilon}{2}\mathcal{H}^{n+1}$ is a positive measure.

For $\rho > 0$ consider the function $G_0^\varepsilon - \rho$. Then

$$\Delta\left(G_0^\varepsilon - \rho\right) = \Delta G_0^\varepsilon$$

and $W'(G_0^\varepsilon - \rho)$ converges uniformly on N to $W'(G_0^\varepsilon)$ as $\rho \to 0$. Therefore we can find a sufficiently small $\rho_0 \in (0,1)$ (depending on ε; in fact, we may choose $\rho_0 \approx \varepsilon^2$) such that for all sufficiently small ε we have

$$-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}\left(G_0^\varepsilon - \rho_0\right) \geq \frac{\mu_\varepsilon}{3}\mathcal{H}^{n+1}. \tag{36}$$

Let C_N be the constant in Lemma A.3. We are going to work with ε sufficiently small to ensure (in addition to the previous conditions identified so far in this proof) that $2\,\varepsilon\,C_N < \mu_\varepsilon/20$. From now we work at fixed ε (satisfying the smallness conditions just imposed).

Let η_δ be the mollifiers defined in Appendix A for $\delta < \delta_0$, where $\delta_0 > 0$ depends only on the geometry of N. Then the (smooth) function $-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(G_0^\varepsilon - \rho_0) \star \eta_\delta$ defined in (A6) is positive for all δ, more precisely for all sufficiently small δ (one needs $\frac{1}{12} > |O(\delta^2)|$, where $O(\delta^2)$ appears in (A1))

$$\left(-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}\left(G_0^\varepsilon - \rho_0\right)\right) \star \eta_\delta \geq \frac{\mu_\varepsilon}{4}. \tag{37}$$

This follows from (36) and (A1), (A6). We now mollify $(G_0^\varepsilon - \rho_0)$ as in (A2). We have $|G_0^\varepsilon - \rho_0| < 2$, since $|G_0^\varepsilon| \leq 1$. From Lemma A.1, part (i), we obtain that the functions $(G_0^\varepsilon - \rho_0) \star \eta_\delta$ converge uniformly on N to $(G_0^\varepsilon - \rho_0)$ as $\delta \to 0$. Therefore (for δ sufficiently small $-2 < (G_0^\varepsilon - \rho_0) \star \eta_\delta < 2$ since the same bound holds for $G_0^\varepsilon - \rho_0$),

$$\begin{aligned} &\left\| W'\left(\left(G_0^\varepsilon - \rho_0\right) \star \eta_\delta\right) - W'(G_0^\varepsilon - \rho_0)\right\|_{C^0(N)} \\ &\quad \leq \|W''\|_{C^0([-2,2])} \|(G_0^\varepsilon - \rho_0) \star \eta_\delta - \left(G_0^\varepsilon - \rho_0\right)\|_{C^0(N)} \to 0 \end{aligned} \tag{38}$$

as $\delta \to 0$. The function $W'(G_0^\varepsilon - \rho_0)$ belongs to $W^{1,\infty}(N)$; therefore by Lemma A.1, part (i), we get $\|W'(G_0^\varepsilon - \rho_0) \star \eta_\delta - W'(G_0^\varepsilon - \rho_0)\|_{C^0(N)} \to 0$. By the triangle inequality we therefore have

$$\left\| W'((G_0^\varepsilon - \rho_0) \star \eta_\delta) - W'(G_0^\varepsilon - \rho_0) \star \eta_\delta \right\|_{C^0(N)} \to 0 \tag{39}$$

as $\delta \to 0$. By Lemma A.3 there exists C_N (depending only on the geometry of N) such that for all $\delta < \delta_0$ we have

$$\|\Delta((G_0^\varepsilon - \rho_0) \star \eta_\delta) - \Delta(G_0^\varepsilon - \rho_0) \star \eta_\delta\|_{L^\infty(N)} \leq C_N \|G_0^\varepsilon - \rho_0\|_{L^\infty(N)} \leq 2C_N.$$

Therefore the modulus of the difference of the two (smooth) functions

$$\begin{aligned} \varepsilon\,\Delta((G_0^\varepsilon - \rho_0) \star \eta_\delta) - \tfrac{W'((G_0^\varepsilon - \rho_0) \star \eta_\delta)}{\varepsilon} + \mu_\varepsilon \\ \text{and} \quad \left(-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(G_0^\varepsilon - \rho_0)\right) \star \eta_\delta \end{aligned} \tag{40}$$

is at most $2\,\varepsilon\,C_N + O_\delta(1)$, where the infinitesimal of δ is given by the norm in (39) plus $O(\delta^2)\mu_\varepsilon$. Recall (37) and the smallness condition imposed on ε. Then for sufficiently small δ, writing

$m = (G_0^\varepsilon - \rho_0) \star \eta_\delta$, we have

$$\varepsilon\,\Delta m - \frac{W'(m)}{\varepsilon} + \mu_\varepsilon \geq \frac{\mu_\varepsilon}{5}. \tag{41}$$

Finally, note that for sufficiently small δ we also have $m < g_{t_0+1}$, since $G_0^\varepsilon - \rho_0 < g_{t_0+1}$ and $(G_0^\varepsilon - \rho_0) \star \eta_\delta$ converges uniformly to $G_0^\varepsilon - \rho_0$ as $\delta \to 0$ (Lemma A.1). □

Remark 7.5 (Choice of ε_2, again). We will assume that Lemma 7.4 is valid for all $\varepsilon < \varepsilon_2$, where once again we change the choice of ε_2 if necessary.

Flow from m. We consider now the negative gradient flow of $(2\sigma)\mathcal{F}_{\varepsilon,\mu_\varepsilon}$, with initial condition given by the smooth function m, that is, the solution m_t to the PDE

$$\begin{cases} \varepsilon \frac{\partial}{\partial t} m_t = \varepsilon\,\Delta m_t - \frac{W'(m_t)}{\varepsilon} + \mu_\varepsilon, \\ m_0 = m, \end{cases} \tag{42}$$

where Δ is the Laplace–Beltrami operator on N. This semilinear parabolic problem has a solution for $t \in [0, \infty)$ and $m_t \in C^\infty(N)$ for all $t > 0$, as we will now sketch.

Short-time existence and uniqueness for a weak solution in $W^{1,2}(N)$ are valid by standard semilinear parabolic theory (rewrite the problem as an integral equation, then use a fixed point theorem). To see why we get global existence in our case, integrate (42) on any interval $[0, T]$ on which the weak solution is defined: we get

$$\begin{aligned} &\varepsilon \int_0^T \left(\int_N \left| \frac{\partial}{\partial t} m_t \right|^2 \right) dt + \frac{\varepsilon}{2} \int_N |\nabla m_T|^2 \\ &= \frac{\varepsilon}{2} \int_N |\nabla m_0|^2 - \frac{1}{\varepsilon} \int_N (W(m_T) - \varepsilon\,\mu_\varepsilon m_T) \\ &\quad + \frac{1}{\varepsilon} \int_N (W(m_0) - \varepsilon\,\mu_\varepsilon m_0). \end{aligned} \tag{43}$$

With our choice of W that is quadratic on $(\pm 2, \pm\infty)$ we can ensure that $\frac{W(u)}{\varepsilon} - \mu_\varepsilon u$ is bounded below. Then (43) gives a priori bounds $\int_N |\nabla m_t|^2 \leq C_{m_0,\varepsilon,W}$ independently of $t \in [0, T]$. Again from (43), moving the term

$$\frac{1}{\varepsilon} \int_N (W(m_T) - \varepsilon\,\mu_\varepsilon m_T)$$

to the left-hand side and recalling $|u|^2 \leq C_{W,\varepsilon} \max\{2, \frac{W(u)}{\varepsilon} - \mu_\varepsilon u\}$, we also get an a priori L^2-bound on m_t. In conclusion,

$$\|m_t\|_{W^{1,2}(N)} \leq C, \qquad \int_0^T \left(\int_N \left| \frac{\partial}{\partial t} m_t \right|^2 \right) \leq C, \tag{44}$$

with C independent of t. This first bound in (44) provides the assumption under which short-time existence can be iterated to lead global existence for a weak solution to (42) in $W^{1,2}(N)$.

Writing the PDE in the form $\frac{\partial}{\partial t}m_t - \Delta m_t = -\frac{1}{\varepsilon^2}W'(m_t) + \frac{\mu_\varepsilon}{\varepsilon}$, we treat the right-hand side as the nonhomogeneous term f_t of a linear parabolic PDE. Thanks to the quadratic growth of W, there exists a constant C_W (depending only on W) such that $\|f_t\|_{W^{1,2}(N)} \le C_W\|m_t\|_{W^{1,2}(N)}$ and $\|\frac{\partial f_t}{\partial t}\|_{L^2(N)} \le C_W\|\frac{\partial m_t}{\partial t}\|_{L^2(N)}$, which are, respectively, L^∞ and L^2 in t by (44). From parabolic regularity $m_t \in W^{2,2}(N)$ and $\frac{\partial m_t}{\partial t} \in W^{1,2}(N)$ for all t, with $\|m_t\|_{W^{2,2}(N)}$ and $\|\frac{\partial m_t}{\partial t}\|_{W^{1,2}(N)}$ bounded uniformly in time (see, e.g., [7, sec. 7.2.3, theorem 6]). Bootstrapping gives smoothness of m_t.

