Current Developments in Mathematics, 2017

Regularity and compactness for stable codimension 1 CMC varifolds

Neshan Wickramasekera

Abstract. According to the Allard regularity theory, the set of singular points (i.e. non $C^{1,\alpha}$-embedded points) of an integral n-varifold with generalized mean curvature locally in L^p for some $p > n$ is a nowhere dense (closed) subset of the support of the varifold. A well-known codimension 1 example due to Brakke shows that not much can be said about the Hausdorff measure of the singular set; it need not have zero n-dimensional measure. We survey recent work that shows nonetheless, that in codimension 1, all is well whenever those parts of the varifold that are regular (in certain specific ways) satisfy further hypotheses, namely: (a) that the orientable portions of the $C^{1,\alpha}$ embedded part and the C^2 immersed part are respectively stationary (or equivalently CMC) and stable with respect to the area functional for volume preserving deformations, and (b) that there is appropriate control on two types of singularities—called classical and touching singularities—that are formed by $C^{1,\alpha}$ embedded pieces of the varifold coming together in a regular fashion. This work builds on and extends the recent codimension 1 theory for zero mean curvature stable varifolds with no classical singularities, and the earlier fundamental curvature estimates of Schoen–Simon–Yau and of Schoen–Simon. We include a brief discussion of these previous works and their role in (different approaches to) the existence theory for minimal hypersurfaces in compact Riemannian manifolds. The main focus of the survey is on the novel aspects of the CMC regularity and compactness theory and the associated curvature estimates.

Contents

1. Introduction

Let N be a smooth $(n+1)$-dimensional Riemannian manifold and let $k \in \{1, 2, \ldots, n\}$. A basic geometric functional on the space of k dimensional submanifolds (with or without boundary) of N is the k-dimensional area, i.e. the mapping $\mathcal{A} : M \mapsto \mathcal{H}^k(M)$, where $\mathcal{H}^k$ is the k-dimensional

Hausdorff measure on N induced by the Riemannian metric. It is of interest to study critical points of $\mathcal{A}$, sometimes restricted to families of submanifolds satisfying additional topological or geometric constraints such as fixed genus, fixed relative homology class or, when it is meaningful, fixed enclosed volume. For certain large families of submanifolds whose typical elements may exhibit little geometric control, finding a critical point of a functional such as $\mathcal{A}$ (restricted to the family) provides a way of choosing a geometrically well-behaved representative whose special properties may be exploited in applications.

Critical points of $\mathcal{A}$ have geometric properties that can be described in terms of their mean curvature in some way. For instance, submanifolds that are critical points of $\mathcal{A}$ for unconstrained ambient space (infinitesimal) deformations away from their boundary—which we shall call *stationary submanifolds* of N—are characterized by the simple condition that they are *minimal submanifolds*, i.e. they have zero mean curvature. When $k = 1$, these critical points are geodesics of N, so minimal submanifolds are a higher dimensional analogue of geodesics. In a similar vein, in case $k = n$ a critical point of $\mathcal{A}$ among oriented hypersurfaces that enclose a region of fixed volume, or equivalently, an oriented hypersurface that is an unconstrained critical point of the functional

$$\mathcal{J}(M) = \mathcal{A}(M) + \lambda vol(M)$$

for some constant λ where vol is the volume enclosed by M, is a CMC hypersurface, i.e. one that has constant scalar mean curvature relative to a continuous unit normal on M. In fact an oriented hypersurface is stationary with respect to $\mathcal{J}$ if and only if its scalar mean curvature with respect to an appropriate choice of unit normal is λ. More generally, if $g : N \to \mathbb{R}$ is a smooth function, then an oriented hypersurface is a critical point of the functional

$$\mathcal{J}_g(M) = \mathcal{A}(M) + vol_g(M),$$

where vol_g is the *relative enclosed g-volume* (see [**BelWic-1**]), if and only if its scalar mean curvature with respect to a choice of unit normal is equal to g everywhere.

1.1. Existence of critical points of area. Little is known concerning the general question of when one can find a k dimensional properly embedded smooth submanifold that is a critical point of $\mathcal{A}$, with or without other constraints, in a given smooth compact Riemannian manifold N of dimension $> k$. What is known in several instances of this question is that there are critical points that are smoothly embedded away from a possible (closed) singular set of lower Hausdorff dimension that can be bounded from above by an integer depending only on k (and in particular independent of N and the metric on N). Except in the case of hypersurfaces in low dimensional manifolds, not much is known about how typical it is (e.g. with respect to metrics on N) that such a singular set cannot be avoided. There are however various explicit examples and constructions that do exhibit singular behaviour in stationary submanifolds (see Sect. 2.2 for a brief discussion on this); these examples show in particular that the known existence results yield solutions that are "optimally regular," meaning that the dimension upper bounds on their singular sets are the best possible general bounds independent of the ambient space.

In line with a widely utilised strategy in PDE theory, a fruitful approach to proving existence of optimally regular k-dimensional critical points of $\mathcal{A}$ starts with a construction, in the first instance, of "weak (or generalized) solutions" belonging to a space of generalized k-dimensional submanifolds whose general elements are allowed to be highly singular. This is followed by an independent study of regularity of the weak solutions to bound the size of their (essential) singular set. The space of generalised k-dimensional submanifolds to work in in this process needs to be chosen with care. On the one hand, construction of weak solutions involves taking limits of generalized submanifolds, and hence requires choosing a topology on the space of generalized submanifolds with respect to which appropriate closure properties hold for the space. On the other hand, in order to be able to do analysis on the weak solutions to prove their regularity (starting with extending the first variation formula—a first order computation—to obtain an analytically useful characterizaton of weak solution), it is necessary that the generalized k-dimensional submanifolds correspond to objects having k-dimensional tangent spaces $\mathcal{H}^k$ a.e., i.e. k-rectifiable sets. Compatibility between these two competing requirements necessitates allowing multiplicities on the rectifiable sets, and consequently also extending the functional $\mathcal{A}$ to the k-dimensional mass functional—i.e. k-dimensional Hausdorff measure weighted by multiplicity. Powerful theorems in Geometric Measure Theory guarantee that in so far as establishing existence of weak solutions is concerned it suffices to consider positive *integer* multiplicities. Having to allow multiplicity ≥ 2 however leads to very serious difficulties concerning regularity of the weak solutions.

The pioneering work of Federer–Fleming published in 1960 ([**FedFle60**]) implemented the construction-of-weak-solution part of this approach for k-th relative homology classes of N, by establishing a fundamental closure theorem for integer multiplicity k-rectifiable currents (roughly speaking, k-rectifiable sets with positive integer-valued multiplicity functions and oriented tangent spaces), taken with the topology induced by weak convergence of k-dimensional currents. The Federer–Fleming theorem implies the existence, in any given relative homology class, of an integer multiplicity rectifiable current that is mass minimizing within the class; in the absence of boundary, this result gives an absolutely mass minimizing cycle in its integral homology class.

When the k-th homology is trivial (e.g. when N is an n-dimensional topological sphere with a Riemannian metric), there are of course no non-trivial absolutely mass minimizing k-cycles. Thus in that case the Federer–Fleming compactness theorem is inadequate to produce non-trivial stationary k-cycles. A few years later, Almgren ([**Alm65**]) developed a deep Morse theoretic alternative applicable in such situations that is designed to capture k-cycles that are unstable critical points. Almgren's work vastly extended the ideas in the earlier work of Birkhoff ([**Bir17**]) on the existence of closed geodesics in 2-spheres. For this approach it is important that the mass functional is continuous with respect to the topology on the space of generalised submanifolds. Since mass is only lower semicontinuous with respect to weak convergence of currents, Almgren's theory required an alternative topology. Giving up the orientation on the tangent spaces, it uses the more general space of integral k-varifolds, i.e. the space of integer multiplicity k-rectifiable sets with the topology of measure theoretic convergence (induced by the weak convergence of Radon measures on $\{(x, S) : x \in N, \quad S \in G_k(T_x N)\}$ where $G_k(T_x N)$ is the space of k-dimensional subspaces of the tangent space $T_x N$, rather than the weak convergence of Radon measures on N, so as to ensure that not only the mass functional but also the first variation is well-behaved). A key outcome of Almgren's theory of varifolds is the existence of a non-trivial stationary integral k-cycle in N when N is compact, which arises as a min-max critical point of the mass functional. Almgren's method, strengthened further in the subsequent work of Pitts ([**Pit77**]), has come to be known as the Almgren–Pitts min-max theory. In the hypersurface case, with the help of the regularity theory for locally area minimizing hypersurfaces and the compactness theories of Schoen–Simon–Yau ([**SSY75**]) and of Schoen–Simon ([**SchSim81**]) for stable minimal hypersurfaces, it gave the first proof of the existence of an optimally regular critical point in any given compact manifold. The Almgren–Pitts min-max theory has recently attracted much attention since the beautiful work of Marques–Neves ([**MarNev14a**]) applying it to resolve the Willmore conjecture. It has been extended in various ways in the hypersurface case (see e.g. [**MarNev16**] [**IMN18**], [**ChaLio16**], [**ZhoZhu17**]), including in the recent work of Song ([**Son18**]) that proves, building on [**MarNev14b**], Yau's conjecture in low dimensions: there are infinitely many minimal hypersurfaces in any compact Riemannian manifold.

Also in recent years, a fundamentally different, PDE theoretic alternative to the Almgren–Pitts existence theory has emerged. This is the result of the combined work of Guaraco ([**Gua15**]), Hutchinson–Tonregawa ([**HutTon00**]), Tonegawa ([**Ton05**]) and Tonegawa–Wickramasekera ([**TonWic12**]), and it relies on the regularity theory of [**Wic14a**] (discussed below) extending the Schoen–Simon theory. In this method, the basic idea is to obtain a minimal hypersurface as a weak limit of level sets of solutions u_j to a sequence of singularly perturbed (elliptic) Allen–Cahn equations with perturbation parameter $\epsilon_j \to 0^+$. The key difference in this approach is that it uses, in place of the varifold min-max construction of Almgren–Pitts for the area functional, a more elementary and abstract min-max existence theorem for critical points of functionals, on Hilbert spaces, satisfying the Palais–Smale compactness condition. This theorem applied to the Allen–

Cahn functional on $W^{1,2}(N)$ produces u_j as above with uniformly bounded Allen–Cahn energy and uniformly bounded Morse index with respect to the Allen–Cahn energy ([**Gua15**]). This amounts to a dramatic simplification on the part of the min-max construction. An optimally regular minimal hypersurface is obtained from the u_j's by first establishing that in the varifold limit as $j \to \infty$, the level sets of u_j concentrate on a codimension 1 stationary integral varifold ([**HutTon00**]), and then applying a general varifold regularity theorem ([**Wic14a**]) to the limit ([**TonWic12**]). Thus, in this approach the existence of minimal hypersurfaces is a fairly direct, clean corollary of PDE principles and GMT regularity and compactness principles of broader significance. We shall present a more detailed discussion of this method in Sect. 5.2 below.

1.2. Regularity theory: a brief overview. The justification of a variational approach to the existence question for minimal or related submanifolds ultimately rests, of course, on the strength of the regularity results that can be applied to the weak solutions. Although much still remains to be understood in this direction, there has been a great deal of progress in the regularity theory in the period spanning the past seven or so decades. This body of work has lead to several optimal theorems concerning various classes of minimal submanifolds. Equally importantly, the work has lead to the discovery of a number of powerful general tools and techniques of far reaching significance beyond their immediate utility for minimal submanifolds. Indeed, these techniques have had a profound impact on several other areas of Geometric Analysis and PDE including harmonic maps, free boundary problems, geometric flows and general relativity, to name just a few.

A brief summary of known regularity results that imply optimal regularity of the Federer–Fleming mass minimizing cycles, the codimension 1 Almgren–Pitts min-max stationary varifolds or the codimension 1 stationary varifolds arising from the Allen–Cahn approximation is as follows: (1) the regularity theory for mass minimizing integer multiplicity currents when $k = n$ (i.e. in the hypersurface case); this asserts the embeddedness of minimizers away from the boundary and away frrom a closed singular set of Hausdorff dimension at most $(n - 7)$—a result that has been achieved in a step by step fashion over a decade or so starting with the pioneering work of De Giorgi ([**DeG61**]) and of Reifenberg ([**Rei60**]) establishing almost everywhere regularity followed by successively better control on the singular set established in the work of Federer ([**Fed70**]), Fleming ([**Fle62**]), Almgren ([**Alm66**]) and finally of Simons ([**JSim68**]) that established the optimal $n - 7$ Hausdorff dimension bound; (2) the regularity theory of Almgren ([**Alm83**]) for mass minimizing integer multiplicity currents of dimension $k < n$ establishing embeddedness of the minimizers away from the boundary and away from a closed singular set of Hausdorff dimension $\leq (k - 2)$ which again is the optimal dimension bound (see also

[**DeLSpa14**], [**DeLSpa16a**], [**DeLSpa16b**]); (3) the compactness theory of Schoen–Simon–Yau ([**SSY75**]) for stable minimal hypersurfaces with no singularities and dimension $k \leq 5$, extended in the work of Schoen–Simon ([**SchSim81**]) to general dimensions allowing also an a priori small singular set, giving subsequential convergence (in C^m_{loc} for every m if $k \leq 6$, or, for general k, in the varifold topology and in C^m_{loc} for every m away from the singularities) of locally uniformly mass bounded sequences of such hypersurfaces.

The combination of the Schoen–Simon compactness theory and the regularity theory for codimension 1 mass minimizing integer multiplicity currents yields embeddedness of the Almgren–Pitts min-max solutions of dimension $k = n \geq 2$ away from a closed set of Hausdorff dimension $\leq n - 7$.

The optimal regularity of the stationary varifolds arising in the Allen–Cahn approach is obtained by applying a recently developed optimal regularity theory for stable, stationary codimension 1 integral varifolds ([**Wic14a**]). This theory replaces the a priori singular set size hypothesis of the Schoen–Simon theory with a milder (necessary) structural hypothesis that just requires ruling out a very specific type of singularity—called *classical singularities* (see the defintion in Sect. 1.3 below or Definition 4.1). This condition is readily checked in the context of the the limit varifolds arising from the Allen–Cahn solutions ([**TonWic12**], [**Gua15**]) as well as in other applications, e.g. as in [**Wic14b**]. (Incidentally, since locally area minimizing hypersurfaces cannot have classical singularities, this theory also subsumes the aforementioned regularity theory for codimension 1 area minimizing integer multiplicity currents). A further generalisation of this theory to CMC varifolds, summarised below in Sect. 1.3 and discussed in more detail in Sects. 3–10, shall be our primary focus in this article.

To finish this quick overview of regularity theories, we mention the fundamental regularity theory of Allard ([**All72**]) which implies that a stationary integral k-varifold, or more generally any integral k-varifold (of any codimension) with generalised mean curvature locally summable to a power $p > k$, is regular (i.e. smoothly embedded in the stationary case, or $C^{1,\alpha}$ embedded in the general case for $\alpha = 1 - \frac{k}{p}$ with $p < \infty$) on a dense (open) subset of its support. In particular, the Almgren–Pitts min-max stationary cycles of higher codimension in compact manifolds are smoothly embedded on an open dense subset of the support of the cycle. Beyond this, no further regularity of these min-max stationary varifolds is known, even in the two dimensional case.

1.3. Codimension 1 regularity theory: recent developments. In a series of works over the past several years, beginning with [**Wic14a**] and continued with the work [**BelWic18**] (of C. Bellettini and the author) and the work [**BCW18**] (of C. Bellettini, O. Chodosh and the author), the compactness theories of Schoen–Simon ([**SchSim81**]) and of Schoen–Simon–Yau ([**SSY75**]) have been strengthened and extended. This recent work considers the class of codimension 1 integral n-varifolds whose generalized mean curvature is locally summable to a power $p > n$. The Allard regularity theory implies that the C^1 embedded part of such a varifold is a dense subset of its support. However, no useful control of the dimension of the singular set can follow from the L^p hypothesis on the mean curvature alone, even when $p = \infty$, as shown by an example due to Brakke ([**Bra78**]) (refined recently by Menne–Kolasiński [**KolMen15**]) which has a singular set of positive n-dimensional measure.

The work [**BelWic18**] makes (additional) variational hypotheses on the regular parts of such a varifold requiring (slightly imprecisely speaking) that the regular parts are stationary and stable with respect to the mass functional for volume preserving deformations supported away from the singular set. (The earlier work [**Wic14a**] assumed stationarity everywhere and stability on the regular part for unconstrained compactly supported deformations). The main new discovery in [**Wic14a**], [**BelWic18**] is that for integral codimension 1 varifolds with genralized mean curvature in L^p_{loc} for some $p > n$ and satisfying these variational hypotheses, certain *structural conditions involving only the parts of the varifold that are made up of regular (i.e. $C^{1,\alpha}$) embedded pieces coming together in a regular fashion*, in contrast to a more stringent assumptions on the size of the entire singular set such as in [**SchSim81**], are sufficient (and necessary) for the same sharp $n - 7$ bound on the Hausdorff dimension of the (genuine) singular set. (Most recently, this work has been generalised even further to allow stability and stationarity to hold for deformations that are constrained in a more general way, or equivalently, to allow the mean curvature of the regular part to be prescribed in a more general way ([**BelWic-1**])).