Lemma 7.6 (Mean convexity of m_t). *The positivity condition* $-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m_t) = \varepsilon\,\Delta m_t - \frac{W'(m_t)}{\varepsilon} + \mu_\varepsilon > 0$ *holds for all* $t \ge 0$.

Proof. For notational convenience, we write for this paragraph $F_t = \varepsilon\,\Delta m_t - \frac{W'(m_t)}{\varepsilon} + \mu_\varepsilon$ (right-hand side of the first line in (42)). By the previous discussion, F_t is smooth on N for all $t \in [0,\infty)$. Differentiating $F_t = \varepsilon\,\Delta m_t - \frac{W'(m_t)}{\varepsilon} + \mu_\varepsilon$ (and using $\varepsilon\,\partial_t m_t = F_t$) we get the evolution of F_t, given by $\partial_t F_t = \Delta F_t - \frac{W''(m_t)}{\varepsilon^2}F_t$. So F_t solves $\partial_t\gamma = \Delta\gamma - \frac{W''(m_t)}{\varepsilon^2}\gamma$, and the constant $\gamma = 0$ is also a solution to the same PDE. The condition $F_t > 0$ is therefore preserved by the maximum principle, since it holds at $t = 0$ by Lemma 7.4. □

Lemma 7.6 implies in particular that $m_t : N \to \mathbb{R}$ is increasing in t since $\partial_t m_t = -\frac{2\sigma}{\varepsilon}\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m_t) > 0$, therefore $\lim_{t\to\infty} m_t = m_\infty$ is well-defined pointwise on N. The $W^{1,2}(N)$-norm of m_t is bounded uniformly in t by (44); therefore $m_t \to m_\infty$ in $W^{1,2}$-weak. Moreover, $\|W'(m_t)\|_{W^{1,2}(N)}$ is also uniformly bounded in t, since

$$|\nabla(W'(m_t))| = |W''(m_t)||\nabla m_t| \le \|W''\|_{C^0([-2,2])}|\nabla m_t|$$

(one can check that $-2 \le m_t \le 2$ for all t by the maximum principle). Therefore $W'(m_t) \to W'(m_\infty)$ in $W^{1,2}$-weak. By the second bound in (44) we have L^1-summability in time on $t \in (0,\infty)$ for $\|\frac{\partial}{\partial t}m_t\|_{L^2(N)}$ and therefore there exists $t_j \to \infty$ such that the function $\frac{\partial}{\partial t}m_t : N \to \mathbb{R}$ has $L^2(N)$-norm that tends to 0 along the sequence t_j. These facts imply that the weak formulation of the PDE in (42) passes to the limit as $t_j \to \infty$ and gives that m_∞ solves $-\mathcal{F}'_{\varepsilon,\mu_\varepsilon} = 0$ in the weak sense. Standard elliptic theory (or passing parabolic estimates for m_t to the $t \to \infty$ limit) then show that $m_\infty \in C^\infty$ solves $-\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m_\infty) = 0$ in the strong sense.

Lemma 7.7 (Stability of m_∞). *The limit m_∞ of the flow m_t (as $t \to \infty$) is a stable solution of* $\mathcal{F}'_{\varepsilon,\mu_\varepsilon} = 0$.

Proof. This is a consequence of the mean-convexity of m_t (Lemma 7.6) and of the maximum principle. We give the explicit argument. Recall from the previous discussion that m_∞ is stationary, that is, $\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m_\infty) = 0$. Also recall that the second variation at $u : N \to \mathbb{R}$ of the functional $(2\sigma)\mathcal{F}_{\varepsilon,\mu}$ (for a constant μ) on the test function ϕ is given by the quadratic form $Q(\phi,\phi) = \int_N \varepsilon\,|\nabla\phi|^2 + \frac{W''(u)}{\varepsilon}\phi^2$ (the term involving μ disappears because it is linear) and the Jacobi operator is given by $-\varepsilon\,\Delta\phi + \frac{W''(u)}{\varepsilon}\phi$.

Let ρ_1 be its first eigenfunction; then ρ_1 is (strictly) positive and smooth on N. Consider, for $s \in (-\delta, \delta)$ (for some small positive δ), the functions $m_\infty - s\rho_1$. Then their first variation satisfies

$$\frac{\partial}{\partial s}\left(-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m_\infty - s\rho_1)\right) = -\varepsilon\,\Delta\rho_1 + \frac{W''(m_\infty - s\rho_1)}{\varepsilon}\rho_1.$$

If m_∞ were unstable, then the first eigenfunction would satisfy

$$-\varepsilon\,\Delta\rho_1 + \frac{W''(m_\infty)}{\varepsilon}\rho_1 = \lambda_1\rho_1$$

for some $\lambda_1 < 0$ and therefore

$$\left.\frac{\partial}{\partial s}\right|_{s=0}\left(-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m_\infty - s\rho_1)\right) = \lambda_1\rho_1 < 0$$

on N. Then we could choose $s_0 > 0$ sufficiently small so that

$$\begin{aligned} &-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m_\infty - s\rho_1) \\ &\quad = \varepsilon\,\Delta(m_\infty - s\rho_1) - \frac{W'((m_\infty - s\rho_1))}{\varepsilon} + \mu_\varepsilon < 0 \end{aligned} \tag{45}$$

on N for $s \in [0, s_0]$. Note that $m_\infty - s\rho_1$ is smooth on N.

Since $-(2\sigma)\mathcal{F}'_{\varepsilon,\mu_\varepsilon}(m_t) > 0$, at any $t \in [0,\infty)$ we have $m_t > m_0$, in particular $m_\infty > m_0$. Choose s sufficiently small so that $s < s_0$ and $m_\infty - s\rho_1 > m_0$. Let $\tau > 0$ be the first time for which m_τ has a point x such that $m_\tau(x) = (m_\infty - s\rho_1)(x)$. Then $m_\infty - s\rho_1 - m_\tau$ is a smooth nonnegative function on N with a minimum at x, so $\Delta(m_\infty - s\rho_1)(x) \geq \Delta m_\tau(x)$. Moreover, we have $W'(m_\infty - s\rho_1) = W'(m_\tau)$ at x. Recalling that $\varepsilon\,\Delta m_\tau - \frac{W'(m_\tau)}{\varepsilon} + \mu_\varepsilon > 0$ on N (preservation of mean convexity) we get $\varepsilon\,\Delta(m_\infty - s\rho_1)(x) - \frac{W'((m_\infty - s\rho_1)(x))}{\varepsilon} + \mu_\varepsilon > 0$, contradicting (45). □

Proposition 7.8. *If* $\mathrm{Ric}_N > 0$ *then any stable solution to* $\mathcal{F}'_{\varepsilon,\mu} = 0$ *on* N *must be a constant (here* μ *is any given constant.)*

Proof. Let u be a stable solution to $\mathcal{F}'_{\varepsilon,\mu}(u) = 0$. We test the stability inequality $Q(\cdot,\cdot) \geq 0$ on a test function of the form $|\nabla u|\phi$ for $\phi \in C^2(N)$. We get (this expression of Q follows using Bochner's identity; see [5, 6, 36])

$$\int_{N\setminus\{|\nabla u|=0\}} \left(|A_\varepsilon|^2 + \mathrm{Ric}_N\left(\frac{\nabla u}{|\nabla u|}, \frac{\nabla u}{|\nabla u|}\right)\right)\varepsilon\,|\nabla u|^2\phi^2 \leq \int_N \varepsilon\,|\nabla u|^2|\nabla\phi|^2,$$

where $|A_\varepsilon|^2 = |D^2 u|^2 - |\nabla|\nabla u||^2 \geq 0$. We plug in $\phi = 1$ so the positiveness of Ric_N gives $\nabla u \equiv 0$. □