Thus, this regularity theory has the advantage that control on the *size* of the singular set is not a hypothesis but is entirely a conclusion. With less to check a priori on the singular set, it can be applied in a wider variety of situations. Indeed, the main results of [**Wic14a**], in combination with the work of Ilmanen ([**Ilm96**]) and of Solomon–White ([**SW89**]), have lead to a strong maximum principle and a unique continuation theorem for stationary codimension 1 integral varifolds ([**Wic14b**]), providing an optimal answer to a long studied question on which several partial results had previously been known ([**Mir67**], [**Mos77**], [**Sim87**], [**SW89**], [**Ilm96**]). More relevant to the themes of the present article is its application, mentioned above, to the regularity question for minimal submanifolds arising as limit-interfaces corresponding to sequences of stable (or more generally, bounded Morse index) solutions of the singularly perturbed Allen–Cahn equation ([**TonWic12**], [**Gua15**]). The latter application in turn has made way for a considerably simpler and streamlined PDE theoretic alternative to the Almgren–Pitts existence theory for minimal hypersurfaces. We shall discuss this new existence theory briefly in Sect. 5.2 below.

In the remainder of this introduction we wish to describe the main content of the sharp regularity theory in more detail, for which we need the following definitions:

(1) A *classical singularity* of a set $S \subset N$ is a point $Y \in S$ such that, for some $\alpha \in (0,1)$ and some small $\rho > 0$, $S \cap B_\rho(Y)$ is the union of a finite number of at least three $C^{1,\alpha}$ embedded hypersurfaces-with-boundary M_j that meet pairwise, with at least one pair transversely, only along a $C^{1,\alpha}$ free boundary containing Y and common to all M_j.

(2) A *touching singularity* of a set $S \subset N$ is a point $Y \in S$ that is not a classical singularity of S and near which S is not a C^1 embedded hypersurface but for some $\alpha \in (0,1)$ and some small $\rho > 0$, $S \cap B_\rho(Y) = M_1 \cup M_2$ for precisely two distinct embedded $C^{1,\alpha}$ hypersurfaces M_1, M_2 of $B_\rho(Y)$.

(3) If Y is a touching singularity of a set $S \subset N$ and ρ, M_1, M_2 are as in (2) above, then the *coincidence set* of S in $B_\rho(Y)$ is the set $M_1 \cap M_2$.

Any local hypothesis concerning classical or touching singularities of a (singular) hypersurface S should be viewed as a structural condition on S since by definition, the entire hypersurface near such a singularity has a "$C^{1,\alpha}$ regular structure;" in verifying such a hypothesis one is thus allowed to assume (a form of) regularity of the hypersurface. Therein lies the advantage of theorems (such as those described below) in which control of classical or touching singularities are the only hypotheses on the singular set.

The main content of the works [**Wic14a**], [**BelWic18**] and [**BCW18**] can now be described as follows:

(i) The work [**Wic14a**] concerns stationary hypersurfaces (i.e. codimension 1 integral varifolds with zero generalized mean curvature) that are stable on every orientable portion of their regular part. Subject to the (necessary) condition that the hypersurface has no classical singularities,[1] it gives embeddedness of the hypersurfaces away from a closed set of Hausdorff dimension $\leq n-7$.

(ii) The work [**BelWic18**] builds on and generalises the regularity theory of [**Wic14a**]. It concerns the class of weakly stable CMC hypersurfaces of an open subset of a Euclidean space[2] having no classical singularities and satisfying an additional (necessary) structural condition involving touching singularities. More precisely, this work considers codimension 1 integral n-varifolds of an open set $\mathcal{U} \subset \mathbb{R}^{n+1}$ having generalized mean curvature locally in L^p for some fixed $p > n$ and:

  (a) having any orientable portion of the C^1 embedded part (a dense set by the Allard regularity theory [**All72**]) stationary

[1] In [**Wic14a**], the no-classical-singularities condition was called the α-structural hypothesis.

[2] The generalisation to Riemannian ambient spaces (and to mean curvature prescribed more generally) is discussed in [**BelWic-1**].

with respect to the mass functional for deformations preserving enclosed volume relative to a choice of orientation on that portion;

(b) having any orientable portion[3] of the C^2 immersed part (a classical CMC immersion by assumption (a), elliptic regularity and the denseness of the embedded part) stable with respect to deformations preserving enclosed volume relative to a choice of orientation on that portion;

(c) having no classical singularities and

(d) satisfying the condition that the coincidence set in a neighborhood of each of its touching singularities is $\mathcal{H}^n$-null.

In view of the De Giorgi structure theory ([**DeG61**]), the structural condition (ii)(d) is automatically satisfied in the important special case when the varifold corresponds to the reduced boundary (taken with multiplicity 1) of a Caccioppoli set.

The main regularity theorem of [**BelWic18**] is that if a varifold satisfies the above conditions, then there is a closed set $\Sigma \subset \mathcal{U}$ with Hausdorff dimension $\leq n-7$ such that away from Σ, its support is a classical CMC immersion that is quasi-embedded; this means that either it is an embedded C^2 minimal hypersurface away from Σ, or locally away from Σ it is either a single C^2 embedded CMC disk or the union of precisely two C^2 embedded CMC disks each lying on one side of the other and intersecting only tangentially along a set contained in an $(n-1)$-dimensional embedded submanifold. And moreover, in case the mean curvature is non-zero, away from Σ the hypersurface is an orientable immersion with a continuous choice of unit normal with respect to which the value of the scalar mean curvature is equal to the same constant everywhere. Both structural hypotheses (ii)(c) and (ii)(d) are necessary for these regularity conclusions.

The work [**BelWic18**] also provides an associated compactness theorem which says that any sequence of hypersurfaces in $\mathcal{U}$ which intersect a fixed compact subset of $\mathcal{U}$ and satisfy the above assumptions together with a locally uniform mass bound and a uniform mean curvature bound has a subsequence that converges in the varifold topology to a hypersurface in $\mathcal{U}$ satisfying the same assumptions as above (and hence also regularity conclusions as above).

(iii) The work [**BCW18**] establishes curvature estimates for weakly stable CMC hypersurfaces considered in [**BelWic18**], making the compactness theorem of [**BelWic18**] a posteriori effective; in particular, these curvature estimates imply that the convergence in the preceding compactness statement, in addition to being in the

[3] In case the (constant) mean curvature of the C^1 embedded part is non-zero, the entire C^2 immersed part *is* orientable.

varifold topology in $\mathcal{U}$, is locally in the C^m topology for every m away from the "genuine" singular set Σ (of Hausdorff dimension $\leq n-7$) of the limit hypersurface.

The structural conditions ii(c) and ii(d) are both implied by the assumption $\mathcal{H}^{n-1}(\operatorname{sing} V) = 0$, but this size restriction is not preserved under taking limits. See the opening paragraphs of Sect. 6 for a more detailed discussion on the singular set assumptions (including perils of allowing $\mathcal{H}^{n-1}(\operatorname{sing} V) > 0$ with no further conditions).

Beyond the implication of Allard's regularity theorem that the singular set is nowhere dense, no previously known regularity theorem implies any control of the singular set of a varifold satisfying the hypotheses listed in (ii) (or even in (i)). In particular, in the theorems described above, the regular part a priori could be very small in measure and thus in particular the stability hypothesis a priori is valid only on an (open) set of potentially very small measure; this is in contrast to the setting of [**SchSim81**] where the small-singular-set hypothesis implies the validity of the stability inequality a priori over the entire hypersurface including the singular set, and hence in particular L^2 boundedness on the second fundamental form locally uniformly in the ambient space $\mathcal{U}$.

Indeed, one of the central analytic difficulties overcome by the methods developed in the above works (i), (ii) is the one arising from the lack of information a priori on the behaviour of the second fundamental form of the regular set on approach to the (a priori potentially large) singular set. The analytic machinery developed in the works (i)–(iii) taken together establishes a key pointwise bound on the second fundamental form which can be described follows: *if an integral n-varifold in an $(n+1)$-dimensional ball satisfying a mass bound has generalized mean curvature sufficiently small in L^p for some $p > n$, if the varifold lies sufficiently close to an n-dimensional plane, and if it further satisfies the two structural conditions (ii)(c) and (ii)(d) above and the first and second variation hypotheses (ii)(a) and (ii)(b) on its regular parts, then it is a properly immersed smooth CMC hypersurface in the interior which is embedded except possibly on a set of touching singularities contained in a smooth $(n-1)$-dimensional submanifold; moreover, its second fundamental form in the interior is uniformly bounded from above by a fixed constant times the sum of its L^2 distance to the hyperplane and the absolute value of the (constant) mean curvature.*

In particular, this result rules out branch point singularities, i.e. non-immersed points where one tangent cone is supported on a hyperplane, in varifolds satisfying the hypotheses as in (ii) above.

Estimates of this nature had previously been established in instances where either the mass of the varifold is close to the mass of the unit n-dimensional ball ([**All72**]) (in which case the structural conditions (ii)(c), (ii)(d) as well as the variational hypotheses (ii)(a), (ii)(b) are all unnecessary, the conclusion is a $C^{1,\alpha}$ estmate as opposed to a C^2 estimate, and the result holds in arbitrary codimension) or where the singular set is a priori asumed to be sufficiently small ([**SSY75**], [**SchSim81**]) (a condition stronger than the two structural conditions (ii)(c), (ii)(d) combined). Easy examples show that a curvature estimate as above does not hold if we remove any one of the stated hypotheses (see Sect. 6.2), including the codimension 1 hypothesis.

The theorems established in [**Wic14a**], [**BelWic18**], [**BCW18**] are built upon a wide circle of ideas and results developed in previous works in regularity theory over a period spanning six decades. These include the work of De Giorgi ([**DeG61**]) on regularity of area minimizing hypersurfaces; the work of De Giorgi ([**DeG54**], [**DeG55**]) on the structure of sets of locally finite perimeter; the work of Allard ([**All72**]) on regularity of integral varifolds with appropriately summable generalized mean curvature; the work of Schoen–Simon–Yau ([**SSY75**]) and of Schoen–Simon ([**SchSim81**]) on compactness for stable minimal hypersurfaces with small singular sets; the work of Simon ([**Sim93**]) on asymptotic behaviour near singularities of minimal submanifolds in multiplicity 1 classes; the work of Hardt–Simon ([**HarSim79**]) on boundary regularity for area minimizing hypersurfces; the work of Simons ([**JSim68**]) on classification under curvature conditions of embedded minimal hypercones of Euclidean space, including stable hypercones in low dimensions; the work of Almgren ([**Alm83**]) on parametrization of mean curvature controlled integral varifolds via multiple-valued Lipschitz functions; the work of Almgren ([**Alm83**]) on generalized stratification of singular sets of integral varifolds with appropriately controlled generalized mean curvature; the work of Cheng–Cheung–Zhou ([**CCZ08**]) on the structure of weakly stable minimal hypersurfaces at infinity and the work of White ([**Whi87**]) on curvature estimates for stationary surfaces.

2. First variation, stationarity and the mean curvature

2.1. The smooth setting. Let N be a smooth, $(n+1)$-dimensional Riemannian manifold, $k \in \{1, 2, \ldots, n\}$ and let M be an embedded k-dimensional C^1 submanifold of N with locally finite mass in N (i.e. with $\mathcal{H}^k(M \cap K) < \infty$ for each compact set $K \subset N$). To define the notion of (unconstrained) critical point of the k-dimensional area functional on N, we start with a smooth compactly supported vector field X on N, and note that we can use X to deform M in the following way: choose $\epsilon > 0$ and a smooth map

$$\varphi : (-\epsilon, \epsilon) \times N \to N$$

such that: (i) $\varphi(0, x) = x$ for each $x \in N$; (ii) $\varphi(t, x) = x$ for every $x \in N \setminus \operatorname{spt} X$ and every $t \in (-\epsilon, \epsilon)$; (iii) $\frac{d}{dt}\big|_{t=0} \varphi(t, x) = X(x)$ for every $x \in N$; (iv) the map $\varphi_t : N \to N$ given by $\varphi_t(x) = \varphi(t, x)$ is a diffeomorphism of N for each $t \in (-\epsilon, \epsilon)$. Corresponding to any given X, there are many such maps φ, and each gives rise to a 1-parameter family of submanifolds $\varphi_t(M)$, $t \in (-\epsilon, \epsilon)$, with $\varphi_0(M) = M$.

The *first variation* of M with respect to the area functional in the direction X, denoted $\delta M(X)$, is defined by

$$\delta M(X) = \frac{d}{dt}\bigg|_{t=0} \mathcal{H}^k(\varphi_t(M \cap \operatorname{spt} X)).$$

The first variation of M depends only on X, and not on the choice of φ. In fact, a calculation (see [**Sim83**, Sect. 9]) shows that

$$\delta M(X) = \int_M \operatorname{div}_M X \, d\mathcal{H}^k \tag{1}$$

where $\operatorname{div}_M X$ denotes the tangential divergence of X on M, given, at a point $x \in M$, by

$$\operatorname{div}_M X(x) = \sum_{j=1}^{k} < \tau_j, \nabla_{\tau_j} X(x) > \tag{2}$$

for any orthonormal basis $\{\tau_1, \tau_2, \ldots, \tau_k\}$ for the tangent space $T_x M$ (and is independent of the choice of the orthonormal basis). The identity (1) is known as the *first variation formula*. The submanifold M is said to be *stationary* in N if

$$\delta M(X) = 0$$

for every compactly supported vector field X on N.

Now suppose M is of class C^2. The second fundamental form of M assigns to any pair of tangent vector fields v_1, v_2 to M the vector field

$$A(v_1, v_2) = (\nabla_{v_1} v_2)^{\perp},$$

where $\perp$ denotes the component of a vector normal to M. This is a symmetric bilinear map, and for any $x \in M$, $A(v_1, v_2)(x)$ depends only on the values $v_1(x)$ and $v_2(x)$. The trace of A is called the mean curvature of M, denoted H_M. So for $x \in M$, $H_M(x)$ is a vector normal to M at x, given by

$$H_M(x) = \sum_{i=1}^{k} A(\tau_i, \tau_i)$$

for any orthonormal basis $\{\tau_1, \ldots, \tau_k\}$ for $T_x M$. The submanifold M is said to be *minimal* of its mean curvature vanishes everywhere. If M is 1-dimensional, this condition means that M is a geodesic of N.

With the vector field X as above, for each $x \in M$, we can write $X(x) = X^{\top}(x) + X^{\perp}(x)$, the sum of components of $X(x)$ tangential and normal to M at x. It follows from the definition of H_M that $\operatorname{div}_M X^{\perp} = - < H_M, X >$. If $\overline{M}$ is a submanifold (of class C^2) with boundary $\partial M = \overline{M} \setminus M$, then by applynig the divergence theorem we obtain that

$$\int_M \operatorname{div}_M X \, d\mathcal{H}^k = - \int_M < H_M, X > \, d\mathcal{H}^k + \int_{\partial M} < X, \nu > \, d\mathcal{H}^{k-1} \tag{3}$$

where ν is the outward pointing unit vector normal to ∂M and tangential to M. Thus for C^2 submanifolds M such that $\overline{M}$ is a submanifold possibly with boundary of class C^2, the *variational property* of being stationarity in N with respect to the area functional is characterized by the absence of boundary together with the simple *geometric property* of being minimal, i.e. having vanishing mean curvature. This is analogous to the equivalence, for C^2 functions on a domain, of the variational property of being stationary (away from the boundary of the domain) with respect to the Dirichlet energy and the analytic property of being harmonic; it sparks nonetheless a degree of surprise at first sight.

2.2. Some explicit singular and non-singular examples. 1-dimensional minimal submanifolds are geodesics. Among the simplest 2-dimensional embedded examples in $\mathbb{R}^3$ are the plane, Catenoid, Helicoid and Scherk surfaces which for centuries have been known to be minimal. There are a number of other known explicit examples in $\mathbb{R}^3$, some discovered relatively recently such as the gyroid and the family of Costa surfaces.

Let $\mathbb{S}^m(r) = \{x \in \mathbb{R}^{m+1} \,:\, |x| = r\}$. The Clifford torus $T_{1,1} = \mathbb{S}^1(1/\sqrt{2}) \times \mathbb{S}^1(1/\sqrt{2}) \subset \mathbb{S}^3(1) \subset \mathbb{R}^4$ is an embedded minimal surface in the round 3-sphere $\mathbb{S}^3(1)$. More generally, the product of spheres $T_{p,q} = \mathbb{S}^p\left(\sqrt{\frac{p}{p+q}}\right) \times \mathbb{S}^q\left(\sqrt{\frac{q}{p+q}}\right) \subset \mathbb{S}^{p+q+1}(1) \subset \mathbb{R}^{p+q+2} \equiv \mathbb{R}^{p+1} \times \mathbb{R}^{q+1}$ is a minimal hypersurface of the round sphere $\mathbb{S}^{p+q+1}(1)$ for any pair of positive integers p and q.

It is easy to check that Σ is a minimal submanifold of $\mathbb{S}^n(1)$ if and only if the cone over Σ, given by $C(\Sigma) = \{\lambda x \,:\, x \in \Sigma, \lambda > 0\}$, is a minimal submanifold of $\mathbb{R}^{n+1}$. This fact provides the simplest examples of *singular* stationary submanifolds in Euclidean spaces. Of course $C(\Sigma)$ is smooth if Σ is, and by the preceding assertion if Σ is minimal in $\mathbb{S}^n$, $C(\Sigma)$ satisfies $\delta_X\left(C(\Sigma)\right) = 0$ for every vector field X with compact support $\subset \mathbb{R}^{n+1} \setminus \{0\}$. If $n \geq 2$, then $C(\Sigma)$ is in fact stationary in $\mathbb{R}^{n+1}$ i.e. $\delta_X(C(\Sigma)) = 0$ for every compactly supported vector field X in $\mathbb{R}^{n+1}$ (with X not necessarily vanishing near the origin). Thus, stationarity of $C(\Sigma)$ holds even for deformations that move its singularity at the origin.

There are no known examples of two dimensional stationary varifolds with isolated singularities in $\mathbb{R}^3$. Whether such examples can be ruled out is in fact a major open problem of central importance in the regularity theory of stationary varifolds. For $n \geq 3$, the work of Caffarelli–Hardt–Simon [**CHS84**] gives a general construction of non-conical stationary hypersurfaces with isolated singularities in $\mathbb{R}^{n+1}$, providing a large family of such examples; each of these hypersurfaces has a single singular point and is asymptotic, at the singularity, to a given singular minimal cone with an isolated singularity.

If M is stationary in $\mathbb{R}^n$, then $M \times \mathbb{R}$ is stationary in $\mathbb{R}^{n+1}$. This fact and the aforementioned connection between the stationary cones in a Euclidean space and the minimal submanifolds of the unit sphere show that $\left(C\left(T_{p,q}\right) \times \mathbb{R}\right) \cap \mathbb{S}^{p+q+2}(1)$ is a singular stationary hypersurface of $\mathbb{S}^{p+q+2}(1)$, providing simple examples in compact manifolds of non-totally geodesic singular stationary hypersurfaces.

Another large family of explicit examples of singular stationary varifolds of codimension > 1 is provided by (singular) holomorphic subvarieties of $\mathbb{C}^n$ which are in fact locally area minimizing. In particular, some of these examples contain branch point singularities, i.e. non-immersed points at which the varifold has a planar tangent cone. A simple (real) 2-dimensional example of this type of singularity is the origin in $\{(z, w) : z^2 = w^3\} \subset \mathbb{C} \times \mathbb{C} \cong \mathbb{R}^4$. Branched minimal surfaces also arise as the image of certain almost conformal harmonic maps of two dimensional manifolds, which may be of codimension 1 or higher. A PDE theoretic construction of a large family of n-dimensional branched minimal immersions in $\mathbb{R}^{n+1}$ is given in [**SimWic07**]. This construction has been streamlined and generalised to arbitrary codimension in [**Krum**].

2.3. Varifolds. While the pointwise meaning of the integrand on the right hand side of identity (3) requires the submanifold M to be of class C^2, the integrand on the left hand side of (3) makes sense pointwise when M is of class C^1. Recall that the integral on the left hand side of (3) is the first variation of M with respect to area. A close look at the derivation of the first variation formula (1) shows that it holds under far less regularity requirements on M than C^1; all that is necessary is that M is a $\mathcal{H}^k$ measurable subset with locally finite $\mathcal{H}^k$ measure, and that a (unique) measure-theoretic tangent plane T_xM to M at x exists for $\mathcal{H}^k$ a.e. $x \in M$, i.e. that M is a *k-rectifiable subset* of N. In that case (1) holds with $\mathrm{div}_M X(x)$ given by the same expression (2) for $\mathcal{H}^k$ a.e. $x \in M$ ([**Sim83**]).

A k-rectifiable set can of course be extremely singular and need not correspond to a C^1 submanifold anywhere. In analogy with various PDE contexts, the fact that the first variation carries over to such a singular setting is extremely helpful in many situations including for establishing existence of stationary submanifolds. It is necessary, in fact, for existence theorems for stationary submanifolds and other applications, to consider first variation in an even more general setting: namely, in the space of *rectifiable k-varifolds*. This is the space of ordered pairs $V = (M, \theta)$ where M is a (countably) k-rectifiable set in N and $\theta : M \to \mathbb{R}^+$ is a locally $\mathcal{H}^k$-integrable function on M. The area functional extends to the space of rectifiable varifolds as the mass functional given by $\mathbb{M}(V \llcorner K) = \int_{M\cap K} \theta \, d\mathcal{H}^k$ for each compact set $K \subset N$. (Since k-rectifiable sets and locally $\mathcal{H}^k$ integrable functions on them are defined only up to sets of $\mathcal{H}^k$ measure zero, a rectifiable k-varifold is in fact an equivalence class of pairs (M, θ) as above, where the equivalence relation is $(M, \theta) \sim (M', \theta')$ whenever $\mathcal{H}^k\left((M \setminus M') \cup (M' \setminus M)\right) = 0$ and $\theta(x) = \theta'(x)$ for $\mathcal{H}^k$ a.e. $x \in M \cap M'$.) For X a C^1 vector field on N with compact support, the first variation of the rectifiable varifold $V = (M, \theta)$ with respect to the mass functional in the direction X is defined by $\delta V(X) = \frac{d}{dt}\big|_{t=0} \mathbb{M}(\varphi_{t\,\#} V \llcorner \mathrm{apt}\, X)$ where $\varphi_{t\,\#} V = (\varphi_t(M), \theta \circ \varphi_t^{-1})$ and

φ_t is as in Sect. 2.1. This leads to a first variation formula generalizing (1) which we shall discuss in the Sect. 2.4.

For the purpose of taking limits of sequences of rectifiable varifolds in a way that facilitates closure theorems for stationary and other classes of varifolds, it is important to take an even more general point of view, namely, to embed the rectifiable k-varifolds defined in this way into the space $\mathcal{V}_k(N)$ of non-negative Radon measures on

$$G_k(N) = \{(x, S) \,:\, x \in N, \;\; S \text{ is a } k\text{-dimensional subspace of } T_x N\}$$

(equipped with the metric induced by the usual product metric on $N \times G(k,n)$ where $G(k,n)$ is the space of k dimensional subspaces of $\mathbb{R}^{n+1}$). This is done by assigning, to any given rectifiable varifold $V = (M, \theta)$, the unique Radon measure on $G_k(N)$ which we shall continue to call V, whose action on $C^0_c(G_k(N))$ is given by

$$V(\phi) = \int_M \phi(x, T_x\, M)\theta(x)\, d\mathcal{H}^k(x) \quad \forall \phi \in C^0_c(G_k(N)). \tag{4}$$

Thus in full generality we define a *k-dimensional varifold* (or a *k-varifold* for short) on N simply as a non-negative Radon measure on $G_k(N)$. We equip the space $\mathcal{V}_k(N)$ of k-varifolds on N with the topology of weak convergence of Radon measures on $G_k(N)$. A k-varifold V is rectifiable if (4) holds for some k-rectifiable set M and a locally $\mathcal{H}^k$ integrable function $\theta : M \to \mathbb{R}^+$; V is *integral* if V is rectifiable and the multiplicity function θ takes (positive) integer values $\mathcal{H}^k$ almost everywhere on M. The space of integral k-varifolds on N is denoted by

$$\mathcal{I}V_k(N).$$

Integral varifolds in fact suffice for many applications including the existence theory for stationary submanifolds.

A k-varifold V on N naturally induces a Radon measure on N, called the *weight measure* or the *mass measure* associated with V, denoted $\|V\|$ (μ_V in [**Sim83**]) and defined by

$$\|V\|(A) = V(A \times G(k,n))$$

for $A \subset N$. If V is rectifiable, $\|V\|(K) = \mathbb{M}(V \llcorner K)$ for each compact set $K \subset N$, where $\mathbb{M}$ is the mass functional defined before.

We define the *regular set*

$$\operatorname{reg} V$$

of a k-varifold V on N to be the set of points $x \in \operatorname{spt} \|V\|$ such that there exists $\rho > 0$ with the property that $\operatorname{spt} \|V\| \cap \operatorname{clos} B^N_\rho(x)$ is a smooth, embedded, compact, connected k-dimensional submanifold of $B^N_\rho(x)$ with boundary $\subset \partial\, B^N_\rho(x)$. Here and subsequently, $B^N_\rho(x)$ denotes the open geodesic ball in N with center x and radius ρ, and clos denotes the closure of a set in N.

The *singular set*

$$\operatorname{sing} V$$

is the set of points $x \in \operatorname{spt}\|V\| \cap N$ with $x \notin \operatorname{reg} V$.

2.4. First variation of a varifold and generalized mean curvature. Let $\Gamma_c(N)$ denote the space of C^1 vector fields on N with compact support. The *first variation* of a k-varifold V on N with respect to mass is the linear map $\delta V : \Gamma_c(N) \to \mathbb{R}$ given by

$$\delta V(X) = \left.\frac{d}{dt}\right|_{t=0} \|\varphi_{t\,\#}\, V \llcorner \operatorname{spt} X\|(N) \tag{5}$$

where for a given vector field $X \in \Gamma_c(N)$, $\varphi_t(x) = \varphi(x,t)$ with $\varphi : N \times (-\epsilon, \epsilon) \to N$ any map as in Sect. 2.1, and $\varphi_{t\,\#}$ is the mapping induced on k-varifolds on N by φ_t (see [**Sim83**, Sect. 39]). By the same calculation as in the C^1 submanifold case ([**Sim83**, Sect. 9]), we then have the *first variation formula* which says that

$$\delta V(X) = \int_{G_k(N)} \operatorname{div}_S X(x)\, dV(x,S) \tag{6}$$

for every $X \in \Gamma_c(N)$, where for $x \in N$ and S a k-dimensional subspace of $T_x N$, $\operatorname{div}_S X = \sum_{j=1}^{k} < \tau_j, \nabla_{\tau_j} X >$ where $\{\tau_1, \ldots, \tau_k\}$ is any orthonormal basis for S.

V is *stationary in* N (with respect to mass) if

$$\delta V(X) = 0 \quad \text{for each} \quad X \in \Gamma_c(N). \tag{7}$$

More generally, we say that a function $H_V \in L^1_{\text{loc}}(\|V\|)$ with $H_V(x) \in T_x N$ for $\|V\|$ almost all $x \in N$ is the *generalized mean curvature vector of* V if

$$\delta V(X) = -\int < H_V(x), X(x) > d\|V\|(x)$$

for all $X \in \Gamma_c(N)$. Thus V is stationary if and only if the generalized mean curvature of V is zero. It is clear that the generalized mean curvature of V, if it exists, is uniquely defined (as an $L^1_{\text{loc}}(\|V\|)$ section of TN). Moreover, from the formula (3) it follows that when V is the multiplicity 1 varifold associated with a C^2 submanifold M with locally finite mass in N, the generalized mean curvature of V agrees with the classical mean curvature of M a.e. on M.

2.5. Compactness and regularity: the Allard theory. Let V be an integral k-varifold with generalized mean curvature $H_V \in L^p_{\text{loc}}(\|V\|)$ for some $p \geq 1$. Recall that by defintion (of integral varifold), there is a k-rectifiable set M together with a positive integer valued, locally $\mathcal{H}^k$-integrable multiplicity function θ on M such that V is given by (4). It is a well-known fact that k-rectifiability of an $\mathcal{H}^k$ measurable set M is equivalent to the fact that M has locally finite $\mathcal{H}^k$ measure and $\mathcal{H}^k$ almost all of M is contained in a countable union of C^1 embedded k-dimensional submanifolds ([**Sim83**]).

By way of regularity, what more can we infer from the assumption that $H_V \in L^p_{\text{loc}}(\|V\|)$? First, the assumption that $H_V \in L^p_{\text{loc}}(\|V\|))$, $p \geq 1$ does not imply any "k-manifold like" property for $\operatorname{spt}\|V\|$. In fact for any $\epsilon > 0$, there are simple examples of multiplicity 1 k-varifolds V with $H_V \in L^p_{\text{loc}}(\|V\|)$ for $p = k - \epsilon$ for which $\operatorname{spt}\|V\|$ is equal to all of the ambient space. For instance one can take a union $\cup_{q\in\mathbb{Q}} \cup_{j\in\mathbb{N}} S_{qj}$ with multiplicity 1, where $\mathbb{Q}$ is the set of rational points in an open subset of $\mathbb{R}^n$, $n \geq 3$, and S_{qj} is a round sphere of dimension $k \geq 2$ centered at q with appropriate radius ρ_{qj} with $\rho_{qj} \to 0$ as $j \to \infty$ for each $q \in \mathbb{Q}$. If $p \geq k$ on the other hand, this behavior does not occur (and in fact we have that $\mathcal{H}^k \llcorner \operatorname{spt}\|V\| \leq \|V\|$ in this case ([**All72**])).

If $p > k$, the monotonicty formula ([**Sim83**]) implies that $\mathcal{H}^k((\operatorname{spt}\|V\| \setminus M) \cup (M \setminus \operatorname{spt}\|V\|)) = 0$, so $\operatorname{spt}\|V\|$ is k-recrifiable. Thus in this case there is a closed rectifiable set representing V. Moreover, in this case the celebrated work of Allard ([**All72**]) gives an embryonic level of regularity for V, and a compactness theorem under uniform mass and L^p mean curvature bounds. The Allard regularity theorem implies that if $V \in \mathcal{IV}_k(N)$ and $H_V \in L^p_{\text{loc}}(\|V\|)$ for some $p > k$, then there is an dense open subset $\operatorname{reg}_1 V$ of $\operatorname{spt}\|V\|$ near each point of which $\operatorname{spt}\|V\|$ is an embedded k-dimensional C^1 submanifold. In fact $\operatorname{reg}_1 V$ is of class $C^{1,1-\frac{k}{p}}$ if $n < p < \infty$ and is of class $C^{1,\alpha}$ for any $\alpha \in (0,1)$ if $p = \infty$. Allard's intregral varifold compactness theorem implies that any family of varifolds $V \in \mathcal{IV}_k(N)$ satisfying locally uniform mass bounds in N and locally uniform $L^p(\|V\|)$ bounds on H_V for some fixed $p > k$ is compact in the topology of varifold convergence (i.e. convergence as Radon measures on $G_k(N)$).

Although no regularity for $\operatorname{spt}\|V\|$ follows from the existence of generalized mean curvature, a remarkably general theorem of Menne ([**Men13**], see [**Sch04**] for earlier partial results) implies that an integral varifold $V = (M, \theta)$ with generalized mean curvature is C^2-rectifiable, i.e. $\mathcal{H}^k$ almost all of M is contained in a countable union of C^2 submanifolds M_j, $j = 1, 2, 3, \ldots$, and moreover, for each j, the classical mean curvature of M_j agrees with H_V at $\mathcal{H}^k$ a.e. point on M_j.

2.6. The Brakke example. The conclusion of Allard's regularity theorem—that $\operatorname{reg}_1 V$ is a dense open subset of $\operatorname{spt}\|V\|$ if $V \in \mathcal{IV}_k(N)$ with $H_V \in L^p_{\text{loc}}(\|V\|)$ for some $p > k$—cannot be improved to $\mathcal{H}^k$ a.e. regularity of $\operatorname{spt}\|V\|$ under the same hypotheses. There is an explicit example due to Brakke ([**Bra78**]) of a 2-dimensional integral varifold W in $\mathbb{R}^3$ with $H_W \in L^\infty_{\text{loc}}(\|W\|)$, $\operatorname{reg}_1 W = \operatorname{reg} W$ and $\mathcal{H}^2(\operatorname{sing} W) > 0$.

A refinement of this example, due to Kolasinski–Menne ([**KolMen15**]), ensures that W has the additional property that $\int_{\operatorname{reg} W \cap K} |A| < \infty$ for each compact set $K \subset \mathbb{R}^3$ but there is a set $\Sigma \subset \operatorname{sing} W$ with $\|W\|(\Sigma) > 0$ such that $\int_{\mathrm{reg}\, W\cap B_\rho(y)} |A|^q = \infty$ for each $q > 1$, $\rho > 0$ and $y \in \Sigma$. Here A denotes the second fundamental form of $\mathrm{reg}\, W$.