Lemma 7.7 and Proposition 7.8 give that m_∞ is a constant. There exist exactly two stable constant solutions of $\mathcal{F}'_{\varepsilon,\mu} = 0$. Indeed, any constant k satisfying $\mathcal{F}'_{\varepsilon,\mu} = 0$ must satisfy $W'(k) = \varepsilon\,\mu$ (and therefore $W(k) \approx c_W^2\,\varepsilon^2\,\mu^2$ for some c_W depending on W), so we obtain three constants, one slightly larger than -1, one slightly larger than $+1$, one slightly smaller than 0, when ε is sufficiently small. It is easily verified that the constant close to 0 is unstable, while the other two are stable. In our case, since $m_\infty > m_0$ and $m_0 > 1/2$ on an open neighbourhood of $\overline{M}$, we conclude that m_∞ is the constant slightly larger than $+1$, which we will denote by k_{μ_ε}:

$$m_\infty \equiv k_{\mu_\varepsilon}. \tag{46}$$

Flow from g_{t_0+1}. We now want to consider the negative $(2\sigma)\mathcal{F}_{\varepsilon,\mu_\varepsilon}$-gradient flow $\{h_t\}$ starting at g_{t_0+1}. We first make the initial datum smooth, by considering mollifiers η_δ for $\delta \in (0,\overline{\delta}]$ as in Appendix A and $\overline{\delta}$ sufficiently small to preserve the strict inequality with $m = m_0$, that is, to ensure $g_{t_0+1} \star \eta_\delta > m_0$ for $\delta \in (0,\overline{\delta}]$. The family

$$\delta \in (0,\overline{\delta}] \to g_{t_0+1} \star \eta_\delta \in W^{1,2}(N) \tag{47}$$

is continuous in δ and extends by continuity at $\delta = 0$ with value g_{t_0+1} (see Remark A.2). Continuity is also valid for $\delta \in (0,\overline{\delta}] \to g_{t_0+1} \star \eta_\delta \in C^0(N)$. As a consequence, $\mathcal{E}_\varepsilon(g_{t_0+1} \star \eta_\delta)$ varies continuously with δ and therefore, upon choosing $\overline{\delta}$ possibly smaller, we also have, in addition to (47) and to $g_{t_0+1} \star \eta_{\overline{\delta}} > m_0$, that the following holds for all $\delta \in (0,\overline{\delta}]$,

$$\mathcal{E}_\varepsilon(g_{t_0+1} \star \eta_\delta) \le \mathcal{E}_\varepsilon(g_{t_0+1}) + \frac{1}{4}\tau. \tag{48}$$

We now let $h_0 = g_{t_0+1} \star \eta_{\overline{\delta}}$ be the initial condition for the negative $(2\sigma)\mathcal{F}_{\varepsilon,\mu_\varepsilon}$-gradient flow:

$$\begin{cases} \varepsilon \dfrac{\partial}{\partial t} h_t = \varepsilon \Delta h_t - \dfrac{W'(h_t)}{\varepsilon} + \mu_\varepsilon, \\ h_0 = g_{t_0+1} \star \eta_{\overline{\delta}}. \end{cases} \tag{49}$$

By the maximum principle, since $m_0 < h_0$, the two flows (42) and (49) preserve $m_t < h_t$ for all t.[13] Since $g_{t_0+1} \le 1$ by construction, we also have $h_0 < k_{\mu_\varepsilon}$; therefore $h_t < k_{\mu_\varepsilon}$ for all $t > 0$ by the maximum principle. On the other hand, we saw that $m_t \to k_{\mu_\varepsilon}$ as $t \to \infty$; therefore (with smooth convergence, in particular we have continuity in t for $t \in [0,\infty] \to h_t \in W^{1,2}(N)$)

$$h_t \to k_{\mu_\varepsilon} \quad \text{as } t \to \infty. \tag{50}$$

Evaluation of $\mathcal{E}_\varepsilon$ *on the path* h_t. Let us estimate the value of $\mathcal{E}_\varepsilon$ along this path. For this, note that $\mathcal{F}_{\varepsilon,\mu_\varepsilon}$ is decreasing along the flow $\{h_t\}$; therefore $\mathcal{E}_\varepsilon(h_t) \le \mathcal{E}_\varepsilon(h) + 2\frac{\mu_\varepsilon}{2\sigma}\mathcal{H}^{n+1}(N)$ for all t (where we used $h_t < 2$ for all t). This implies that $\mathcal{E}_\varepsilon$ is bounded above independently of ε; more precisely, recalling that $\mathcal{E}_\varepsilon(h_0) \le 2\mathcal{H}^n(M) - \tau + O(\varepsilon\,|\log\varepsilon|)$, we can absorbe $\frac{\mu_\varepsilon}{\sigma}\mathcal{H}^{n+1}(N)$ in the error $O(\varepsilon\,|\log\varepsilon|)$ for ε sufficiently small. In other words, we obtain, for $\varepsilon_2 \le \varepsilon_1$ sufficiently small, the upper bound

$$\mathcal{E}_\varepsilon(h_t) \le 2\mathcal{H}^n(M) - \frac{3}{4}\tau + O(\varepsilon\,|\log\varepsilon|) \tag{51}$$

for all t and for all $\varepsilon < \varepsilon_2$.

To complete the path, we connect $h_\infty = k_{\mu_\varepsilon}$ to $+1$ (through constant functions):

$$k_t = (1-t)k_{\mu_\varepsilon} + t \tag{52}$$

[13] We have smoothed the initial data in order to use basic linear parabolic theory to obtain smoothness at all times and thus use the classical maximum principle. The other option is to use g_{t_0+1} as initial condition and prove that it becomes smooth after a short time.

for $t \in [0,1]$. The energy $\mathcal{E}_\varepsilon(k_t)$ is decreasing in $t \in [0,1]$, since W is an increasing function on $[1, k_{\mu_\varepsilon}]$. Therefore the same upper bound that we had in (51) holds:

$$\mathcal{E}_\varepsilon(k_t) \leq 2\mathcal{H}^n(M) - \frac{3}{4}\tau + O(\varepsilon\,|\log\varepsilon|) \tag{53}$$

for all t and for all $\varepsilon < \varepsilon_2$.

8 | CONCLUSION OF THE PROOF OF THEOREMS 1.3, 1.1, AND 1.8

In the previous sections we exhibited (given M as in Theorem 1.3, which also fixed B and τ by Remark 5.4 and Lemma 6.4) for all sufficiently small ε (namely $\varepsilon < \varepsilon_2$) the following six continuous paths in $W^{1,2}(N)$: (24) reversed, (26), (31), (47), (49), and (52). In the order just given, the endpoint of each partial path matches the starting point of the next one, therefore their composition in the same order provides a continuous path in $W^{1,2}(N)$ for all $\varepsilon < \varepsilon_2$, that starts at the constant -1 and ends at the constant $+1$ and such that

$$\mathcal{E}_\varepsilon \text{ along this path is} \leq 2\mathcal{H}^n(M) - \min\left\{\frac{3}{4}\tau, \frac{3\mathcal{H}^n(B)}{2}\right\} + O(\varepsilon\,|\log\varepsilon|),$$

thanks to (25, 30, 34, 48, 51, 53). Choosing ε_3 sufficiently small to ensure that $\varepsilon < \varepsilon_3 \Rightarrow |O(\varepsilon\,|\log\varepsilon|)| \leq \min\{\frac{\tau}{4}, \mathcal{H}^n(B)\}$ the above bound gives, for all $\varepsilon < \varepsilon_3$, that the maximum of $\mathcal{E}_\varepsilon$ on the path is at most $2\mathcal{H}^n(M) - \min\{\frac{\tau}{2}, \frac{\mathcal{H}^n(B)}{2}\}$.

The path is in the admissible class for the minmax construction in [14]; therefore the maximum on this specific path controls from above the minmax value c_ε achieved by the index-1 solution u_ε obtained from [14] (for all $\varepsilon < \varepsilon_3$). Summarising, for every $M \subset N$ as in Theorem 1.3 there exist $\varepsilon_3 > 0$, $\tau > 0$, and $B \subset M$ (nonempty) such that for all $\varepsilon < \varepsilon_3$

$$c_\varepsilon = \mathcal{E}_\varepsilon(u_\varepsilon) \leq 2\mathcal{H}^n(M) - \min\left\{\frac{\tau}{2}, \frac{\mathcal{H}^n(B)}{2}\right\}. \tag{54}$$

This concludes the proof of the strict inequality in Theorem 1.3.

For Theorem 1.1 it suffices to observe that the integral varifold V produced in [14] is (thanks to [37, 40]) such that each connected component of reg_V (the smoothly embedded part of spt $\|V\|$) has the properties needed so that it can be used in place of M in Theorem 1.3 or in (54) above; moreover, the mass $\|V\|(N)$ of V is $\lim_{\varepsilon_i \to 0} c_{\varepsilon_i}$ (see Section 2.1). Letting M be any connected component of reg_V and denoting by $\theta \in \mathbb{N}$ its (constant) multiplicity, using (54) we get $\theta\mathcal{H}^n(M) \leq \|V\|(N) < 2\mathcal{H}^n(M)$. This implies $\theta = 1$ and the multiplicity assertion in Theorem 1.1 is proved.

The fact that the minimal hypersurface is two-sided then follows immediately, since under multiplicity-1 convergence (and by the lower energy bounds in [14]) we have that $u_{\varepsilon_i} \to u_\infty$ in $BV(N)$, where u_∞ is a nonconstant function that takes values in $\{-1, +1\}$, and moreover, V is the multiplicity-1 varifold associated to the reduced boundary of the set (of finite perimeter) $\{u_\infty = +1\}$ (there is no "hidden boundary" in the limit). We therefore have a global normal on reg_V (the interior- or the exterior-pointing normal for $\partial\{u_\infty = +1\}$). Theorem 1.8 is therefore proved.

Remark 8.1. Note that reg_V has to be connected, since each connected[14] component of it is unstable (because it is two-sided and $\mathrm{Ric}_N > 0$), and therefore the Morse index of reg_V is at least the number of its connected components. On the other hand, by multiplicity-1 convergence (or by [9, 15]) the Morse index of reg_V is ≤ 1. An alternative argument for the connectedness, that does not rely on two-sidedness, can be given by means of the maximum principle for stationary varifolds [17, 40] and the Frankel property[15] for $\mathrm{Ric}_N > 0$ (using the regularity results [37, 40]).

Remark 8.2. In view of discussing Remark 1.6, we collect the three instances in which the curvature assumption $\mathrm{Ric}_N > 0$ was used in the proof of Theorems 1.1, 1.3, and 1.8. The first was in obtaining the sign condition $\Delta d_{\overline{M}} \leq 0$ in Lemma 3.6 and the area bounds in Lemma 4.1. The second, in Lemma 5.1, was to conclude that $\iota : \widetilde{M} \to N$ is unstable as a minimal immersion. The third, in Proposition 7.8, was to conclude that every stable solution to $\mathcal{F}'_{\varepsilon,\mu} = 0$ on N (for μ constant) is a constant function.

The weaker assumptions stated in Remark 1.6 are easily seen to be sufficient for the proof. Lemma 3.6 only requires $\mathrm{Ric}_N \geq 0$. Similarly, in Lemma 4.1 it suffices to use $\mathrm{Ric}_N \geq 0$ to conclude $\mathcal{H}^n(\Gamma_t) \leq 2\mathcal{H}^n(M)$; this inequality is enough for the estimates that follow Lemma 4.1 and lead to (8).

Let us assume that $\{\mathrm{Ric}_N = 0\}$ has vanishing $\mathcal{H}^n$-measure. To carry out the proof of Lemma 5.1, in particular to obtain the negativity of $-\int_M \mathrm{Ric}_N(\nu,\nu)d\mathcal{H}^n$, it suffices to notice that the integrand is negative on a subset of M of full measure. For Proposition 7.8 the conclusion will be in a first instance that ∇u vanishes except possibly on $\{\mathrm{Ric}_N = 0\}$; the smoothness of any solution to $\mathcal{F}'_{\varepsilon,\mu} = 0$ and the fact that $\{\mathrm{Ric}_N = 0\}$ has empty interior then imply that ∇u vanishes identically.

Let us assume now that $\{\mathrm{Ric}_N = 0\} \subset \bigcup_{i=1}^{\infty} A_i$ as in Remark 1.6. Then for a stationary varifold ($2|M|$ in our case) the support cannot be contained in $\bigcup_{i=1}^{\infty} A_i$ (and therefore $M \setminus \{\mathrm{Ric}_N = 0\}$ has positive measure). This follows from the maximum principle [18], using the boundaries of A_i as barriers (e.g., flowing them by mean curvature until they touch the support of the varifold). Then we follow Lemma 5.1 and the negativity of the term $-\int_M \mathrm{Ric}_N(\nu,\nu)d\mathcal{H}^n$ follows from the previous observation. For Proposition 7.8 the conclusion will be in a first instance that ∇u vanishes except possibly on $\{\mathrm{Ric}_N = 0\}$. On an arbitrary connected component of $N \setminus \{\mathrm{Ric}_N = 0\}$ then, u has to be a constant k; this constant must also be a solution to $\mathcal{F}'_{\varepsilon,\mu} = 0$ on N. Then k and u are both solutions to $\mathcal{F}'_{\varepsilon,\mu} = 0$, and they coincide on a nonempty open set; taking the difference of the two PDEs. By unique continuation we obtain $u - k \equiv 0$, in particular u is constant.

ACKNOWLEDGMENTS

This work is partially supported by EPSRC under the grant EP/S005641/1. I am thankful to Neshan Wickramasekera for introducing me to the Allen–Cahn functional and its geometric impact. I would like to thank Felix Schulze for helpful conversations about parabolic PDEs and mean curvature flow. I am grateful to Otis Chodosh for a mini-course on geometric features of the Allen–Cahn equation, held at Princeton University in June 2019, and for related discussions. The insight that I gained at that time proved very valuable when I addressed the problem discussed in this work. These lectures took place while I was a member of the Institute for Advanced Study, Princeton: I gratefully acknowledge the excellent research environment and the support provided by

[14] We point out that connectedness of reg_V and of $\mathrm{spt}\,\|V\|$ are in fact the same thing by the varifold maximum principle.

[15] The proof of the Frankel property can be adapted because we have local stability for V and so the shortest geodesic between two connected components of $\mathrm{spt}\,\|V\|$ must have endpoints on the smooth parts (not on the singular set), by the same reasoning used in Lemma 3.1, see also [42, Theorem 2.10].

the Institute and by the National Science Foundation under Grant No. DMS-1638352. Further thanks to Kobe Marshall-Stevens and Myles Workman for comments on a preliminary version of the manuscript, and to Alessio Figalli for an interesting discussion related to some aspects of this work.

REFERENCES

1. G. Alberti, *On the structure of singular sets of convex functions*, Calc. Var. Partial Differential Equations **2** (1994), 17–27.
2. F. J. Almgren, Jr. *Theory of varifolds*, Mimeographed notes, Princeton University, 1965.
3. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variations and Free Discontinuity Problems*, Oxford Science Publications, Oxford Math. Monogr. The Clarendon Press, Oxford University Press, New York, 2000, xviii+434 pp.
4. C. Bellettini, *Generic existence of multiplicity-1 minmax minimal hypersurfaces via Allen–Cahn*, Calc. Var. Partial Differential Equations **61** (2022), no. 4, paper no. 149, 25 pp.
5. C. Bellettini and N. Wickramasekera, *The inhomogeneous Allen–Cahn equation and the existence of prescribed-mean-curvature hypersurfaces*, arXiv:2010.05847, 2020.
6. O. Chodosh and C. Mantoulidis, *Minimal surfaces and the Allen–Cahn equation on 3-manifolds: Index, multiplicity, and curvature estimates*, Ann. Math. **191** (2020), no. 1, 213–328.
7. L. C. Evans, *Partial Differential Equations*, 2nd ed., American Mathematical Society, Providence, RI, 2010.
8. L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, revised edition, CRC Press, Boca Raton, FL, 2015, xiv+299 pp.
9. P. Gaspar, *The second inner variation of energy and the morse index of limit interfaces*, J. Geom. Anal. **30** (2020), 69–85.
10. P. Gaspar and M. A. M. Guaraco, *The Weyl Law for the phase transition spectrum and the density of minimal hypersurfaces*, Geom. Funct. Anal. **29** (2019), no. 2, 382–410.
11. A. Gray, *Tubes*, 2nd ed., Progr. Math., 221 Birkhäuser Verlag, Basel, 2004, xiv+280 pp.
12. A. Gray and L. Vanhecke, *Riemannian geometry as determined by the volumes of small geodesic balls*, Acta Math. **142** (1979), 157–198.
13. M. Gromov, *Paul Levy's isoperimetric inequality*. Available at https://www.ihes.fr/~gromov/wp-content/uploads/2018/08/1133.pdf.
14. M. A. M. Guaraco, *Min-max for phase transitions and the existence of embedded minimal hypersurfaces*, J. Differential Geom. **108** (2018), no. 1, 91–133.
15. F. Hiesmayr, *Spectrum and index of two-sided Allen–Cahn minimal hypersurfaces*, Comm. Partial Differential Equations **43** (2018), no. 11, 1541–1565.
16. J. E. Hutchinson and Y. Tonegawa, *Convergence of phase interfaces in the van der Waals–Cahn–Hilliard theory*, Calc. Var. Partial Differential Equations **10** (2000), no. 1, 49–84.
17. T. Ilmanen, *Convergence of the Allen-Cahn equation to Brakke's motion by mean curvature*, J. Differential Geom. **38** (1993), no. 2, 417–461.
18. T. Ilmanen, *A strong maximum principle for singular minimal hypersurfaces*, Calc. Var. Partial Differential Equations **4** (1996), 443–467.
19. D. Ketover, F. C. Marques, and A. Neves, *The catenoid estimate and its geometric applications*, J. Differential Geom. **115** (2020), no. 1, 1–26.
20. R. V. Kohn and P. Sternberg, *Local minimisers and singular perturbations*, Proc. Roy. Soc. Edinburgh Sect. A **111** (1989), no. 1–2, 69–84.
21. C. Mantegazza and A. Mennucci, *Hamilton–Jacobi equations and distance functions on Riemannian manifolds*, Appl. Math. Opt. **47** (2002), 1–25.
22. F. C. Marques and A. Neves, *Min-max theory and the Willmore conjecture*, Ann. Math. (2) **179** (2014), no. 2, 683–782.
23. F. C. Marques and A. Neves, *Morse index of multiplicity one min-max minimal hypersurfaces*, Adv. Math. **378** (2021), paper no. 107527, 58 pp.
24. F. C. Marques and A. Neves, *Existence of infinitely many minimal hypersurfaces in positive Ricci curvature*, Invent. Math. **209** (2017), no. 2, 577–616.