3. Second variation and stability

3.1. Second variation formula. Let M be a k-dimensional embedded C^2 minimal submanifold of N with locally finite mass in N. Consider a deformation of M, as described in Sect. 2.1, by a smooth compactly supported vector field X in N. The *second variation* of M with respect to the area functional in the direction X, denoted $\delta^2 M(X, X)$, is defined by

$$\delta^2 M(X,X) = \frac{d^2}{dt^2}\bigg|_{t=0} \mathcal{H}^k(\varphi_t(M))$$

where φ_t, $t \in (-\epsilon, \epsilon)$ is a family of diffeomorphisms of N as described in Sect. 2.1. $\delta^2 M(X, X)$ is a quadratic expression in X, and if $X \llcorner M$ is normal to M with $\operatorname{spt} X \llcorner M$ a compact subset of M, we obtain the following formula:

$$\begin{aligned} &\delta^2 M(X,X) \\ &\quad = \int_M \left(|\nabla^\perp X|^2 - \sum_{i,j=1}^k (< A(\tau_i,\tau_j), X >)^2 \right. \\ &\qquad \left. - \sum_{j=1}^k \mathrm{Riem}^N(\tau_j, X, \tau_j, X) \right) d\mathcal{H}^k \end{aligned} \tag{8}$$

where $\nabla^\perp$ denotes the covariant derivative on the normal vector fields, $\{\tau_1, \ldots, \tau_k\}$ is any local orthonormal frame for M, A denotes the second fundamental form of M and Riem^N denotes the Riemannian curvature of N. If $\delta^2 M(X, X) \geq 0$ for all such compactly supported normal vector fields X, we say that M is *stable*. Stability of M is equivalent to non-negativity of the Dirichlet eigenvalues of a linear elliptic operator, called the Jacobi operator, acting on normal vector fields.

In case M is a two-sided hypersurface and $X \llcorner M$ is normal to M with compact support, we can write $X \llcorner M = \zeta\nu$ where ζ is a smooth function on M with compact support and ν is choice of continuous unit normal to M. In this case, the above formula (with $\delta^2 M(\zeta\nu, \zeta\nu)$ denoted $\delta^2 M(\zeta, \zeta)$) reduces to

$$\delta^2 M(\zeta,\zeta) = \int_M \left(|\nabla^M \zeta|^2 - \left(|A|^2 + \mathrm{Ric}^N(\nu,\nu)\right)\zeta^2 \right) d\mathcal{H}^n$$

where Ric^N is the Ricci curvature of N. Thus stability of M in this case is the requirement that

$$\int_M \left(|\nabla^M \zeta|^2 - \left(|A|^2 + \mathrm{Ric}^N(\nu,\nu)\right)\zeta^2 \right) d\mathcal{H}^n \geq 0 \tag{9}$$

for all $\zeta \in C^\infty_c(M)$. Furthermore, the Jacobi operator $\mathcal{L}$ in this case acts on smooth functions on M and has the simple form

$$\mathcal{L}\,\zeta = -(\Delta_M\,\zeta + \left(|A|^2 + \mathrm{Ric}^N\,(\nu,\nu)\right)\zeta \tag{10}$$

where Δ_M is the Laplace–Beltrami operator on M.

3.2. Schoen–Simon–Yau and Schoen–Simon a priori estimates. Let M be an n-dimensional embedded stable minimal hypersurface $M \subset N$ with locally finite mass in N. An immediate consequence of the stability inequality (9) above is that if the singular set $\mathrm{sing}\, M = \overline{M} \setminus M$ is sufficiently small ($\mathcal{H}^{n-2}\,(\mathrm{sing}\, M \cap K) < \infty$ for each compact $K \subset N$ suffices), then the norm of the second fundamental form A of M satisfies a mass-dependent uniform interior L^2 bound. That is to say, for each open set $U \subset N$ with $\overline{U}$ compact, each compact set $K \subset U$ and each $\Lambda > 0$, there is a constant $C = C(K, U, \Lambda)$ such that if $\mathcal{H}^n\,(M \cap U) \leq \Lambda$, then

$$\int_{K\cap M} |A|^2 \leq C.$$

The fundamental work of Schoen–Simon–Yau ([**SSY75**]) and of Schoen–Simon ([**SchSim81**]) improved this, subject to further hypotheses, to an interior *pointwise* curvature bound:

$$\sup\nolimits_{X\in M\cap K} |A|(X) \leq C.$$

The Schoen–Simon–Yau result ([**SSY75**, Theorem 3]) holds in dimensions $n \leq 5$ for properly immersed M. The Schoen–Simon curvature estimate ([**SchSim81**, Theorem 1]) is for embedded M with a small singular set, and is valid either when $n \leq 6$ or (if $n \geq 7$) on regions of M that are sufficiently flat in a certain weak sense. In fact for any n, given an appropriate "area-like" functional F on hypersurfaces of an open ball $B^{n+1}_\rho(0)$ in the Eucildean space $\mathbb{R}^{n+1}$ and a number $\Lambda > 0$, the Schoen–Simon theorem says that the above pointwise estimate is valid with $K = \overline{B^{n+1}_{\rho/2}(0)}$ and $C = C(n, \rho, \Lambda, F)$ for C^2 embedded hypersurfaces M of $B^{n+1}_\rho(0)$ with $0 \in \overline{M}$ provided that M is stationary and stable with respect to F, M is Hausdorff close to a hyperplane, $\mathcal{H}^n(M) \leq \Lambda\rho^n$ and that

$$\dim_{\mathcal{H}}(\,\mathrm{sing}\, M) \leq n - 7. \tag{11}$$

In particular, given any $X_0 \in N$, F may be taken to be the pull back, to the tangent space $T_{X_0}\,N \approx \mathbb{R}^{n+1}$ via the exponential map, of the Riemannian area functional on hypersurfaces in a sufficiently small neighborhood of X_0 in N. In that case, it is part of the conclusion that $\overline{M}$ is embedded near the origin.

The curvature estimate of [**SSY75**] implies strong compactness, i.e. compactness with respect C^m_{loc} convergence for any m, of any locally uniformly mass bounded family of properly immersed stable minimal hypersurfaces of dimension $n \leq 5$. This follows from general PDE arguments. In dimension $n = 6$, by the curvature estimate of [**SchSim81**], this continues to hold for properly embedded hypersurfaces. In dimensions $n \geq 7$ such strong compactness fails ([**HS85**]), but the curvature estimate of [**SchSim81**], which is conditional upon a flatness assumption as described above, implies by standard Geometric Measure Theory principles compactness with respect to the varifold topology of any locally uniformly mass bounded family of stable minimal hypersurfaces satisfying (11).

REMARK 3.1. The Schoen–Simon curvature estimate in fact holds subject to a weaker a priori size bound on the singular set than (11). However all of the key difficulties in its proof are present under hypothesis (11); moreover, the recent work (to be described in the remainder of this article) establishing a general codimension 1 varifold theory—which makes no *size* hypothesis on the singular set and gives, subject to appropriate variational and *structural* hypotheses, the sharp conclusion that the "genuine" singular set has Hausdorff dimension $\leq n - 7$—depends on the Schoen–Simon estimate but only in the form where the singular set hypothesis is (11).

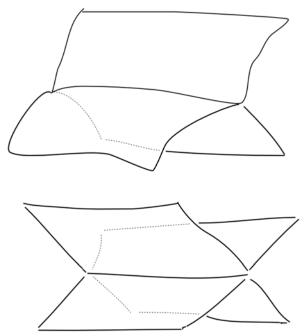
Thus the Schoen–Simon estimate can be viewed as providing a key first step—an a priori estimate subject to the sharp size bound (11) on the singular set—of a general regularity and compactness theory for stable varifolds which uses the a priori estimate to remove the assumption on the size of the singular set altogether. Seen in this light, the overall approach to the regularity theory for codimension 1 stable varifolds parallels a well known strategy in PDE theory of proving regularity of solutions based on a priori estimates for smooth solutions. In the PDE setting the argument usually involves a regularization scheme to approximate the (a priori singular) solution and an application of priori estimates to the approximations to pass to their limit. In the geometric setting there is no direct approximation of the varifold by stable hypersurfaces with small singular sets; however, the proof does involve a certain indirect approximation via a higher multiplicity linearization scheme. The Schoen–Simon estimate is a key ingredient in establishing this linear theory. See [**Wic14c**, Sect. 7] for a discussion of its role in the linear theory corresponding to stable stationary (i.e. zero mean curvature) varifolds, and [**BelWic18**, Sect. 4] for the extension of the procedure to the CMC setting.

4. Special structures: Classical and touching singularities

In the remainder of this article we shall describe the series of recent works [**Wic14a**], [**BelWic18**], [**BCW18**] that has lead to a general, sharp regularity and compactness theory that extends the Schoen–Simon–Yau and Schoen–Simon theories. This work concerns a class of codimension 1 integral n-varifolds with generalised mean curvature locally in L^p for some $p > n$. The varifolds in this class are required to satisfy two structural conditions and appropriate variational hypotheses on the regular parts. Each of the two structural conditions involves a special type of singularity (called respectively a classical singularity or a touching singularity) that is formed when *regular* pieces of the varifold come together in a regular fashion.

DEFINITION 4.1. A *classical singularity* of a set $Y \subset N$ is a point $p \in Y$ such that for some $\sigma > 0$ and some $\alpha \in (0,1)$, $Y \cap B^N_\sigma(p)$ is the union of three or more embedded $C^{1,\alpha}$ hypersurfaces-with-boundary meeting only along their common boundary Γ, with $p \in \Gamma$ and with at least one pair meeting transversely (along Γ).

A classical singularity of a varifold V in N is defined to be a classical singularity of $\operatorname{spt} \|V\|$; we shall denote by $\operatorname{sing}_C V$ the set of all classical singularities of V.



DEFINITION 4.2. A *touching singularity* of a set $Y \subset N$ is a point $p \in Y$ such that p is not a classical singularity of Y and there exist $\sigma > 0$ and $\alpha \in (0,1)$ such that $Y \cap B^N_\sigma(p) = M_1 \cup M_2$ where M_1, M_2 are distinct embedded $C^{1,\alpha}$ hypersurfaces of $B^N_\sigma(p)$ with $\partial M_j \cap B^N_\sigma(p) = \emptyset$ and $p \in M_1 \cap M_2$.

If p is a touching singularity of Y and σ, M_1, M_2 are as above, then the *coincidence set of* Y *in* $B^N_\sigma(p)$ is the set $M_1 \cap M_2$.

A touching singularity of a varifold V in N is defined to be a touching singularity of $\operatorname{spt} \|V\|$; we shall denote by $\operatorname{sing}_T V$ the set of all touching singularities of V. For a varifold V and for $p \in \operatorname{sing}_T V$, the coincidence set of V in a ball $B^N_\sigma(p)$ is the coincidence set of $\operatorname{spt} \|V\|$ in $B^N_\sigma(p)$.

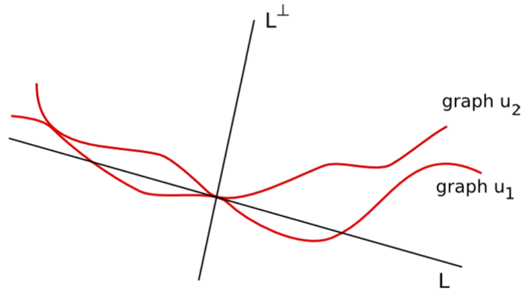
REMARK 4.1 (*Graph structure near a touching singularity*). If p is a touching singularity of a set $Y \subset \mathbb{R}^{n+1}$, then each of the two distinct $C^{1,\alpha}$-hypersurfaces M_1, M_2 corresponding to p (as in Definition 4.2) are tangential to each other at p (because p is not a classical singularity). Let L be the common tangent plane to M_1 and M_2 at p. Upon possibly choosing a smaller σ we then have that

$$Y \cap B^N_\sigma(p) = (\operatorname{graph} u_1 \cup \operatorname{graph} u_2) \cap B^N_\sigma(p)$$

for functions

$$u_1, u_2 \,:\, \left(B^{n+1}_\sigma(p) \cap L\right) \to L^\perp$$

of class $C^{1,\alpha}$ such that $u_1(p) = u_2(p)$ and $Du_1(p) = Du_2(p) = 0$. Note that $u_1 \neq u_2$ since M_1 and M_2 are distinct.



REMARK 4.2 (*terminology*). More generally, for any integer $\ell \geq 2$ one may speak of an *ℓ-fold touching singularity* of a set $Y \subset N$: a point $p \in Y$ is an ℓ-fold touching singularity of Y if there exists $\sigma > 0$ such that

$$Y \cap B^N_\sigma(p) = \cup_{i=1}^{\ell} M_i$$

where M_i are distinct $C^{1,\alpha}$ embedded hypersurfaces of $B^N_\sigma(p)$ with $\partial M_i \cap B_\sigma(p) = \emptyset$, $X \in M_i$ for every $i \in \{1, \ldots, \ell\}$ and with $T_p M_i = T_p M_j$ for any $i, j \in \{1, \ldots, \ell\}$.

Denote by $\text{sing}^{\ell}_T V$ the set of all ℓ-fold touching singularities of a varifold V. For the varifolds in the theorems described below (as well as those in [**BelWic-1**] concerning prescibed-mean-curvature hypersurfaces), the only type of touching singularities on which we need to make any assumption are those in $\text{sing}^2_T V$, and we will in fact a posteriori rule out the occurrence of ℓ-fold touching singularities in V for $\ell \geq 3$. For this reason, we will just refer to a 2-fold touching singularity simply as a "touching singularity" and write $\text{sing}_T V$ for $\text{sing}^2_T V$.

The theory developed in [**Wic14a**], [**BelWic18**], [**BCW18**] gives conditions on a codimension 1 integral n-varifold V that imply that the (genuine) sigular set of V has Hausdorff dimension $\leq n-7$. A key advantage of this theory is that it requires no hypothesis on the singular set beyond verification of the following two structural conditions: (i) absence of classical singularities, and (ii) the $\mathcal{H}^n$-nullity of the coincidence set near every touching singularity. These conditions are in fact necessary and sufficient for the conclusions. In paricular, singular behaviour such as that exhibited by the Brakke example discussed above, where the singularities form a set of positive n-dimensional measure, is not a priori controlled in any way but is ruled out as part of the main conclusions.

5. Regularity and compactness for stable codimension 1 stationary integral varifolds

5.1. Main theorem. The first part of the recent general theory is the work [**Wic14a**] which concerns the case of stationary (i.e. zero mean curvature), stable codimension 1 integral varifolds V. In this case it is easy to see by the Hopf boundary point lemma that if V has no classical singularities then V has no touching singularities.

Subject to the condition that V has no classical singularities, the work [**Wic14a**] shows that the singular set of V has codimension 7.

THEOREM 5.1 ([**Wic14a**]). *If a stationary codimension 1 integral n-varifold V has no classical singularities and if any orientable part of its regular set* $\mathrm{reg}\, V$ *is stable, then its support is smoothly embedded except on a (closed) set of Hausdorff dim* $\leq n-7$. *Furthermore, any uniformly area bounded subfamily of such varifolds is compact in the topology of varifold convergence.*

REMARK 5.1. In verifying the no-classical-singularities assumption, one may leave out any set of zero $(n-1)$-dimensional Hausdorff measure. More precisely, if there is a subset $Z \subset \mathrm{spt}\,\|V\|$ (not assumed closed) with $\mathcal{H}^{n-1}(Z) = 0$ such that no point $Y \in \mathrm{spt}\,\|V\| \setminus Z$ is a classical singularity of V, then V has no classical singularities anywhere. This follows directly from the definition of classical singularity. This is a very useful feature of the theorem, and in one application—namely, the PDE approach to the existence theory for minimal hypersurfaces based on the Allen–Cahn equation (discussed below)—this observation is capitalised on very directly.

We refer the reader to the survey article [**Wic14c**] dedicated to a discussion of the main aspects of the proof of Theorem 5.1 as well as some applications of Theorem 5.1, some questions arising from it and background related to it. One key application (seen completion since publication of [**Wic14c**]) is a PDE theoretic alternative to the Almgren–Pitts existence theory for minimal hypersurfaces. We discuss this briefly below.

5.2. A PDE-based approach to existence of minimal hypersurfaces. This new approach to the existence of minimal hypersurfaces in Riemannian manifolds is the outcome of the combined work of Guaraco, Hutchinson, Tonegawa and the author ([**Gua15**], [**HutTon00**], [**Ton05**], [**TonWic12**], [**Wic14a**]). The basic idea of this approach is to obtain a minimal hypersurface as a weak limit of level sets of a sequence of functions $u_j = u_{\epsilon_j} : N \to \mathbb{R}$, $j = 1, 2, 3, \ldots$ solving the (elliptic) Allen–Cahn equation, i.e. the equation

$$\epsilon \Delta u - \epsilon^{-1} W'(u) = 0 \tag{12}$$

with $\epsilon = \epsilon_j$ where $\epsilon_j \to 0^+$. Here $W : \mathbb{R} \to \mathbb{R}$ is a fixed non-negative double-well potential with precisely two non-degenerate minima at ± 1 with $W(\pm 1) = 0$, one local maximum at a point between -1 and 1 and no other critical points. One standard choice is $W(t) = \frac{1}{4}(t^2 - 1)^2$. The equation (12) is the Euler–Lagrange equation of the Allen–Cahn energy functional

$$E_\epsilon(u) = \int_N \frac{\epsilon}{2} |\nabla u|^2 + \epsilon^{-1} W(u).$$

The first step (chronologically the last step) is an elegant, simple and clean min-max construction due to Guaraco ([**Gua15**]). This construction produces for each small $\epsilon > 0$ a solution u_ϵ to (12) with $\sup_N |u_\epsilon| \leq 1$, Morse index (with respect to E_ϵ) of $u_\epsilon \leq 1$ and satisfying uniform upper and (positive) lower bounds on $E_\epsilon(u_\epsilon)$ independent of ϵ.

The key to this construction is the fact that for each fixed ϵ the functional E_ϵ satisfies the Palais–Smale compactness condition. This enables application of a general mountain-pass theorem to give a critical point u_ϵ of E_ϵ attaining the min-max value over continous paths in $W^{1,2}(N)$ connecting the constant functions 1 and -1, the two global minima of E_ϵ. i.e. a solution u_ϵ to (12) with

$$E_\epsilon(u_\epsilon) = \inf_{\gamma} \max_{t\in[0,1]} E(\gamma(t))$$

where $\gamma : [0,1] \to W^{1,2}(N)$ is continuous with $\gamma(0) = -1$ and $\gamma(1) = 1$. The general theorem also guarantees that Morse index of u_ϵ is ≤ 1. Moreover, by a simple truncation, we can also arrange that $\sup_N |u_\epsilon| \leq 1$. The uniform positive lower bound on $E_\epsilon(u_\epsilon)$ is based on the fact that any continuous path in $W^{1,2}(N)$ connecting the constant functions $-1, 1$ must pass through a function u with $\int_N u = 0$, and that any such u must have high energy. The uniform upper bound on $E_\epsilon(u_\epsilon)$ follows from showing that there is an ϵ-independent upper bound on the energy $E_\epsilon(\gamma(t))$ along a certain explicit path γ in $W^{1,2}(N)$ between the constant functions $-1, 1$; this path is constructed, roughly speaking, by composing signed distance to an appropriate sweepout of N by hypersurfaces with a scaled solution to the 1-dimensional Allen–Cahn equation. See [**Gua15**] for details.

As discussed below, the Morse index bound on u_ϵ is the key advantage in this construction compared to the Almgren–Pitts varifold min-max method. In the latter, serious difficulties need to be overcome because of the lack of clear analogues of Palais–Smale condition or Morse index in the limiting process.

To produce a minimal hypersurface from the solutions u_ϵ, the method appeals to several earlier works concerning the limiting behaviour of u_ϵ as $\epsilon \to 0$. The first of these is a result of Hutchinson–Tonegawa ([**HutTon00**]) which implies that along some subsequence $\epsilon_{j'}$, the level sets of $u_{\epsilon_{j'}}$ converge measure-theoretically to a stationary codimension 1 integral varifold V_{ac}; moreover, the Allen–Cahn energy along the subsequence converges, after normalizing by a fixed constant, to the mass $\|V_{ac}\|(N)$. (The Hutchinson–Tonegawa theorem in fact holds for general uniformly bounded, energy-bounded sequences of critical points that are not required to have uniformly bounded Morse index. It is the elliptic analogue of the pioneering work of Ilmanen ([**Ilm93**]) in the parabolic case, and is a generalisation of the earlier work of Modica ([**Mod87**]) and Sternberg ([**Ste88**]) that studied convergence properties of energy minimizers.) In particular, V_{ac} is non-trivial because of the uniform positive lower bound on $E_\epsilon(u_\epsilon)$.

Regularity of V_{ac} is obtained by appealing to Theorem 5.1. The reason why Theorem 5.1 is applicable to V_{ac} are the combined results of [**Ton05**] (with modifications to suit Riemannian ambient spaces as discussed in [**FSV13**], [**Gua15**]) and [**TonWic12**], which can be summarised as follows:

($\star$) for some open set $U \subset N$ and all j, if u_{ϵ_j} is stable in U (with respect to E_{ϵ_j}), then $\operatorname{reg} V_{ac}$ is stable in U ([**Ton05**], [**FSV13**], [**Gua15**]) and moreover, there is a set $Z \subset \operatorname{spt} \|V_{ac}\| \cap U$ with $\dim_{\mathcal{H}}(Z) \leq n-2$ such that V_{ac} has no classical singularities in $U \setminus Z$ ([**TonWic12**]), and hence (see Remark 5.1) V_{ac} has no classical singularities in U.

This set Z should be thought of as the "L^2 concentration set" of the second fundamental form of the level sets of u_{ϵ_j} as $j \to \infty$ (see [**TonWic12**, Sect. 4]). A simple scaling argument gives that $\dim_{\mathcal{H}}(Z) \leq n-2$. It is also not difficult to see that V_{ac} has no classical singularities in $U \setminus Z$ for if a point $Y \in \operatorname{spt} \|V_{ac}\| \setminus Z$ is a classical singularity, then for each sufficiently large j, the (embedded) level sets of u_{ϵ_j} must bend significantly near Y, contradicting non-concentration of the second fundamental form of the level sets at any point in $\operatorname{spt} \|V_{ac}\| \setminus Z$ (see [**TonWic12**] for details). The stability of $M = \operatorname{reg} V_{ac}$ is verified by showing that the stability inequality (9) holds, which in tern is obtained by taking $\varphi = |\nabla u_{\epsilon_j}|\zeta$ in the second variation inequality $\frac{d^2}{dt^2}\Big|_{t=0} E_{\epsilon_j}(u_{\epsilon_j} + t\varphi) = \int_N \epsilon_j |\nabla \varphi|^2 - \epsilon_j^{-1} W''(u_{\epsilon_j})\varphi^2 \geq 0$ and passing to the limit as $j \to \infty$ ([**Ton05**], [**FSV13**], [**Gua15**]).

Now, since the Morse index of u_ϵ constructed in [**Gua15**] is ≤ 1, by a well-known general observation following directly from the definition of Morse index, it follows that for any given open set $U \subset N$, there is a subsequence $\epsilon_{j'} \to 0$ such that $u_{\epsilon_{j'}}$ are all stable in U or in $N \setminus \overline{U}$. By this fact and ($\star$) above, it follows that either (i) for each $y \in \operatorname{spt} \|V_{ac}\|$ there is $\rho_y > 0$ and such that in $B^N_{\rho_y}(y)$, $\operatorname{reg} V_{ac}$ is stable and V_{ac} has no classical singularities or (ii) there is a point $y_0 \in \operatorname{spt} \|V_{ac}\|$ such that in $(N \setminus \{y_0\})$, $\operatorname{reg} V_{ac}$ is stable and V_{ac} has no classical singularities. If (i) holds, it follows from Theorem 5.1 that $\dim_{\mathcal{H}}(\operatorname{sing} V_{ac}) \leq n-7$. If (ii) holds then since $n \geq 2$, it follows that $\operatorname{reg} V_{ac}$ is stable everywhere, and that V_{ac} has no classical singularities anywhere, so again by Theroem 5.1 we see that $\dim_{\mathcal{H}}(\operatorname{sing} V_{ac}) \leq n-7$.

Thus, compared to the Almgren–Pitts approach, in this method the burden of proof shifts more to the regularity side making way for considerable simplicity on the min-max construction part. This PDE method has subsequently been extended to a multi-parameter min-max construction in the work of Gaspar–Guaraco ([**GasGua1**], [**GasGua2**]). Furthermore, the work of Le ([**Le11**], [**Le15**]), Hiesmayr ([**Hie**]) and Gaspar ([**Gas**]) have established upper semi-continuity of the eigenvalues of the Jacobi operator and lower semi-continuity of the Morse index with respect to the measure-theoretic convergence of level sets of u_{ϵ_j} to the minimal hypersurface. This provides Morse index bounds for the limit minimal hypersurface, again more straightforwardly than in the Almgren–Pitts method (for which Morse index bounds have been obtained by Marques–Neves ([**MarNev16**])).

With additional work aimed at better understanding the convergence of the level sets of the Allen–Cahn solutions, one can expect to gain more information about the limit minimal hypersurface and make the existence theory even more PDE theoretic. For the case $n = 2$, this has been achieved in the very recent work of Chodosh–Mantoulidis ([**ChoMan**]). Building on the work of Wang–Wei ([**WanWei**]) and a theorem of Ambrosio–Cabré (resolving the De Giorgi conjecture concerning entire solutions of the Allen–cahn equation on $\mathbb{R}^3$, [**AmbCab00**]), they show $C^{2,\alpha}$ convergence of the level sets if u_{ϵ_j} are stable with respect to the Allen–Cahn energy. This provides an entirely PDE theoretic proof of the existence and regularity of the limit, as well as a number of new results on the multiplicity, Morse index and the area of the limit, at least for certain ambient metrics forming a dense set in C^∞. We mention that still in the case $n = 2$, different but again PDE theoretic min-max approaches giving existence of branched minimal immersions have been developed in the works of Colding–Minicozzi ([**ColMin08-1**], [**ColMin08-2**]), Riviére ([**Riv15**]) and Pigati–Riviére ([**PigRiv17**]), the latter two being for general codimension.

6. A more general regularity and compactness theory: weakly stable codimension 1 CMC integral varifolds

We now begin to discuss a generalization of the regularity theory of [**Wic14a**] (discussed in Sect. 5) to the CMC setting, following [**BelWic18**] and [**BCW18**]. Just as in [**BelWic18**], [**BCW18**], we will restrict attention to the case where the ambient Riemannian manifold is an open subset of the Euclidean space $\mathbb{R}^{n+1}$.

In [**BelWic-1**] the theory has been further extended to general Riemannian ambient spaces and to the case where the mean curvature (of the regular part of the varifold) is prescribed by an arbitrary ambient function g of class $C^{1,\alpha}$. It is natural to address both these generalizations simultaneously, as done in [**BelWic-1**], because of the way the enclosed volume functional on hypersurfaces of a Riemannian manifold N transforms, locally near a point $X_0 \in N$ under the exponential map, to a functional on hypersurfaces of a ball in the tangent space $T_{X_0} N \approx \mathbb{R}^{n+1}$. We shall not discuss this further generalization here except to say that in that generality, when g vanishes on a large set but not everywhere, there are some additional subtleties—of a topological nature—that need to be addressed. All of the key analytic difficulties of the problem are however present in the case when g is constant.

The hypotheses of the main theorems discussed here are guided by the well-known variational characterization of the CMC condition on an oriented hypersurfaces M of class C^2 as *stationarity of M with respect to the hypersurface-area functional for compactly supported deformations that preserve the volume enclosed by M*. This in turn is equivalent to stationarity of M with respect to the functional $\mathcal{J}(M) = \mathcal{A}(M) + \lambda vol(M)$ for some constant λ for arbitrary compactly supported deformations, where $\mathcal{A}$ and vol are the area and enclosed volume functionals. There is a vast literature on CMC hypersurfaces in the classical (i.e. C^2) setting. The varifold regularity theory in the stable zero mean curvature case and its various applications including the PDE approach to the existence theory for minimal hypersurfaces described in the preceding section motivate developing a CMC (and more generally prescribed mean curvature) regularity theory for codimension 1 integral varifolds.

Such a theory would be most useful if it is a local theory, with hypotheses and conclusions that concern a varifold only in a neighborhood of a point. The first difficulty one faces in aiming for a local varifold CMC theory is related to the above variational characterization of the CMC condition. Specifically, the notion of volume preserving deformations does not have a clear meaning for a varifold in a neighborhood of an arbitrary point. For an *oriented* C^1 hypersurface, regardless of whether it bounds a region of the ambient space, this notion makes sense (for instance in case of Euclidean ambient space, the enclosed volume is given by the explicit formula (15)) but the absence of an orientation prevents its extension to varifolds.

Concerning the hypotheses on the singular set, we wish the spirit of the minimal hypersurface theory (discussed in the preceding section) to carry over to the CMC case, i.e. to make no hypothesis on the singular set beyond easy-to-check conditions on specific types of singularities that are formed when regular pieces come together in a regular fashion. This is indeed possible, as it turns out. This aspect of the theory is preferable to an a priori size hypothesis on the entire singular set because the latter may be more difficult to check. A more fundamental reason why it is (more) important in the CMC case is compactness. Even the weakest possible size restriction on the singular set, namely $\mathcal{H}^{n-1}(\text{sing}\, V) = 0$, that would, as it turns out, imply "good" regularity need not be preserved under taking limits. For instance, consider two disjoint unit cylinders with parallel axes that come together to form a singular set of positive $(n-1)$-dimensional measure in the limit. (If we were to allow the possibility that $\mathcal{H}^{n-1}(\text{sing}\, V) > 0$ with no other assumptions but stationarity and stability, then the singular behavior can be quite complicated as branch points may occur. What can be said in full generality concerning the singular set in this case remains a serious open question. Some partial results are obtained in [**Wic**]). The desire to keep the singular set hypotheses to the weakest that would imply both compactness and good regularity, and the presence of higher multiplicity (also necessitated by compactness considerations) combine to create considerable additional difficulties in the CMC case.

Nonetheless, the work [**BelWic18**], [**BCW18**] in the end shows that there is indeed a satisfactory regularity and compactness theory under a surprisingly mild set of hypotheses. On the variational side, the requirement is that stationarity and stability (with respect to area for volume preserving deformations) need only be checked on the orientable portions of the regular parts (respectively, on the $C^{1,\alpha}$ embedded part and on the C^2 immersed part) of the varifold, where these conditions make sense classically. Concerning the singular set, all that is required is that there are no classical singularities (definition 4.1) and that the coincidence set near any touching singularity (defintion 4.2) is $\mathcal{H}^n$-null. Thus each of these hypotheses need only be verified in regions of the varifold where regularity (at least of class $C^{1,\alpha}$ in some form) is known. The only assumption that concerns the varifold in its entirety is that its generalized mean curvature is locally in L^p for some $p > n$, which guarantees, by the Allard theory, that the $C^{1,\alpha}$ regular part is non-empty. Subject to these hypotheses, the work provides sharp regularity conclusion that the support of the varifold is a classical CMC immersion away from a closed set of Hausdorff dimension $\leq n-7$, together with compactness and curvature estimates (analogous to the Schoen–Simon–Yau and Schoen–Simon estimates) under uniform area and mean curvature bounds.

6.1. Definitions and the statements of the main results. To describe this theory more precisely, let $\mathcal{U} \subset \mathbb{R}^{n+1}$ be open, and let V be an integral n-varifold on $\mathcal{U}$. We shall continue to denote by $\|V\|$ the weight measure associated with V, by $\operatorname{reg} V$ the smoothly embedded part of $\operatorname{spt} V$, by $\operatorname{sing} V$ the singular set $(\operatorname{spt}\|V\| \setminus \operatorname{reg} V) \cap \mathcal{U}$, by $\operatorname{sing}_C V$ the set of classical singularities of V and by $\operatorname{sing}_T V$ the set of touching singularities of V.

Definition 6.1 (*C^1-regular set $\operatorname{reg}_1 V$*). Let $\operatorname{reg}_1 V$ to be the set of points $X \in \operatorname{spt}\|V\|$ with the property that there is $\sigma > 0$ such that $\operatorname{spt}\|V\| \cap B^{n+1}_\sigma(X)$ is an embedded hypersurface of $B^{n+1}_\sigma(X)$ of class C^1.

We now precisely state other necessary hypotheses (items labeled **(I)**-**(V)** below) on V together with some comments related to them:

- **(I):** The first variation of V is locally bounded in $\mathcal{U}$ and is absolutely continuous with respect to $\|V\|$, and the generalized mean curvature of V is in $L^p_{\mathrm{loc}}(\|V\|)$ for some $p > n$.

Under the conditions (**I**) the monotonicity formula [**Sim83**, 17.6] holds and implies that the density $\Theta(\|V\|, X) := \lim_{\rho \to 0} \frac{\|V\|(B^{n+1}_\rho(X))}{\omega_n \rho^n}$ exists for every $X \in \mathcal{U}$, is upper-semi-continuous and that $\Theta(\|V\|, X) \geq 1$ for every $X \in \operatorname{spt}\|V\|$. Moreover, Allard's regularity theorem [**All72**] gives an embryonic level of regularity, namely, the existence of a dense open subset of $\operatorname{spt}\|V\|$ in which $\operatorname{spt}\|V\|$ agrees with an embedded C^1 hypersurface (which in fact is of class $C^{1,\alpha}$ where $\alpha = 1 - \frac{n}{p}$ if $p \in (n, \infty)$ and α is any number $\in (0,1)$ if $p = \infty$). This $C^{1,\alpha}$-embedded part of $\operatorname{spt}\|V\|$ coincides with the set $\operatorname{reg}_1 V$ as in Definition 6.1. The density $\Theta(\|V\|, X)$ is a locally constant integer for $X \in \operatorname{reg}_1 V$ ([**BelWic18**, Lemma A1]).

- **(II):** $\operatorname{sing}_C V = \emptyset$, i.e. V has no classical singularities (see Definition 4.1).
- **(III):** For each $X \in \operatorname{sing}_T V$ there exists $\rho > 0$ such that the coincidence set of V in $B^{n+1}_\rho(X)$ (see Definition 4.2) has zero $\mathcal{H}^n$-measure.