25. L. Modica and S. Mortola, *Un esempio di Γ-convergenza*, Boll. Un. Mat. Ital. B (5) **14** (1977), no. 1, 285–299.
26. A. Pigati and T. Rivière, *A proof of the multiplicity one conjecture for min-max minimal surfaces in arbitrary codimension*, Duke Math. J. **169** (2020), no. 11, 2005–2044.
27. A. Pigati and D. Stern, *Minimal submanifolds from the abeilan Higgs model*, Invent. Math. **223** (2021), no. 3, 1027–1095.
28. J. Pitts, *Existence and regularity of minimal surfaces on Riemannian manifolds*, Princeton University press, 1977.
29. A. Ramírez-Luna, *Orientability of min-max hypersurfaces in manifolds of positive Ricci curvature*, arXiv:1907.12519, 2019.
30. T. Rivière, *A viscosity method in the min-max theory of minimal surfaces*, Publ. Math. Inst. Hautes Études Sci. **126** (2017), 177–246.
31. T. Sakai, *Riemannian Geometry*, American Mathematical Society, Providence, RI, 1996, xiv+358 pp.
32. R. Schoen, L. Simon, and S. T. Yau, *Curvature estimates for minimal hypersurfaces*, Acta Math. **134** (1975), no. 3-4, 275–288.
33. R. Schoen and L. Simon, *Regularity of stable minimal hypersurfaces*, Comm. Pure Appl. Math. **34** (1981), no. 6, 741–797.
34. L. Simon, *Lectures on geometric measure theory*, Proc. Centre Math. Anal., vol. 3, Canberra, 1984, pp. VII+272.
35. A. Song, *Existence of infinitely many minimal hypersurfaces in closed manifolds*, Ann. Math. **197** (2023) no. 3, 859–895.
36. Y. Tonegawa, *On stable critical points for a singular perturbation problem*, Comm. Anal. Geom. **13** (2005), no. 2, 439–459.
37. Y. Tonegawa and N. Wickramasekera, *Stable phase interfaces in the van der Waals–Cahn–Hilliard theory*, J. Reine Angew. Math. **668** (2012), 191–210.
38. K. Wang, *A new proof of Savin's theorem on Allen-Cahn equations*, J. Eur. Math. Soc. **19** (2017), no. 10, 2997–3051.
39. K. Wang and J. Wei, *Finite Morse index implies finite ends*, Comm. Pure Appl. Math. **72** (2019), no. 5, 1044–1119.
40. N. Wickramasekera, *A general regularity theory for stable codimension 1 integral varifolds*, Ann. Math. (2) **179** (2014), no. 3, 843–1007.
41. X. Zhou, *Min-max minimal hypersurface in* (M^{n+1}, g) *with* $\mathrm{Ric}_g > 0$ *and* $2 \le n \le 6$, J. Differential Geom. **100** (2015), no. 1, 129–160.
42. X. Zhou, *Minmax hypersurface in manifold of positive Ricci curvature*, J. Differential Geom. **105** (2017), no. 2, 291–343.
43. X. Zhou, *On the multiplicity one conjecture in min-max theory*, Ann. Math. (2) **192** (2020), no. 3, 767–820.

APPENDIX A: MOLLIFIERS

We explain in detail the mollification procedure used in Section 7.5. For this appendix, notation is reset. Let (N, g) be a closed Riemannian manifold of dimension $n+1$ and $f : N \to \mathbb{R}$ in $W^{1,\infty}(N)$. We are going to produce, for every $\delta > 0$ sufficiently small, a smooth function $f_\delta : N \to \mathbb{R}$ such that $f_\delta \to f$ strongly in $W^{1,2}(N)$ as $\delta \to 0$ (even $W^{1,p}$ for every $p < \infty$, but we will not need this). The function f_δ is defined as a convolution $f \star \eta_\delta$, for a suitable mollifier η_δ. Moreover, we will check that, if additionally $\nabla f \in BV(N)$, then we have, for all δ sufficiently small, that $\Delta f_\delta = (\Delta f) \star \eta_\delta + E_\delta$, where $(\Delta f) \star \eta_\delta$ is the convolution of the Radon measure Δf with the mollifier η_δ and hence it is identified with its (smooth) density with respect to $\mathcal{H}^{n+1}$, and E_δ is a smooth function bounded in L^∞ by a constant that depends only on N. It would not suffice for our scopes in Section 7.5 to have a convolution procedure that gives $\Delta f_\delta \to \Delta f$ as measures; therefore we give an ad hoc construction here.

We begin with the definitions. The standard smooth bump function on $\mathbb{R}$ is $\eta(x) = e^{-\frac{1}{1-x^2}}$ for $|x| < 1$ and $\eta(x) = 0$ for $|x| \ge 1$. In the following, $\delta < \mathrm{inj}(N)$. We then let $\eta_\delta : N \times N \to \mathbb{R}$ be defined as

$$\eta_\delta(x,y) = \begin{cases} \frac{1}{c_n}\frac{1}{\delta^{n+1}}\eta\left(\frac{d(x,y)}{\delta}\right) & \text{for } d(x,y) < \delta, \\ 0 & \text{for } d(x,y) \geq \delta; \end{cases}$$

here d is the Riemannian distance on N (note that in the first line y belongs to the geodesic ball centred at x with radius δ), and we set

$$c_n = \int_{B_1^{n+1}(0)} \eta(|x|)\, d\mathcal{L}^{n+1} = (n+1)\omega_{n+1}\int_0^1 \eta(s)s^n\, ds,$$

where the integration is with respect to the Lebesgue $(n+1)$-dimensional measure. Therefore for every x, using normal coordinates centred at x, the function $\eta_\delta \circ \exp_x$ integrates to 1 in the ball of radius δ in the tangent space to N at x, endowed with the Euclidean metric.

The sectional curvatures of N are bounded in modulus since N is compact. Recalling Riccati's equation and the Bishop-Günther inequalities (see the final inequality in the proof of [11, theorem 3.17], combined with [11, (3.23)] in the case $P = \{x\}$) there exist $\delta_0 < \text{inj}(N)$ and $C_N > 0$ such that for all $x \in N$ and for $\delta \leq \delta_0$ we have

$$|\mathcal{H}^n(\partial B_\delta(x)) - (n+1)\omega_{n+1}\delta^n| \leq C_N(n+1)\omega_{n+1}\delta^{n+2},$$

where ω_{n+1} is the Euclidean volume of the unit ball in $\mathbb{R}^{n+1}$.

Moreover, denoting by $B_\delta(x)$ the geodesic ball centred at x, by picking a possibly smaller $\delta_0 \in (0, \text{inj}(N))$, we can further ensure that there exists $C_N > 0$ such that, for all $x \in N$ and for all $\delta \leq \delta_0$,

$$\int_{B_\delta(x)} \eta_\delta(x,y)d\mathcal{H}^{n+1}(y) = 1 + O(\delta^2), \tag{A1}$$

where $|O(\delta^2)| \leq C_N\delta^2$. (The constant C_N depends only on the curvature of N, more precisely on the maximum of the modulus of the sectional curvature.)

Proof of (A1). This follows by using the coarea formula in $B_\delta(x)$ for the function $d(x,\cdot)$, for which $|\nabla d(x,\cdot)| = 1$. By the choice of δ_0 above we have a constant $C_N > 0$ such that for all $x \in N$ and for $s \leq \delta_0$,

$$|\mathcal{H}^n(\partial B_s(x)) - (n+1)\omega_{n+1}s^n| \leq C_N(n+1)\omega_{n+1}s^{n+2}.$$

Then by the coarea formula we get

$$\int_{B_\delta(x)} \eta_\delta(x,y)d\mathcal{H}^{n+1}(y)$$

$$= \frac{1}{c_n}\frac{1}{\delta^{n+1}}\int_0^\delta \mathcal{H}^n(\partial B_s(x))\eta\left(\frac{s}{\delta}\right)ds$$

$$\leq \frac{1}{c_n}\frac{1}{\delta^{n+1}}(n+1)\omega_{n+1}\int_0^\delta s^n\eta\left(\frac{s}{\delta}\right)ds$$
$$+\frac{1}{c_n}\frac{1}{\delta^{n+1}}C_N(n+1)\omega_{n+1}\int_0^\delta s^{n+2}\eta\left(\frac{s}{\delta}\right)ds,$$

and using $s^2 \leq \delta^2$ in the second term we conclude that $\int_{B_\delta(x)}\eta_\delta(x,y)d\mathcal{H}^{n+1}(y)$ is bounded above by

$$\frac{1}{c_n}(n+1)\omega_{n+1}\int_0^1 t^n\eta(t)dt+\delta^2\frac{1}{c_n}\frac{1}{\delta^{n+1}}C_N(n+1)\omega_{n+1}\int_0^\delta s^n\eta\left(\frac{s}{\delta}\right)ds$$
$$=1+C_N\delta^2.$$

For the other inequality, namely $\int_{B_\delta(x)}\eta_\delta(x,y)d\mathcal{H}^{n+1}(y)\geq 1-C_N\delta^2$, one proceeds similarly.