Remark 6.1. Note that by a standard covering argument, hypothesis **(III)** implies that

$$\mathcal{H}^n(\operatorname{sing}_T V) = 0.$$

Let us now discuss the first and second variation hypotheses.

Stationarity. By the divergence theorem, in the case that M is a compact embedded C^1 hypersurface that bounds a region, the volume enclosed by M is given by $\frac{1}{n+1}\int_M \vec{x}\cdot\hat{\nu}\, d\mathcal{H}^n$, where $\vec{x} = (x_1, \ldots, x_{n+1})$ and $\hat{\nu}$ is the outward unit normal on M. For M not necessarily a boundary but merely orientable, we take this as the definition of enclosed volume. For a varifold $V \in IV_n(\mathcal{U})$ and open set $\mathcal{O} \subset \mathcal{U} \setminus (\mathrm{spt}\,\|V\| \setminus \mathrm{reg}_1 V)$ such that $\mathrm{spt}\,\|V \llcorner \mathcal{O}\| \subset \mathrm{reg}_1 V$, if $\mathrm{reg}_1 V \cap \mathcal{O}$ is orientable, we define the enclosed volume of $V \llcorner \mathcal{O}$ as

$$\mathrm{vol}_{\mathcal{O}}(V) = \frac{1}{n+1}\int_{\mathrm{reg}_1 V \cap \mathcal{O}} \vec{x}\cdot\hat{\nu}\, d\|V\|, \tag{13}$$

where $\vec{x} = (x_1, \ldots, x_{n+1})$ and $\hat{\nu}$ is a continuous choice of unit normal on $\mathrm{reg}_1 V \cap \mathcal{O}$. Note that this is a signed volume: a change of sign in the choice of the normal induces a change in the sign of the enclosed volume. (Geometrically, for a C^1 embedded hypersurface D of small size, $|\mathrm{vol}(D)|$ is the volume of the cone on D and vertex at the origin).

We can now specify the stationarity assumption that we will need. Given a vector field $X \in C^1_c(\mathcal{O})$ (where, as above, $\mathrm{spt}\,\|V \llcorner \mathcal{O}\| \subset \mathrm{reg}_1 V$ and $\mathrm{reg}_1 V \cap \mathcal{O}$ is orientable), choose $\epsilon > 0$ and a C^1 map $\Psi : (-\epsilon, \epsilon) \times \mathcal{O} \to \mathcal{O}$ such that $\Psi(0, x) = x$ for $x \in \mathcal{O}$, $\Psi(t, x) = x$ for $(t, x) \in (-\epsilon, \epsilon) \times (\mathcal{O} \setminus \mathrm{spt}\, X)$, $\frac{d}{dt}\big|_{t=0} \Psi(t, x) = X(x)$ for $x \in \mathcal{O}$ and $\Psi_t(\cdot) = \Psi(t, \cdot)$ is a diffeomorphism of $\mathcal{O}$ for $t \in (-\epsilon, \epsilon)$. Such a map Ψ is called a variation. The variation Ψ is called volume-preserving if additionally $\mathrm{vol}_{\mathcal{O}}\left((\Psi_t)_\# V\right)$ is constant for $t \in (-\varepsilon, \varepsilon)$. The stationarity condition on V is the requirement that V is a critical point of the area functional for volume-preserving variations, i.e. $\frac{d}{dt}\big|_{t=0} \|(\Psi_t)_\# V\| = 0$ for any $X \in C^1_c(\mathcal{O}; \mathbb{R}^{n+1})$ and any associated variation Ψ that is volume-preserving. A natural condition on X that guarantees the existence of an associated volume-preserving variation is $\int_{\mathrm{reg}_1 V \cap \mathcal{O}} X \cdot \hat{\nu}\, d\|V\| = 0$ (see [**BarDoC84**, Lemma 2.4] the proof of which, notice, only requires C^1-regularity of the hypersurface; note also that multiplicity is constant, by [**BelWic18**, Lemma A1], on each connected component of $\mathrm{reg}_1 V$). As explained in [**BarDoC84**] the first variation depends only on X and not otherwise on the choice of variation Ψ.

Equivalently, we can encode the fixed-enclosed-volume constraint by introducing a Lagrange multiplier [**BarDoC84**]: for $W \in IV_n(\mathcal{U})$ and $\lambda \in \mathbb{R}$, we consider the functional

$$J_{\mathcal{O}}(W) = A_{\mathcal{O}}(W) + \lambda \mathrm{vol}_{\mathcal{O}}(W),$$

where $A_{\mathcal{O}}(W) = \|W\|(\mathcal{O})$, and require that $V \llcorner \mathcal{O}$ is stationary for $J_{\mathcal{O}}$ with respect to arbitrary deformations, i.e. $\frac{d}{dt}\big|_{t=0} J_{\mathcal{O}}((\Psi_t)_\# V) = 0$ for every $X \in C^1_c(\mathcal{O})$. Again the first variation depends only on X. Thus our stationarity assumption is the following:

(IV): Whenever $\mathcal{O} \subset (\mathcal{U} \setminus (\operatorname{spt}\|V\| \setminus \operatorname{reg}_1 V))$ is such that $\operatorname{reg}_1 V \cap \mathcal{O}$ is orientable, there exists an orientation $\hat{\nu}$ on $\operatorname{reg}_1 V \cap \mathcal{O}$ such that

$$\left.\frac{d}{dt}\right|_{t=0} \|(\Psi_t)_\# V\| = 0$$

for every $X \in C^1_c(\mathcal{O})$ with $\int_{\operatorname{reg}_1 V \cap \mathcal{O}} X \cdot \hat{\nu}\, d\|V\| = 0$ and every variation Ψ with $\left.\frac{d}{dt}\right|_{t=0} \Psi_t = X$, or equivalently, there exists $\lambda \in \mathbb{R}$ such that

$$\left.\frac{d}{dt}\right|_{t=0} J_{\mathcal{O}}((\Psi_t)_\# V) = 0$$

for every $X \in C^1_c(\mathcal{O})$ and every variation Ψ with $\left.\frac{d}{dt}\right|_{t=0} \Psi_t = X$.

Discussion. Let us now analyse the local and global consequences of hypothesis ***(IV)***. Since the multiplicity on each connected component of $\operatorname{reg}_1 V$ is constant by [**BelWic18**, Lemma A1], every connected component of $\operatorname{reg}_1 V$ can locally be expressed as a graph of a $C^{1,\alpha}$ function u (over a tangent plane) which, taken with multiplicity 1, is stationary for $J_{\mathcal{O}}$; this yields that u satisfies, in a weak sense, the CMC equation

$$\operatorname{div}\left(\frac{Du}{\sqrt{1+|Du|^2}}\right) = \lambda$$

for a constant λ, where $u \in C^{1,1-\frac{n}{p}}(B^n_R(0))$. Standard elliptic theory yields that u is of class C^∞, and therefore that $\operatorname{reg}_1 V$ is a smooth hypersurface and thus $\operatorname{reg}_1 V = \operatorname{reg} V$. Moreover the equation is equivalent to the condition that $\vec{H} = \lambda\hat{\nu}$, where $\vec{H}$ is the mean curvature of $\operatorname{reg} V$, i. e. $\operatorname{graph} u$ is a smooth CMC hypersurface with scalar mean curvature $h_0 := \lambda$. Note that at this stage the value h_0 of the mean curvature, while constant on each connected component, might still depend on the chosen connected component of $\operatorname{reg}_1 V = \operatorname{reg} V$. Note also that, unless the mean curvature is zero,[4] the fact that the mean curvature vector $\vec{H}$ is parallel implies that *each connected component of* $\operatorname{reg}_1 V$ *is orientable.* (We wish to emphasise that the preceding derivation only requires the local orientability of $\operatorname{reg}_1 V$, which is always true, and that either of the two possible choices of orientation leads to the same conclusion.)

Let us next discuss the presence of distinct connected components. Note that the volume preserving condition, without a preferred orientation for the varifold, is ambiguous when we are dealing with two distinct connected components of $\operatorname{reg}_1 V = \operatorname{reg} V$. However, as we have seen, local considerations imply the existence of a (non-zero) parallel mean curvature vector on $\operatorname{reg} V$ and hence a canonical global orientation. This allows us to choose $\mathcal{O}$ to cover multiple connected components of $\operatorname{reg} V$.

[4]Note that the orientability of each connected component is obtained here as a consequence of the non-vanishing of the mean curvature. However our work covers the case $H = 0$ as well, see Remark 6.9.

We will next show that assumption **(IV)** implies that $\hat{\nu} = +\frac{\vec{H}}{|\vec{H}|}$ and $\hat{\nu} = -\frac{\vec{H}}{|\vec{H}|}$ are the only choices of orientation for which that assumption can possibly hold; moreover, there exists a constant h such that $\vec{H} = h\hat{\nu}$.

By the previous discussion $\mathrm{reg}_1 V = \mathrm{reg}\, V$ and moreover, for the chosen $\hat{\nu}$, on each connected component $\mathcal{R}$ of $\mathrm{reg}_1 V$, $\vec{H} = h_{\mathcal{R}}\hat{\nu}$ for a constant $h_{\mathcal{R}} \in \mathbb{R}$. Now consider a volume-preserving variation (with respect to the chosen orientation $\hat{\nu}$) that is supported on the union of two distinct connected components $\mathcal{R}_1$ and $\mathcal{R}_2$ of $\mathrm{reg}_1 V = \mathrm{reg} V$ and not separately volume-preserving on each of them. As we recalled earlier, such a volume-preserving variation can be induced by any $\zeta \in C^1_c(\mathcal{R}_1 \cup \mathcal{R}_2)$ such that $\int_{\mathcal{R}_1} \zeta\theta_1 d\mathcal{H}^n + \int_{\mathcal{R}_2} \zeta\theta_2 d\mathcal{H}^n = 0$, where $\theta_1, \theta_2 \in \mathbb{N}$ denote respectively the (constant) density on $\mathcal{R}_1$ and on $\mathcal{R}_2$, and such that $\int_{\mathcal{R}_1} \zeta\theta_1 d\mathcal{H}^n \neq 0$, $\int_{\mathcal{R}_2} \zeta\theta_2 d\mathcal{H}^n \neq 0$. Then by [**Sim83**, §16] the first variation of area δV evaluated on the vector field $\zeta\hat{\nu}$ is given by

$$
\begin{aligned}
\delta V(\zeta\hat{\nu}) &= \int \vec{H} \cdot \zeta\hat{\nu}\, d\|V\| \\
&= \int_{\mathcal{R}_1} h_{\mathcal{R}_1}\zeta\theta_1 d\mathcal{H}^n + \int_{\mathcal{R}_2} h_{\mathcal{R}_2}\zeta\theta_2 d\mathcal{H}^n \\
&= h_{\mathcal{R}_1}\int_{\mathcal{R}_1} \zeta\theta_1 d\mathcal{H}^n + h_{\mathcal{R}_2}\int_{\mathcal{R}_2} \zeta\theta_2 d\mathcal{H}^n.
\end{aligned}
\tag{14}
$$

This implies that $h_{\mathcal{R}_1} = h_{\mathcal{R}_2}$ and hence there exists a constant $h \in \mathbb{R}$ such that $\vec{H} = h\hat{\nu}$. Thus the assertion holds.

Stability and stability inequalities. Let us now discuss the stability hypothesis, i.e. non-negativity for the second variation of V with respect to the area functional for volume-preserving deformations. In our theorems, the stability assumption will be made only on the smoothly immersed part of V, which we shall call the "generalized regular set" of V:

DEFINITION 6.2 (*Generalized regular set*). Let $V \in IV_n(\mathcal{U})$. A point $X \in \mathrm{spt}\,\|V\|$ is a generalized regular point if either (i) $X \in \mathrm{reg}\, V$ or (ii) $X \in \mathrm{sing}_T V$ and we may choose *smooth* functions u_1 and u_2 corresponding to X (as in Definition 4.2) such that $u_1 \geq u_2$. The set of generalised regular points will be denoted by gen-reg V.

REMARK 6.2. Under assumption ***(II)*** gen-reg V is open in $\mathrm{spt}\,\|V\|$. By definition gen-reg V can be realised as a smooth immersion in $\mathcal{U}$ of an abstract n-dimensional manifold (possibly with many connected components).

REMARK 6.3. It is important to note the following. Assume ***(I)***, ***(II)*** and ***(IV)***. Locally near any point $X \in$ gen-reg V we have that $\mathrm{spt}\,\|V\|$ is a smooth embedded hypersurface or the union of exactly two smooth embedded hypersurfaces. By Allard's regularity theorem, $\mathrm{reg}\, V$ is dense in $\mathrm{spt}\,\|V\|$ and in particular any $y \in \mathrm{sing}\, V$ is a limit point of $\mathrm{reg}\, V$. Therefore the mean curvature is necessarily constant on each smooth embedded hypersurface describing gen-reg V; in other words gen-reg V is a C^2 CMC immersion. This condition is equivalent [**BarDoC84**, Proposition 2.7] to the fact that the immersion is stationary for the (multiplicity 1) area measure under volume preserving variations. The definition of enclosed volume can be given for any oriented immersion [**BarDoC84**, (2.2)] and we will discuss it in detail after Remark 6.4.

REMARK 6.4 (Maximum principle and the measure of gen-reg $V \setminus \text{reg}\, V$). At any $y \in$ gen-reg $V \setminus \text{reg}\, V$ by definition $\text{spt}\, \|V\|$ is locally given by the union of two smooth CMC hypersurfaces that intersect tangentially at y. Set coordinates such that $y = (0, 0)$ and $u_j : B^n_\sigma(0) \to \mathbb{R}$ are smooth and satisfy the CMC equation and describe $\text{spt}\, \|V\|$ around y, with $u_1 \leq u_2$, $u_1(0) = u_2(0)$ and $Du_1(0) = Du_2(0) = 0$; here each function u_1 and u_2 solves one of the following PDEs

$$\text{div}\left(\frac{Du}{\sqrt{1+|Du|^2}}\right) = +|h| \quad \text{or} \quad \text{div}\left(\frac{Du}{\sqrt{1+|Du|^2}}\right) = -|h|.$$

If the mean curvature is 0 then gen-reg $V \setminus \text{reg}\, V = \emptyset$ by the maximum principle. When $h \neq 0$, analysing case by case and considering the possibilities for the signs on the right-hand side of the equation and writing the PDE for the difference $v = u_1 - u_2$ we conclude, again by the maximum principle, that u_1 must necessarily solve the PDE with $-|h|$ on the right-hand side and u_2 must solve it with $+|h|$ on the right-hand side. This means, in other words, that the mean curvature vector $\vec{H}$ of $\text{graph}\, u_1$ is such that $\vec{H} \cdot \hat{e}_{n+1} < 0$ and the mean curvature vector $\vec{H}$ of $\text{graph}\, u_2$ is such that $\vec{H} \cdot \hat{e}_{n+1} > 0$. Moreover, observing the Hessian of v, there must exist an index $\ell \in \{1, \ldots, n\}$ such that $D^2_{\ell\ell} v(0) \neq 0$ (because of the non-vanishing of the mean curvature). The implicit function theorem then gives that the set $\{D_\ell v = 0\}$ is, locally around 0, an $(n-1)$-dimensional submanifold. The set of points $\{u_1 = u_2\}$ is contained in the set $\{Du_1 = Du_2\}$ by assumption **(II)** and therefore $\{u_1 = u_2\} \subset \{Dv = 0\} \subset \{D_\ell v = 0\}$. This implies in particular the the set gen-reg $V \setminus \text{reg}\, V$ has locally finite $\mathcal{H}^{n-1}$ measure.

Since gen-reg V is a C^2 CMC immersion (possibly with many connected components), it is orientable and is stationary (as an immersion) for volume-preserving variations, where the enclosed volume for an oriented immersion $i : M^n \to \mathcal{U}$ is given by the formula [**BarDoC84**, (2.2)]

$$vol(i) = \frac{1}{n+1} \int_{M^n} \vec{i} \cdot \hat{\nu}\, dM, \tag{15}$$

where dM is the metric induced on M^n by the immersion into $\mathcal{U}$, $\hat{\nu}$ is the unit normal chosen by the orientation and $\vec{i}$ is the vector $(i_1, \ldots, i_{n+1})$. (For the varifold V under study this quantity is equal to

$$\text{vol}_{\mathcal{U} \setminus (\text{sing} V \setminus \text{gen-reg}\, V)}(|\text{reg}\, V|) = \frac{1}{n+1} \int_{\text{reg}\, V} \vec{i} \cdot \hat{\nu}\, d\mathcal{H}^n;$$

note that $\mathcal{H}^n(\text{gen-reg}\, V \setminus \text{reg}\, V) = 0$.) We stress that, when we consider volume-preserving variations of gen-reg V as an immersion, we allow a one-parameter family $\iota_t : M^n \to \mathcal{U}$ of immersions, $t \in (-\varepsilon, \varepsilon)$, with $\iota_0 = \iota$, $\iota_t(x) = \iota_0(x)$ for x outside a fixed compact subset of M and $vol(\iota_t) = vol(\iota_0)$. Such a deformation is not necessarily induced by an ambient vector field in $\mathcal{U}$: in a neighbourhood of a point in gen-reg $V \setminus$ reg V the two touching sheets will generally be moved independently of each other by such variations (while preserving $vol(\iota_t)$). We will not require the stability for all possible volume-preserving variations as an immersion, but only for those induced by an ambient test function; more precisely, we only need to test the stability for variations with initial normal speed given by $\phi\nu$, where ν is the chosen unit normal and ϕ is an arbitrary ambient smooth function compactly supported in $\mathcal{U} \setminus (\text{sing} V \setminus \text{gen-reg}\, V)$ such that $\int_{\text{gen-reg}\, V} \phi \, d\mathcal{H}^n = 0$ (as we will discuss below, the last integral condition is necessary and sufficient for the existence of a volume-preserving variation with initial normal speed $\phi\nu$).