Final choice of δ_0. By picking a possibly yet smaller $\delta_0\in(0,\mathrm{inj}(N))$, we can further ensure the following (see [11, (3.35)] or [12, lemma 12.1]). For all $x\in N$ and for $\delta\leq\delta_0$, denoting by $H_{x,\delta}$ the mean curvature function on the geodesic sphere of radius δ around the point x (with respect to the outward-pointing normal, hence $H_{x,\delta}\leq 0$), we have ($-\frac{n}{\delta}$ is the Euclidean mean curvature of the sphere of radius δ in $\mathbb{R}^{n+1}$)

$$\left|H_{x,\delta}+\frac{n}{\delta}\right|\leq C_N\delta \quad \text{on } \partial B_\delta(x).$$

From now on we take $\delta\leq\delta_0$. The convolution of an L^∞ function $f:N\to\mathbb{R}$ with η_δ is the function $f\star\eta_\delta : N\to\mathbb{R}$ defined as follows:

$$(f\star\eta_\delta)(x)=\int_N f(y)\eta_\delta(x,y)d\mathcal{H}^{n+1}(y). \tag{A2}$$

This is a smooth function thanks to the smoothness of η_δ in (x,y). (Smoothness can be checked in charts using standard properties of convolutions.) Note that we have chosen a convolution kernel that does not integrate exactly to 1; however, (A1) suffices to ensure:

Lemma A.1. *Let $f\in W^{1,\infty}(N)$. Then*

(i) $f\star\eta_\delta\to f$ uniformly on N;
(ii) $f\star\eta_\delta\to f$ in $W^{1,2}(N)$.

Proof of Lemma A.1(i). For all x we rewrite

$$\int_N|f(y)-f(x)|\eta_\delta(x,y)d\mathcal{H}^{n+1}(y)$$

as

$$\int_N|f(y)-f(x)|\frac{\eta_\delta(x,y)}{1+O(\delta^2)}\,d\mathcal{H}^{n+1}(y)$$
$$+\int_N\frac{O(\delta^2)\ |f(y)-f(x)|\eta_\delta(x,y)}{1+O(\delta^2)}\,d\mathcal{H}^{n+1}(y);$$

here $O(\delta^2)$ is the function in (A1). Writing L_f for the Lipschitz constant of f, the first term is bounded by $L_f \int_N |x-y| \frac{\eta_\delta(x,y)}{1+O(\delta^2)} d\mathcal{H}^{n+1}(y) \le L_f \delta$. The second term is bounded in absolute value by $\widetilde{C}_N L_f \delta^2$ for all sufficiently small δ. Therefore $\int_N |f(y)-f(x)|\eta_\delta(x,y) d\mathcal{H}^{n+1}(y)$ tends to 0 uniformly in x. Then we compute, recalling (A1),

$$\begin{aligned}
&(f \star \eta_\delta)(x) - f(x) \\
&= \int_N f(y)\eta_\delta(x,y) d\mathcal{H}^{n+1}(y) - \int_N f(x) \frac{\eta_\delta(x,y)}{1+O(\delta^2)} d\mathcal{H}^{n+1}(y) \\
&= \int_N (f(y)-f(x))\eta_\delta(x,y) d\mathcal{H}^{n+1}(y) \\
&\quad + \int_N f(x) \frac{O(\delta^2)}{1+O(\delta^2)} \eta_\delta(x,y) d\mathcal{H}^{n+1}(y).
\end{aligned}$$

The last term is bounded in absolute value by $C_N \|f\|_{C^0(N)} \delta^2$ for all sufficiently small δ. Therefore

$$(f \star \eta_\delta) \to f \quad \text{uniformly on } N. \tag{A3}$$

□

Proof of Lemma A.1(ii). We can choose a finite cover of N by geodesic balls of radius δ_0 in which we fix a local orthonormal frame. In each ball $U \subset N$, we let $\{v_\ell\}_{\ell=1}^{n+1}$ denote the g-orthonormal frame. We can make the nonrestrictive assumption that the collection of open sets $\widetilde{U}$ obtained by setting

$$\widetilde{U} = \{x \in U : \mathrm{dist}(x, \partial U) \ge \delta_0/2\}$$

still constitutes a finite cover of N. Our final aim is to prove that for each U and for every ℓ we have $\int_{\widetilde{U}} |(\nabla(f \star \eta_\delta) - \nabla f) \cdot v_\ell|^2 \to 0$ as $\delta \to 0$. There are only finitely many open sets $\widetilde{U}$, so this implies that $\int_N |\nabla(f \star \eta_\delta) - \nabla f|^2 \to 0$. (Here $|\cdot|$ stands for the g-norm, ∇ for the metric gradient, and $\cdot$ for the g-scalar product of vectors.)

We divide the proof into two parts. In Step 1 we will show that, writing v for one of the v_ℓ, we have $(\nabla f \cdot v) \star \eta_\delta \to (\nabla f \cdot v)$ in $L^2(\widetilde{U})$ (by the choice of $\widetilde{U}$, these convolutions can be defined by staying inside U for $\delta < \delta_0/2$). In Step 2 we will prove that $(\nabla f \cdot v) \star \eta_\delta - \nabla(f \star \eta_\delta) \cdot v$ tends to 0 in $L^\infty(\widetilde{U})$. The two steps together then give

$$\int_{\widetilde{U}} |\nabla(f \star \eta_\delta) \cdot v - \nabla f \cdot v|^2 \to 0$$

as $\delta \to 0$, which is our aim.

Step 1. The first observation is that if $q \in L^\infty(N)$, then for $\mathcal{H}^{n+1}$ − a.e. x we have

$$\int_N |q(y) - q(x)| \eta_\delta(x,y) d\mathcal{H}^{n+1}(y) \to 0 \quad \text{as } \delta \to 0. \tag{A4}$$

This follows by writing, as in Lemma A.1(i),

$$\int_N |q(y)-q(x)|\eta_\delta(x,y)d\mathcal{H}^{n+1}(y)$$
$$= \int_N |q(y)-q(x)|\frac{\eta_\delta(x,y)}{1+O(\delta^2)}d\mathcal{H}^{n+1}(y)$$
$$+ \int_N |q(y)-q(x)|\frac{O(\delta^2)}{1+O(\delta^2)}\eta_\delta(x,y)d\mathcal{H}^{n+1}(y).$$

The second term tends to 0 as argued earlier. The first term tends to 0 if x is a Lebesgue point of q (hence for almost all x). Then we have

$$\int_{\widetilde{U}} |((\nabla f\cdot v)\star\eta_\delta)(x)-(\nabla f\cdot v)(x)|^2 d\mathcal{H}^{n+1}(x)$$
$$= \int_{\widetilde{U}} \Bigg| \int_U ((\nabla f\cdot v)(y)-(\nabla f\cdot v)(x))\eta_\delta(x,y)d\mathcal{H}^{n+1}(y)$$
$$+\left((\nabla f\cdot v)(x)\int_U \frac{O(\delta^2)\eta_\delta(x,y)}{1+O(\delta^2)}d\mathcal{H}^{n+1}(y)\right|^2 d\mathcal{H}^{n+1}(x) \quad \underbrace{\le}_{|a+b|^2\le 2a^2+2b^2}$$
$$\le 2\int_{\widetilde{U}} \underbrace{\left|\int_U ((\nabla f\cdot v)(y)-(\nabla f\cdot v)(x))\eta_\delta(x,y)d\mathcal{H}^{n+1}(y)\right|^2}_{\to 0 \text{ by (A4) for a.e. } x} d\mathcal{H}^{n+1}(x)$$
$$+ \widetilde{C}_N\|\nabla f\|^2_{L^\infty(N)}\delta^4.$$

(In the last term, we have included $\mathcal{H}^{n+1}(\widetilde{U})\le\mathcal{H}^{n+1}(N)$ in the constant $\widetilde{C}_N$.) The braced integrand in the first term tends to 0 for a.e. x by (A4), used with $\nabla f\cdot v$ in place of q. Moreover, the braced expression is bounded for every x by $4\|\nabla f\|^2_{L^\infty(N)}(1+O(\delta^2))^2$, which is summable on N. Hence we can use dominated convergence to conclude that the first term tends to 0 as $\delta\to 0$. The second term tends to 0 as well, therefore we conclude that

$$(\nabla f\cdot v)\star\eta_\delta \to (\nabla f\cdot v) \ \text{ in } L^2(\widetilde{U}).$$

Step 2. We compute the difference between the two (smooth) functions

$$(\nabla f\cdot v)\star\eta_\delta \quad\text{and}\quad \nabla(f\star\eta_\delta)\cdot v$$

and prove that it goes to 0 uniformly on $\widetilde{U}$. We work in normal coordinates centred at an arbitrary point $O\in\widetilde{U}$, namely in the ball $D=\{x\in\mathbb{R}^{n+1} : |x|<\delta_0/2\}$, with exponential map $\exp_O : D\to B_{\delta_0/2}(O)\subset U$. We will evaluate the difference of the two functions at O, making sure that the result does not depend on O. Since we are interested in $\nabla(f\star\eta_\delta)\cdot v$, we need to let x vary in a neighbourhood of O before evaluating the derivative; therefore we will assume $x\in\{x\in\mathbb{R}^{n+1} : |x|<\delta_0/4\}$ and $\delta<\delta_0/4$, so that y stays in D.

We use the customary notation g_{ij} for the metric coefficients, $\sqrt{|g|}$ for the volume density induced by g. We denote by h the Lipschitz function on D given by $f\circ\exp_O : D\to\mathbb{R}$ and by

$\rho : D \times D \to \mathbb{R}$ the mollifier $\rho(x, y) = \eta_\delta(\exp_O(x), \exp_O(y))$ for an arbitrary $\delta < \frac{\delta_0}{4}$. We point out that $\rho(0, y) = \frac{1}{c_n \delta^{n+1}} \eta\left(\frac{|y|}{\delta}\right)$ because we are in normal coordinates, where $|\cdot|$ denotes the Euclidean length. We write ∇_g to denote the metric gradient in D, $(\nabla_g)^i = g^{ij}\partial_{x_j}$. Let v_ℓ be represented in the chart by $\sum v_\ell^j \partial_j$. We fix an arbitrary ℓ and write, for notational ease, $v = (v^1, \dots, v^{n+1}) = (v_\ell^1, \dots, v_\ell^{n+1})$. We will write $\cdot$ between two vectors to denote the scalar product induced by g, so $\nabla_g h \cdot v = \sum g_{ij} g^{ia} \partial_{x_a} h\, v^j = \delta_j^a \partial_{x_a} h\, v^j = \partial_{x_j} h\, v^j\ (= dh(v))$. We restrict to $x \in D_{\delta_0/4}$, and we compute for $\delta < \frac{\delta_0}{4}$ the coordinate expression for $\nabla(f \star \eta_\delta) \cdot v$ (integration is in dy unless otherwise specified):

$$
\begin{aligned}
&\partial_{x_j}\left(\int_D h(y)\rho(x,y)\sqrt{|g|}(y)dy\right) v^j(x) \\
&= \underbrace{v^j(x)\int_D h(y)\partial_{x_j}(\rho(0, y-x))\sqrt{|g|}(y)}_{\mathrm{I}} \\
&\quad + \underbrace{v^j(x)\int_D h(y)\partial_{x_j}(\rho(x,y) - \rho(0, y-x))\sqrt{|g|}(y)}_{\mathrm{II}}. \qquad \text{(A5)}
\end{aligned}
$$

Working on the first term, and using the notation $\rho(0, \cdot) = \rho_0(\cdot)$, we have

$$
\begin{aligned}
\mathrm{I} = -\int_D h(y)(\partial_j \rho_0)(y-x)\sqrt{|g|}(y) \underbrace{=}_{y-x=z}& \\
-\int_D h(x+z)(\partial_j \rho_0)(z)\sqrt{|g|}(x+z)dz& \\
= \int_D (\partial_j h)(x+z)\rho_0(z)\sqrt{|g|}(x+z)dz& \\
+ \underbrace{\int_D h(x+z)\rho_0(z)(\partial_j\sqrt{|g|})(x+z)dz}_{\mathrm{III}} \underbrace{=}_{x+z=y}& \\
= \int_D (\partial_j h)(y)\rho_0(y-x)\sqrt{|g|}(y)dy + \mathrm{III}& \\
= \int_D (\partial_j h)(y)\rho(x,y)\sqrt{|g|}(y)& \\
+ \underbrace{\int_D (\partial_j h)(y)(-\rho(x,y) + \rho(0, y-x))\sqrt{|g|}(y)}_{\mathrm{IV}} + \mathrm{III}.&
\end{aligned}
$$

Consider the first term after the last equality sign, recalling that $v^j(x)$ multiplies I in (A5):

$$
v^j(x)\int_D (\partial_j h)(y)\rho(x,y)\sqrt{|g|}(y)
$$

$$= \underbrace{\int_D v^j(y)(\partial_j h)(y)\rho(x,y)\sqrt{|g|}(y)}_{=((\nabla f\cdot v)\star\eta_\delta)(\exp_O(x))}$$

$$+ \underbrace{\int_D (v^j(x)-v^j(y))(\partial_j h)(y)\rho(x,y)\sqrt{|g|}(y)}_{V}.$$

We now evaluate (A5) at $x = 0$ to obtain

$$(\nabla(f\star\eta_\delta)\cdot v)(O) - ((\nabla f\cdot v)\star\eta_\delta)(O)$$
$$= \mathrm{V}|_{x=0} + v^j(0)\,\mathrm{IV}|_{x=0} + v^j(0)\,|_{x=0} + \mathrm{II}|_{x=0}.$$

It is immediate that $\mathrm{IV}_{x=0} = 0$. In V we have $\rho(0,y) = 0$ for $d(0,y) = |y| \geq \delta$; therefore $|v^j(0) - v^j(y)| \leq C|y|$ for some constant C that depends on derivatives of v in U and can be thus chosen independently of U (there are finitely many U's) and of v_ℓ (finitely many smooth vector fields). We therefore get that $V|_{x=0}$ is bounded in modulus by $C\|\nabla f\|_{L^\infty}\delta(1+O(\delta^2)) \leq C'\|\nabla f\|_{L^\infty}\delta$ for some C' that depends only on the choices of charts and vector fields. In $\mathrm{III}_{x=0}$ the integrand is nonzero only for $|z| \leq \delta$. Let $\widetilde{C}_N > 0$ be an upper bound for the modulus of the second derivatives of the volume element in a normal coordinate system of radius δ_0 centred at an arbitrary point in N (such a constant exists by the compactness of N, the smoothness of the metric, and the fact that $\delta_0 < \mathrm{inj}(N)$). Recalling that in normal coordinates the metric coefficients have vanishing first derivatives at 0, we get that $|\mathrm{III}||_{x=0} \leq C\|f\|_{C^0}\delta$ for all $\delta \leq \delta_0$, with a constant C that only depends on the geometric data. For II, recall that $\rho(x,y) = \frac{1}{c_n\delta^{n+1}}\eta(\frac{d(x,y)}{\delta})$, where d is the Riemannian distance (induced by g); so for each y we have $\partial_{x_j}\rho(x,y) = \frac{1}{c_n\delta^{n+2}}\eta'(\frac{d(x,y)}{\delta})\partial_{x_j}d(x,y)$. On the other hand, $\rho(0,y-x) = \frac{1}{c_n\delta^{n+1}}\eta(\frac{|y-x|}{\delta})$, so for each y we have $\partial_{x_j}\rho(0,y-x) = \frac{1}{c_n\delta^{n+2}}\eta'(\frac{|y-x|}{\delta})\partial_{x_j}|y-x|$. At $x = 0$ we have, for every $y \neq 0$, $\partial_{x_j}|y-x| = \partial_{x_j}d(x,y) = -\frac{y_j}{|y|}$, because we are in normal coordinates, and $d(0,y) = |y|$. Therefore $\mathrm{II}_{x=0} = 0$.

We have therefore proved that $|(\nabla(f\star\eta_\delta)\cdot v)(O) - ((\nabla f\cdot v)\star\eta_\delta)(O)| \leq C\delta$ for C independent of O, that is, $|(\nabla f\cdot v)\star\eta_\delta - (\nabla(f\star\eta_\delta)\cdot v)| \to 0$ uniformly on $\widetilde{U}$. □

Remark A.2. Also note that $\delta \in (0,\delta_0] \to (f\star\eta_\delta) \in W^{1,2}(N)$ is continuous, since η_δ changes smoothly with δ (in fact, this curve is differentiable on $(0,\delta_0)$). Similarly, $\delta \in (0,\delta_0] \to (f\star\eta_\delta) \in C^0(N)$ is continuous.

Next we are going to be interested in $\Delta(f\star\eta_\delta)$ under the additional assumption on f that $\nabla f \in BV(N)$. Here Δ denotes the Laplace-Beltrami operator. Recall that $f\star\eta_\delta$ is smooth, so $\Delta(f\star\eta_\delta)$ is smooth on N. We shall compare this function with $(\Delta f)\star\eta_\delta$, where Δf is a Radon measure. For a Radon measure μ on N we define the (smooth) function $\mu\star\eta_\delta : N \to \mathbb{R}$ as follows:

$$(\mu\star\eta_\delta)(x) = \int \eta_\delta(x,y)\,d\mu(y). \tag{A6}$$

Lemma A.3. *Let $f \in W^{1,\infty}(N)$ with $\nabla f \in BV(N)$. There exists $C_N > 0$ (depending only N and δ_0, once $\eta : \mathbb{R} \to \mathbb{R}$ is fixed) such that, for all $\delta < \delta_0$,*

$$\|(\Delta f) \star \eta_\delta - \Delta(f \star \eta_\delta)\|_{L^\infty(N)} \leq C_N \|f\|_{L^\infty(N)}.$$