Our stability assumption precisely is as follows:

(V): For V as above and for every $\phi \in C^1_c(\mathcal{U} \setminus (\text{sing} V \setminus \text{gen-reg}\, V))$ that satisfies

$$\int_{\text{gen-reg}\, V} \phi \, d\mathcal{H}^n = 0,$$

let $\iota_t : M^n \to \mathcal{U}$, with $t \in (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$, be a smooth one-parameter family of immersions such that $\frac{\partial}{\partial t}\big|_{t=0} \iota_t = \phi\nu$, $\iota_0(M^n) =$ gen-reg V, $\iota_t = \iota_0$ outside a fixed compact set for every $t \in (-\varepsilon, \varepsilon)$ and $vol(\iota_t) = vol(\iota_0)$ for $t \in (-\varepsilon, \varepsilon)$. Then

$$\left.\frac{d^2}{dt^2}\right|_{t=0} a(\iota_t) \geq 0,$$

where $a(\iota_t) = \int_{M^n} dM_t$ and dM_t is the metric induced on M^n by the immersion i_t.

Discussion. Let us now discuss hypothesis **(V)** and its consequences. First of all we recall some facts from [**BarDoC84**]. Given a C^2 CMC immersion $i : M = M^n \to \mathcal{U}$ and $\zeta \in C^1_c(M)$ such that $\int_M \zeta dg = 0$, where g is the metric induced on M by $\mathcal{U}$ via the immersion, there exists a volume-preserving (normal) variation ι_t whose variation vector $\frac{d}{dt}\big|_{t=0} \iota_t = \zeta\hat{\nu}$, where $\hat{\nu}$ is a (C^1) choice of the unit normal vector on M [**BarDoC84**, Lemma(2.4)]. Using this fact it is straightforward to show that (see [**BarDoC84**, Proposition 2.10]) stability with respect to volume-preserving variations implies the inequality

(16)

$$\int_M |A|^2 \zeta^2 dg \leq \int_M |\nabla \zeta|^2 dg \quad \text{for any } \zeta \in C^1_c(M) \text{ such that } \int_M \zeta dg = 0,$$

where A denotes the second fundamental form on M induced by the immersion and ∇ is the gradient on M. This is called the *weak stability inequality*. The terminology is used to distinguish it from the *strong stability inequality*, i.e. the same inequality for arbitrary $\zeta \in C^1_c(M)$ (that are not required to satisfy the condition $\int_M \zeta dg = 0$). In fact [**BarDoC84**, Proposition(2.10)] shows that, given an immersed C^2 CMC hypersurface, stability with respect to volume-preserving variations and the validity of the weak stability inequality are actually *equivalent*. Let us outline this argument below.

Given an oriented immersion $\imath : M = M^n \to \mathcal{U}$ with constant mean curvature $h_0\hat{\nu}$, consider the functional

$$J(\imath) = \mathscr{a}(\imath) + h_0 \mathscr{vol}(\imath), \tag{17}$$

where $\mathscr{vol}(\imath)$ is the enclosed volume and $\mathscr{a}(\imath)$ is the area defined above. For any one-parameter variation $\imath_t : M \to \mathcal{U}$ with $\imath_0(M^n) = \text{gen-reg}\, V$, $\imath_t = \imath_0$ for all $t \in (-\varepsilon, \varepsilon)$ outside a fixed compact set, and $\mathscr{vol}(\imath_t) = \mathscr{vol}(\imath_0)$ for $t \in (-\varepsilon, \varepsilon)$, we let $f = \hat{\nu} \cdot \left(\frac{d}{dt}\big|_{t=0} \imath_t\right)$. Then writing $J(t) = J(\imath_t)$, $\mathscr{a}(t) = \mathscr{a}(\imath_t)$ and $\mathscr{vol}(t) = \mathscr{vol}(\imath_t)$ we have, by the constancy of the mean curvature and [**BarDoC84**, Proposition(2.7)], that $J'(0) = 0$. Moreover by [**BarDoC84**, Lemma(2.8)] $J''(0)$ *depends only on* f and

$$J''(0) = \int_M (-|A|^2 f^2 + |\nabla f|^2) dg. \tag{18}$$

(See [**BarDoC84**, Appendix] for the computation; the difficulty in the preceding statement is that the same f can be associated to many distinct variations, not necessarily normal variations.) Once this is established, [**BarDoC84**, Proposition(2.7)] completes the proof of the implication "weak stability inequality $\Rightarrow$ stability for $\mathscr{a}$ under volume-preserving variations" as follows: given any volume-preserving variation $\imath_t$, it easily follows that its normal component $f\hat{\nu}$ is such that $\int_M f dM_0 = 0$ and so by the weak stability inequality (taken with $\zeta = f$) we have $J''(0)(f) \geq 0$. On the other hand $J''(0) = \mathscr{a}''(0) + h_0 \mathscr{vol}''(0) = \mathscr{a}''(0)$ because $\imath_t$ preserves $\mathscr{vol}$, and hence $\mathscr{a}''(0) \geq 0$.

In light of this discussion, assumption **(V)** can be equivalently phrased by requiring that the weak stability inequality

$$\int_{\text{gen-reg}\, V} |A|^2 \phi^2 \, d\mathcal{H}^n \leq \int_{\text{gen-reg}\, V} |\nabla \phi|^2 \, d\mathcal{H}^n \tag{19}$$

holds for every $\phi \in C^\infty_c(\mathcal{U} \setminus (\text{sing} V \setminus \text{gen-reg}\, V))$ such that $\int_{\text{gen-reg}\, V} \phi \, d\mathcal{H}^n = 0$, where ∇ stands for the gradient on gen-reg V.

Let us now come to an additional result (Remark 6.6 below) that will be important for our later purposes. First we need the following:

Remark 6.5 (*weak stability inequality $\Rightarrow$ strong stability inequality at smaller scales*). Let $\mathcal{M} = \imath(M)$ be a C^2 immersed *CMC* hypersurface in the open set $\mathcal{U}$ and g is the metric induced on M by the immersion. assume that it satisfies $\int_M |A|^2 \zeta^2 dg \leq \int_M |\nabla \zeta|^2 dg$ for all $\zeta \in C^1_c(M)$ with $\int_M \zeta dg = 0$. Then whenever $\mathcal{V}_1$ and $\mathcal{V}_2$ are disjoint non-empty open subsets of M it must be true that it (at least) one of them the inequality $\int_{\mathcal{V}_j} |A|^2 \zeta^2 dg \leq \int_{\mathcal{V}_j} |\nabla \zeta|^2 dg$ holds for all $\zeta \in C^1_c(\mathcal{V}_j)$ without the zero-average restriction.

Indeed, if this fails in both $\mathcal{V}_1$ and $\mathcal{V}_2$, then we can find ζ_1 and ζ_2 compactly supported respectively in $\mathcal{V}_1$ and $\mathcal{V}_2$ such that $\int_{\mathcal{V}_j} |A|^2\zeta_j^2 > \int_{\mathcal{V}_j} |\nabla\zeta_j|^2$ for $j \in \{1,2\}$. Then we can find constants $c_1, c_2 \in \mathbb{R}$ such that $c_1\zeta_1+c_2\zeta_2$ satisfies the zero average condition on M, i.e. $c_1\int_{\mathcal{V}_1}\zeta_1 = -c_2\int_{\mathcal{V}_2}\zeta_2$, and since the supports of ζ_1 and ζ_2 are disjoint we then have $(c_1\zeta_1+c_2\zeta_2)^2 = c_1^2\zeta_1^2 + c_2^2\zeta_2^2$ from which it follows that $\int_M |A|^2(c_1\zeta_1+c_2\zeta_2)^2 > \int_M |\nabla(c_1\zeta_1+c_2\zeta_2)|^2$, contradicting the assumption. Thus the weak stability inequality implies the stong one in at least one of two arbitrary disjoint subsets.

By using this fact we can see now that the weak stability inequality assumed on $i(M) \subset \mathcal{U}$ with $p \in i(M)$ actually implies that there exists an ambient open ball B around p in which the strong stability holds (i.e. without the restriction of the zero-average on the test function $\zeta \in C_c^1(B \setminus (\text{sing}V \setminus \text{gen-reg}\, V)))$. To see this, consider, for $R > 0$ fixed such that $B_R^{n+1}(p) \subset \mathcal{U}$ and $0 < r < R$, the ball $B_r^{n+1}(p)$ and the annulus $B_R^{n+1}(p) \setminus \overline{B}_r^{n+1}(p)$. By the previous discussion, the strong stability inequality must hold in at least one of the disjoint open sets $\imath^{-1}(B_R^{n+1}(p)\setminus\overline{B}_r^{n+1}(p))$ and $\imath^{-1}(B_r^{n+1}(p))$. We have either (i) for some r the strong stability inequality holds for all $\zeta \in C_c^1\left(B_r^{n+1}(p) \setminus (\text{sing}V \setminus \text{gen-reg}\, V)\right)$ or (ii) the strong stability holds with any $\zeta \in C_c^1\left(B_R^{n+1}(p) \setminus (\text{sing}V \setminus \text{gen-reg}\, V) \setminus \{p\}\right)$. In the latter case the inequality can be shown to hold for an arbitrary ζ supported in $B_R^{n+1}(p) \setminus (\text{sing}V \setminus \text{gen-reg}\, V)$ by a standard capacity argument, since $n \geq 2$. In either case we reach the same conclusion: there is a ball B around p such that the strong stability inequality holds for all $\zeta \in C_c^1(B \setminus (\text{sing}V \setminus \text{gen-reg}\, V))$.

REMARK 6.6 (*assumption* **(V)** $\Rightarrow$ local strong stability for J). Remark 6.5 says that the requirement that a immersed CMC hypersurface is variationally stable (as an immersion) in an open set with respect to volume-preserving variations induced by ambient test functions (which is equivalent, as mentioned earlier, to the validity of the weak stability inequality in the same open set) implies the validity of the strong stability inequality in a neighbourhood of every point and therefore it gives the non-negativity of $J''(0)$ for any variation supported in that neighbourhood (non necessarily volume preserving) that is induced by an ambient test function. Therefore the (geometrically natural) variational stability of a CMC hypersurface for volume-preserving variations implies that the hypersurface is *locally* a stable critical point for the functional J, for the variations that we allowed. The importance of this observation lies in the fact that J is an admissible functional for the validity of the results in [**SchSim81**].

This concludes the discussion on the "CMC stable" assumptions and we are now ready to state the main regularity result.

THEOREM 6.1 (**regularity of (weakly) stable CMC integral varifolds**). *Let $n \geq 2$ and let V be an integral n-varifold on an open set $\mathcal{U} \subset \mathbb{R}^{n+1}$ such that the hypotheses **(I)-(V)** above hold; specifically:*

(1) *the first variation of V with respect to the area functional is locally bounded in $\mathcal{U}$ and is absolutely continuous with respect to $\|V\|$, and the generalized mean curvature $\vec{H}$ of V is in $L^p_{loc}(\|V\|)$ for some $p > n$;*
(2) *$sing_C V = \emptyset$;*
(3) *whenever $X \in sing_T V$ there exists $\rho > 0$ such that the coincidence set of V in $B^{n+1}_\rho(X)$ has zero $\mathcal{H}^n$-measure;*
(4) *if an open set $\mathcal{O} \subset (\mathcal{U} \setminus (spt\,\|V\| \setminus reg_1 V))$ is such that $reg_1 V \cap \mathcal{O}$ is orientable, then, relative to one choice of orientation on $reg_1 V \cap \mathcal{O}$, $V \llcorner \mathcal{O}$ is stationary with respect to the area functional under volume-preserving variations;*[5]
(5) *for every $\phi \in C^\infty_c(\mathcal{U} \setminus (sing\,V \setminus gen\text{-}reg\,V))$ that satisfies $\int_{gen\text{-}reg\,V} \phi \, d\mathcal{H}^n = 0$, gen-reg V is stable with respect to the area functional under (volume-preserving) variations with initial normal speed $\phi\nu$.*[6]

Then $sing\,V \setminus sing_T\,V$ is empty if $n \leq 6$, discrete if $n = 7$ and is a closed set of Hausdorff dimension at most $n - 7$ if $n \geq 8$. Moreover $sing_T\,V \subset gen\text{-}reg\,V$ (see Definition 6.2 above) and $sing_T\,V$ is locally contained in a smooth submanifold of dimension $(n-1)$, and gen-reg V is a smooth CMC immersion.

REMARK 6.7. It follows directly from the definition of classical singularity that the no-classical-singularities assumption (hypothesis 2) is equivalent to the following: there exists a set $Z \subset \operatorname{spt}\|V\|$ with $\mathcal{H}^{n-1}(Z) = 0$ (not assumed closed) such that $\operatorname{sing}_C V \cap (\operatorname{spt}\|V\| \setminus Z) = \emptyset$.

REMARK 6.8. The preceding regularity result is of local nature, so it suffices to prove it locally around any point of $\operatorname{spt}\|V\|$, i.e. taking $\mathcal{U}$ to be a small open ball around any given point in $\operatorname{spt}\|V\|$. On the other hand, in view of Remark 6.6, around any $X \in \operatorname{spt}\|V\|$ we can find a ball such that gen-reg V is, in that ball, a C^2 CMC immersion that is *strongly* stable with respect to the functional $J = \mathcal{a} + h_0 \mathcal{vol}$ (i.e. stable with respect to J for variations induced by arbitrary ambient test functions not necessarily having zero average), where h_0 is the constant value for the scalar mean curvature (implied by assumption 4, see the discussion after ***(IV)*** above).

In view of Remark 6.8 we see that Theorem 6.1 will be implied by the following theorem in which strong stability is assumed. It turns out that, for the proof, we only need to require the stability with respect to variations with initial speed $f\nu$ where f is a *non-negative* ambient test function.

---

[5] Of course locally on $\operatorname{reg}_1 V$ there is always an orientation; by the discussion following **(IV)** all of $\operatorname{reg}_1 V$ is orientable (and smooth) whenever **(IV)** holds.

[6] The fact that gen-reg V is a CMC C^2-immersion (possibly with several connected components) is not an assumption here, it is an immediate consequence of assumption 4, in view of Remark 6.3. Only the stability is an assumption.

THEOREM 6.2 (**regularity for strongly stable CMC integral varifolds**). *Let $n \geq 2$ and let V be an integral n-varifold on an open set $\mathcal{U} \subset \mathbb{R}^{n+1}$ that satisfies the following assumptions:*

(1) *the first variation of V with respect to the area functional is locally bounded and is absolutely continuous with respect to $\|V\|$, and the generalized mean curvature $\vec{H}$ of V is in $L^p(\|V\|)$ for some $p > n$;*
(2) *$\mathrm{sing}_C V = \emptyset$;*
(3) *whenever $X \in \mathrm{sing}_T V$ there exists $\rho > 0$ such that the coincidence set of V in $B^{n+1}_\rho(X)$ has zero $\mathcal{H}^n$-measure;*
(4) *$\mathrm{reg}_1 V = \mathrm{reg}\, V$ and there exists a continuous choice of unit normal $\hat{\nu}$ on $\mathrm{reg}\, V$ and a constant $h \in \mathbb{R}$ such that $\vec{H} = h\hat{\nu}$ everywhere on $\mathrm{reg}\, V$;*
(5) *for each $f \in C^1_c(\mathcal{U} \setminus (\mathrm{sing}V \setminus \text{gen-reg}\, V))$ with $f \geq 0$, gen-reg V is stable with respect to the functional J defined in (17) under variations (as an immersion) with initial normal speed $f\nu$; equivalently,*
$$\int_{\text{gen-reg}\, V} |A|^2 f^2 \, d\mathcal{H}^n \leq \int_{\text{gen-reg}\, V} |\nabla f|^2 \, d\mathcal{H}^n \text{ for all such } f,$$
with notation as in (19).