Proof. We work in a normal system of coordinates centred at an arbitrary $O \in N$. Let D be the ball centred at $0 \in \mathbb{R}^{n+1}$ of radius δ_0, with $\exp_O : D \to B_{\delta_0}(O)$ denoting the exponential map. We keep notation as in the proof of step 2 of Lemma A.1 (ii), in particular we set $\rho(x, y) = \eta_\delta(\exp_O(x), \exp_O(y))$ and $\rho_0(\cdot) = \rho(0, \cdot)$. The Laplace-Beltrami operator Δ is, in the coordinate chart, $\frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x_i}(\sqrt{|g|} g^{ij} \frac{\partial}{\partial x_j})$, so Δf is $\frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x_i}(\sqrt{|g|} g^{ij} \frac{\partial h}{\partial x_j})$, where $h = f \circ \exp_O$. We compute $\Delta(f \star \eta_\delta)(x)$ in the normal chart, keeping $x \in \{|\cdot| \leq \delta_0/2\}$ and $\delta < \delta_0/2$ so that $y \in D$ in the following computations. Differentiating,

$$\Delta\left(\int_D h(y)\rho(x, y)\sqrt{|g|}(y)dy\right) = \underbrace{\int_D h(y)\Delta(\rho_0(y - x))\sqrt{|g|}(y)}_{\mathrm{I}} + \underbrace{\int_D h(y)\Delta(\rho(x, y) - \rho(0, y - x))\sqrt{|g|}(y)}_{\mathrm{II}}, \tag{A7}$$

where derivatives are taken in x and integration is in dy. We compute, for each y:

$$\begin{aligned}\Delta(\rho_0(y - x)) &= \frac{-1}{\sqrt{|g|(x)}}\partial_{x_i}\left(\sqrt{|g|(x)}g^{ij}(x)\right)(\partial_i\rho_0)(y - x) \\ &\quad + g^{ij}(x)(\partial^2_{ij}\rho_0)(y - x); \\ (\Delta\rho_0)(y - x) &= \frac{1}{\sqrt{|g|(y - x)}}\partial_{x_i}\left(\sqrt{|g|(y - x)}g^{ij}(y - x)\right)(\partial_i\rho_0)(y - x) \\ &\quad + g^{ij}(y - x)(\partial^2_{ij}\rho_0)(y - x).\end{aligned}$$

Therefore

$$\begin{aligned}&\Delta(\rho_0(y - x)) - (\Delta\rho_0)(y - x) \\ &\quad = \left(g^{ij}(x) - g^{ij}(y - x)\right)(\partial^2_{ij}\rho_0)(y - x) \\ &-\left(\frac{1}{\sqrt{|g|(x)}}\partial_{x_i}\left(\left(\sqrt{|g|}g^{ij}\right)(x)\right) + \frac{1}{\sqrt{|g|(x - y)}}\partial_{x_i}\left(\left(\sqrt{|g|}g^{ij}\right)(x - y)\right)\right. \\ &\cdot(\partial_i\rho_0)(y - x)\end{aligned} \tag{A8}$$

and we can rewrite I as follows (so that in the second term we will be able to use (A8)):

$$\int h(x+z)(\Delta\rho_0)(z)\sqrt{|g|}(x+z)dz + \underbrace{\int h(y)(\Delta(\rho_0(y-x)) - (\Delta\rho_0)(y-x))\sqrt{|g|}(y)}_{\mathrm{III}}.$$

We want to evaluate at $x=0$. Let $\tilde\rho_0 = \rho_0 \circ \exp_O^{-1}$ and recall that $f = h\circ \exp_O^{-1}$; then the first term on the right-hand side of the last equality, evaluated at $x=0$, is $\int_N f\Delta\tilde\rho_0\, d\mathcal{H}^{n+1}$. Integrating by parts we rewrite it as $\int_N \Delta f\, \tilde\rho_0\, d\mathcal{H}^{n+1}$, and we get

$$\begin{aligned}\mathrm{I}|_{x=0} &= \int \Delta h(z)\rho_0(z)\sqrt{|g|}(z)dz + \mathrm{III}|_{x=0}\\ &= \int \Delta h(y)\rho(0,y)\sqrt{|g|}(y)dy + \mathrm{III}|_{x=0} = ((\Delta f)\star\eta_\delta)(O) + \mathrm{III}|_{x=0}.\end{aligned}$$

Recall (A7); the statement of Lemma A.3 will therefore follow by estimating $\mathrm{II}|_{x=0}$ and $\mathrm{III}|_{x=0}$, taking care that the estimates should be independent of O. For $\mathrm{III}|_{x=0}$, we use (A8) and the following two facts. Firstly,

$$\int \partial^2_{ij}\rho_0(y)dy = \frac{1}{c_n\delta^2}\left(\int_{B_1}\partial^2_{ij}\eta_1(0,y)dy\right)$$

and

$$\int \partial_i\rho(y)dy = \frac{1}{c_n\delta}\left(\int_{B_1}\partial_i\eta_1(0,y)dy\right)$$

(the two integrals on the right-hand sides are Euclidean and depend only on the explicitly given η, so they will be absorbed into constants). Secondly, since we are in normal coordinates, $g^{ij}(0) = \delta^{ij}$, $\partial_{x_k} g_{ij} = 0$ at 0 for all k; since N is compact, there exists a constant C_{N,δ_0} such that in any normal system of coordinates centred at a point of N and with radius δ_0 ($<\mathrm{inj}(N)$), the second derivatives of the metric coefficients are bounded in modulus by C_{N,δ_0}. Therefore

$$\left|g^{ij}(0) - g^{ij}(y)\right| \le C_{N,\delta_0}|y|^2 \quad \text{and}$$

$$\left|\frac{1}{\sqrt{|g|(-y)}}\partial_{x_i}|_{x=0}\Big((\sqrt{|g|}g^{ij})(x-y)\Big)\right| \le C_{N,\delta_0}|y|.$$

Using these two facts in (A8), and noting that $|y|\le\delta$ on the set where the integrand of $\mathrm{III}|_{x=0}$ does not vanish, we get that $\mathrm{III}|_{x=0}$ is bounded in modulus by $\|f\|_{L^\infty} C_{n,N,\delta_0}\|\eta\|_{C^2(\mathbb{R})} = C\|f\|_{L^\infty}$, with C depending only on fixed geometric data.

For $\mathrm{II}|_{x=0}$ we need to compare, for each $y\neq 0$, $\Delta(\rho(x,y))$ and $\Delta(\rho_0(y-x))$, both evaluated at $x=0$. Let us write $m_\delta(\cdot) = \frac{1}{c_n\delta^{n+1}}\eta(\frac{\cdot}{\delta})$, $m_\delta : \mathbb{R}\to\mathbb{R}$. Then, denoting by d the distance induced by g and by $|\cdot|$ the Euclidean distance, by $|\cdot|_g$ the vector length for g, and by ∇ the g-gradient, we get for each $y \neq 0$ (derivatives with respect to $\cdot$)

$$\begin{aligned}\Delta(\rho(\,\cdot\,,y)) &= \Delta(m_\delta(d(\,\cdot\,,y))) \\ &= m''_\delta(d(\,\cdot\,,y))|\nabla d(\,\cdot\,,y)|^2_g + m'_\delta(d(\,\cdot\,,y))\Delta d(\,\cdot\,,y), \\ \Delta(\rho_0(y-\cdot\,)) &= \Delta(m_\delta(|y-\cdot\,|)) \\ &= m''_\delta(|y-\cdot\,|)|\,\nabla|y-\cdot\,|\,|^2_g + m'_\delta(|y-\cdot\,|)\Delta|y-\cdot\,|.\end{aligned}$$

Evaluating at $\cdot = 0$ we note that $m''_\delta(d(0,y))|\nabla d(0,y)|^2_g = m''_\delta(|y|)|\nabla|y||^2_g$ and, moreover, $m'_\delta(d(0,y)) = m'_\delta(|y|)$, because in normal coordinates we have $d(0,y) = |y|$ and $\nabla|y-\cdot\,| = \nabla d(\,\cdot\,,y) = -\frac{y}{|y|}$ at the point $\cdot = 0$ (for any chosen $y \neq 0$). We therefore need to compare, for any $y \neq 0$, $\Delta d(\,\cdot\,,y)$ and $\Delta|y-\cdot\,|$ at 0. The former is the opposite of the mean curvature at 0 of a geodesic sphere centred at y with radius $d(0,y)$ (as usual, we compute the scalar mean curvature with respect to the outward-pointing normal to the sphere). On the other hand, recall that computing Δ at 0 is the same as computing the Euclidean Laplacian, therefore $-\Delta|y-\cdot\,|$ at 0 is the Euclidean mean curvature at 0 of a Euclidean sphere centred at y with radius $|y|$, hence $\Delta|y-\cdot\,| = -\frac{n}{|y|}$ at 0. The difference $\Delta d(\,\cdot\,,y) - \Delta|y-\cdot\,|$ is therefore bounded in modulus by $C_N|y|$, thanks to the initial choice of δ_0. Since we can take $|y| \leq \delta$ in $\mathrm{II}|_{x=0}$ (because $\rho = 0$ otherwise) we can estimate $(\Delta(\rho(x,y)) - \Delta(\rho_0(y-x)))|_{x=0}$ in modulus by $\|m'_\delta\|_{L^\infty} C_N \delta \leq \frac{1}{\delta^{n+1}} C_\eta C_N$; integrating on $\{|y| \leq \delta\}$ we get that $\mathrm{II}|_{x=0}$ is bounded in modulus by $C_\eta C_N \|f\|_{L^\infty}$.

We have thus obtained $|(\Delta f) \star \eta_\delta - \Delta(f \star \eta_\delta)|(O) \leq C\|f\|_{L^\infty}$ with C depending only on (N, g) and on the fixed entities δ_0, η. The arbitraryness of O gives the result. □