Then $\mathrm{sing}\, V \setminus \mathrm{sing}_T V$ is empty for $n \leq 6$, discrete for $n = 7$ and for $n \geq 8$ it is a closed set of Hausdorff dimension at most $n - 7$. Moreover $\mathrm{sing}_T V \subset$ gen-reg V, in the sense of Definition 6.2 and $\mathrm{sing}_T V$ is locally contained in a smooth submanifold of dimension $(n-1)$, and gen-reg V is a classical CMC immersion.

REMARK 6.9 (*the case $H = 0$*). For the minimal case ($H = 0$) Theorem 6.1 provides the same result as [**Wic14a**] but with weaker assumptions, namely the fact that H is identically 0 is replaced by the requirements (assumption 1) that $H \in L^p(\|V\|)$ and (assumption 3) that $H = 0$ on $\mathrm{reg}_1 V$. The global vanishing of H, which is an assumption in [**Wic14a**], is for us a conclusion. Assumption 2 becomes redundant in the minimal case, as it follows from the remaining assumptions and from the maximum principle that $\mathrm{sing}_T V = \emptyset$. Moreover the variational stability (assumption 5) only needs to be assumed for volume preserving variations rather than for arbitrary ones; note, to this end, that a complete minimal hypersurface in an open ball is always orientable by [**Sam69**], and the argument extends to the case of a singular set having codimension at least 7 (it will be clear from the proof that this is all that is needed).

The class of varifolds in Theorem 6.1 is moreover compact under mass and mean curvature bounds:

THEOREM 6.3 (**compactness for stable CMC integral varifolds**). *Let $n \geq 2$, $\Lambda, H_0 \in \mathbb{R}$ be constants and let $p > n$. Let $\mathcal{U} \subset \mathbb{R}^{n+1}$ be open, and let $K \subset \mathcal{U}$ be compact. The family of integral n-varifolds V in $\mathcal{U}$ with $K \cap \mathrm{spt}\, \|V\| \neq \emptyset$ and satisfying the hypotheses (1)-(5) of Theorem 6.1 and the*

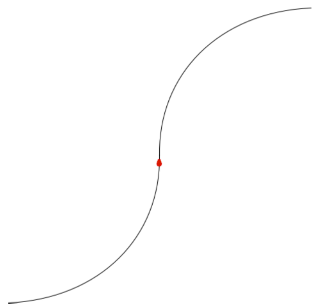


FIGURE 1. The 1-dimensional varifold V depicted here consists of two quarters-of-circle with equal radii joined together in a $C^{1,1}$ fashion, taken with multiplicity 1. Every point of the varifold is in $\text{reg}_1 V$ but V is not of class C^2. Although all of V is orientable, V is not stationarity (for volume preserving deformations) with respect to either choice of orientation; V is stationary only away from the point where the two circular arcs meet

bounds $\|V\|(\mathcal{U}) \leq \Lambda$ *and* $\|\vec{H}_V\|_{L^p(\|V\|)} \leq H_0$ *(where* $\vec{H}_V$ *is the generalized mean curvature of* V *as in assumption 3) is compact in the topology of varifold convergence.*

REMARK 6.10. In the presence of touching singularities one could consider a stronger stability assumption, namely one that allows variations that move the two $C^{1,\alpha}$ hypersurfaces independently at the touching set. Such an assumption would make it possible to employ techniques similar to those used in [**Caf98**] for the so-called obstacle problem, and in particular it would permit the regularity improvement from $C^{1,\alpha}$ to $C^{1,1}$. Note that we are not allowing these variations; we impose only the weaker, classical assumption that stability holds when we already know that the hypersurface is C^2.

6.2. Optimality of the theorems: some examples.

REMARK 6.11. The stationarity assumption must be fulfilled on *any* orientable portion of $\text{reg}_1 V$ for C^2 regularity of $\text{reg}_1 V$ to follow, see Fig. 1.

REMARK 6.12. In the absence of hypothesis 3, the C^2 regularity conclusion of Theorem 6.1 away from a codimension 7 set cannot hold. This is easily seen by the following 1-dimensional example V in $\mathbb{R}^2$ which satisfies hypotheses 1, 2, 4, 5 but not hypothesis 3 of Theorem 6.1, and has one point where it is not C^2 (but is $C^{1,1}$) immersed. (Of course, an n-dimensional example is obtained, with an $(n-1)$-dimensional set where the varifold is not C^2 immersed, by taking the cartesian product of V with $\mathbb{R}^{n-1}$). In this example, V is supported on the set $S \subset \mathbb{R}^2$ defined, with $(x, y) \in \mathbb{R}^2$, by

$$S = \{y \leq 1, x^2 + (y-1)^2 = 1\} \cup \{y \geq -1, x \leq 0, x^2 + (y+1)^2 = 1\}$$

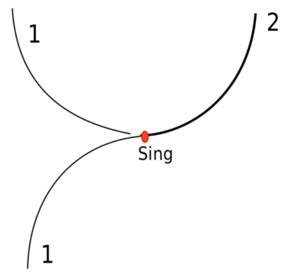


FIGURE 2. Lack of C^2 regularity in the absence of assumption 3, see Remark 6.12

and has multiplicity 2 on the portion $\{(x,y) \in \mathbb{R}^2 : y \leq 1, x \geq 0, x^2 + (y-1)^2 = 1\}$ and multiplicity 1 on the rest. See Fig. 2. Observe that the origin is a touching singularity and the mean curvature is constant on $\text{reg}_1 V = S \setminus \{(0,0)\}$. The stability is also true on $S \setminus \{(0,0)\}$ since we have graphical portions of a CMC curve. However writing the support S at this touching singularity as the union of two graphs on the line $\{y = 0\}$ we are forced to use, for one of the graphs, the function u_1 on $[-1,1]$ that takes the value $\sqrt{1-x^2}+1$ for $x \geq 0$ and the value $\sqrt{1-x^2}-1$ for $x \leq 0$, which is $C^{1,1}$ and enjoys no better regularity (the other graph is the one of the function $u_2 = \sqrt{1-x^2}+1$, that is C^2).

REMARK 6.13. If we drop assumption 2 (absence of classical singularities) we have the examples of two spheres of equal radii crossing along an equator or two transversely intersecting graphical pieces of spheres of equal radii. Both these examples have stable regular parts, and in fact satisfy assumptions 1, 3, 4, 5, but clearly do not satisfy the regularity conclusion. Moreover, once classical singularities are allowed, branch point singularities may develop (as limit points of classical singularities).

REMARK 6.14 (*jumps in the multiplicities at the touching points*). We wish to stress that the stability condition is given only on $\text{spt}\,\|V\|$, i.e. we neglect multiplicities: indeed, with the notation from Definition 6.2 and implicitly restricting to a neighbourhood of $p \in \text{gen-reg}\, V \cap \text{sing}_T V$, we do not generally have that $V = q_1|\text{graph}\, u_1| + q_2|\text{graph}\, u_2|$ for some constants $q_1, q_2 \in \mathbb{N}$, as the following examples show. Consider the 1-dimensional integral varifold V (higher dimensional examples follow by a trivial product with a linear subspace) whose support is given by (see Fig. 3) the set $D \subset \mathbb{R}^2$ defined by (here $(x,y) \in \mathbb{R}^2$)

$$D = \{y \geq -1, x^2 + (y+1)^2 = 1\} \cup \{y \leq 1, x^2 + (y-1)^2 = 1\}$$

with multiplicity 2 on the portions $\{(x,y) \in \mathbb{R}^2 : -1 \leq y \leq 0, x \leq 0, x^2 + (y+1)^2 = 1\}$ and $\{(x,y) \in \mathbb{R}^2 : 0 \leq y \leq 1, x \geq 0, x^2 + (y-1)^2 = 1\}$, and multiplicity 1 on the rest. The support of V agrees with $\text{gen-reg}\, V$, the origin is a touching singularity and all assumptions of Theorem 6.2 are satisfied.

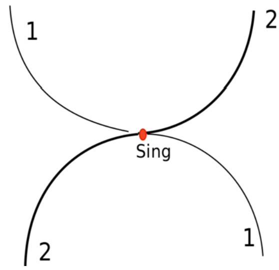


FIGURE 3. A jumps in the multiplicity of each of two C^2 CMC graphs along the touching set, see Remark 6.14

Note that each of the two sheets $\{y = \sqrt{1-x^2} - 1\}$ and $\{y = \sqrt{1-x^2} + 1\}$, taken separately with the assigned multiplicity, is not stationary for the variational problem, as it does not even have generalized mean curvature in L^p, due to the multiplicity jump (the origin belongs to the so-called varifold boundary). For this reason the stability assumption ***(V)*** is stated for spt $\|V\|$.

We can turn the preceding example, which is of local nature, into a global one in the standard sphere S^2, where the support of the varifold is given by the union of four tangential circles of radius $\sqrt{2}/2$, as follows. Let $S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ and

$$C_1 = \left\{x^2 + y^2 = \frac{1}{2}, z = \frac{\sqrt{2}}{2}\right\}, C_2 = \left\{x^2 + z^2 = \frac{1}{2}, y = \frac{\sqrt{2}}{2}\right\},$$

$$C_3 = \left\{x^2 + y^2 = \frac{1}{2}, z = -\frac{\sqrt{2}}{2}\right\}, C_4 = \left\{x^2 + z^2 = \frac{1}{2}, y = -\frac{\sqrt{2}}{2}\right\}.$$

We set multiplicities as follows: on the half-circles $C_1 \cap \{x > 0\}$, $C_2 \cap \{x < 0\}$, $C_3 \cap \{x > 0\}$ and $C_4 \cap \{x < 0\}$ we set the multiplicity equal to 2 and on the remaining four half-circles we set it equal to 1.

REMARK 6.15. In view of Remark 6.14 it is natural to consider the restricted class of varifolds that satisfy the assumptions of Theorem 6.1 with the extra constraint that for $p \in \text{sing}_T V \cap \text{gen-reg}\, V$ the two embedded hypersurfaces going through p have separately constant multiplicity. It will follow from the proof of Theorem 6.3 that this class also enjoys the same compactness result (it is immediate that the regularity theorem holds for this restricted class as well).

REMARK 6.16. The possibility, allowed in the conclusion of Theorems 6.1 and 6.3, that a codimension-7 singular set Σ may be present for $n \geq 7$ is not surprising, in view of the analogous statements for stable minimal hypersurfaces, shown to be optimal by the example of Simons' cone. The recent work [**Irv17**] constructs, based on the construction in [**CHS84**], examples of CMC hypersurfaces with an isolated singularity that are asymptotic to a singular minimal cone. These hypersurfaces are stable when the minimal cone is strictly stable (e.g. Simon's cone), showing the optimality of our conclusion.

6.3. Consequences for Caccioppoli sets. In this subsection we focus on a special class of integral varifolds, namely multiplicity 1 varifolds associated to the reduced boundary of Caccioppoli sets. The latter is a natural class for the variational problem of minimizing boundary area for a fixed enclosed volume, indeed the literature on the subject in the minimizing case is rich and fairly complete, see e.g. [**GMT83**] for the Euclidean case and [**Mor03**] for the extension to Riemannian manifolds. For "stationary Caccioppoli sets" and for "stationary-stable Caccioppoli sets" there is not even a partial local theory available for the variational problem under consideration and the notion of stationarity/stability itself is not immediately clear. In the following we point out that there is a very natural stationarity condition (on a Caccioppoli set) for ambient deformations that fits well with hypotheses 1 and 3 of Theorem 6.1, making Theorem 6.1 also suited to the context of Caccioppoli sets.

REMARK 6.17 (*stationarity for ambient deformations $\Rightarrow$ hypothesis 1*). Any Caccioppoli set E admits a natural notion of enclosed volume, namely $\int \chi_E$, where χ_E denotes the characteristic function of E. In order to make sense of this notion when the enclosed volume is not necessarily finite, one restricts to an arbitrary open set with compact closure. For $\mathcal{O} \subset\subset \mathbb{R}^{n+1}$ we consider the functional (for a certain $\lambda \in \mathbb{R}$)

$$J_{\mathcal{O}}(E) = \|\partial^* E\|(\mathcal{O}) + \lambda \int_{\mathcal{O}} \chi_E, \tag{20}$$

where $\|\partial^* E\|(\mathcal{O})$ denotes the total mass of the boundary measure $\partial^* E = D\chi_E$ in $\mathcal{O}$, and impose the stationarity condition for the varifold $|\partial^* E|$ as follows. For any one-parameter family of deformations ψ_t for $t \in (-\varepsilon, \varepsilon)$ with initial velocity $X \in C^1_c(\mathcal{O}; \mathbb{R}^{n+1})$ we obtain a one-parameter family of Caccioppoli sets $\{E_t = \psi_t(E)\}_{t \in (-\varepsilon, \varepsilon)}$ such that $E_0 = E$ and $E_t \setminus \mathcal{O} = E \setminus \mathcal{O}$ for $t \in (-\varepsilon, \varepsilon)$; we require

$$\left.\frac{d}{dt}\right|_{t=0} J_{\mathcal{O}}(E_t) = 0. \tag{21}$$

This is a stronger hypothesis compared to assumption ***(IV)*** above (we are allowing variations not necessarily supported on reg_1). In this case the stationarity implies automatically that the generalised mean curvature of the multiplicity 1 n-varifold $|\partial^* E|$ associated to the reduced boundary is a constant multiple of the unit normal with no singular part, as we will show now. The first variation of $\int_{\mathcal{O}} \chi_{E_t}$ is equal to $\int_{\partial^* E \cap \mathcal{O}} \nu \cdot X d\mathcal{H}^n \llcorner \partial^* E$ by the divergence theorem on E, where ν denotes the outer normal on $\partial^* E$. The first variation of $\|\partial^* E_t\|(\mathcal{O}) = |\mathcal{H}^n \llcorner (\partial^* E_t \cap \mathcal{O})|$ is, on the other hand, by the first variation formula, given by $\delta_{\partial^* E}(X) = \int_{\partial^* E_t \cap \mathcal{O}} \text{div}_{\partial^* E} X \, d\mathcal{H}^n \llcorner \partial^* E$

and is by definition a continuous linear functional on $C^1_c(\mathcal{O};\mathbb{R}^{n+1})$. The stationarity assumption implies that

$$\delta_{\partial^* E}(X) - \lambda \int \nu \cdot X(d\mathcal{H}^n \llcorner \partial^* E) = 0$$

for every $X \in C^1_c(\mathcal{O};\mathbb{R}^{n+1})$, i.e. $\delta_{\partial^* E}$ and the vector measure $\lambda(d\mathcal{H}^n \llcorner \partial^* E)\nu$ are equal as elements of the dual of $C^1_c(\mathcal{O};\mathbb{R}^{n+1})$. The fact that $(d\mathcal{H}^n \llcorner \partial^* E)\nu$ is a measure implies therefore that $\delta_{\partial^* E}(X) \leq C_E \lambda |X|_{C^0}$ for every $X \in C^1_c(\mathcal{O};\mathbb{R}^{n+1})$, for some constant C_E, in other words the varifold $|\partial^* E|$ has locally bounded first variation in the sense of [**Sim83**, §39] and $\delta_{\partial^* E}$ extends to a continuous linear functional on $X \in C^0_c(\mathcal{O};\mathbb{R}^{n+1})$: this extension necessarily agrees with $\lambda(d\mathcal{H}^n \llcorner \partial^* E)\nu$. The latter is absolutely continuous with respect to the varifold measure $\mathcal{H}^n \llcorner \partial^* E$ and we conclude that the generalized mean curvature of $|\partial^* E|$ in $\mathcal{O}$ is $\vec{H} = \lambda \nu$ for the constant λ, in particular it is L^∞. We point out that the stationarity for $J_{\mathcal{O}}$ for arbitrary ambient deformations is equivalent to the stationarity of the perimeter measure under volume-preserving ambient deformations, see [**Mag12**].

REMARK 6.18 (*hypothesis 1 $\Rightarrow$ hypothesis 3*). When V is the multiplicity 1 varifold naturally associated to the reduced boundary $\partial^* E$ of Caccioppoli set $E \subset \mathbb{R}^{n+1}$, hypothesis 3 in Theorems 6.1 and 6.2 is automatically satisfied in the presence of assumption 1. Indeed, let $\|V\| = \mathcal{H}^n \llcorner \partial^* E$; the assumption that the generalized mean curvature is in $L^p(\|V\|)$ for $p > n$ implies, by the monotonicity formula [**Sim83**, 17.6], that the density $\Theta(\|V\|, x)$ exists everywhere and is ≥ 1 on $\operatorname{spt}\|V\|$. Moreover by [**Sim83**, Theorem 3.15] we have that $\Theta(\|V\|, x) = 0$ for $\mathcal{H}^n$-a.e. $x \in \operatorname{spt}\|V\| \setminus \partial^* E$, so we must have $\mathcal{H}^n(\operatorname{spt}\|V\| \setminus \partial^* E) = 0$. De Giorgi's rectifiability theorem further gives that, for $x \in \partial^* E$, $\Theta(\|V\|, x) = 1$. Since, by the definition of $\operatorname{sing}_T V$, for any $p \in \operatorname{sing}_T V$ we have $\Theta(\|V\|, p) \geq 2$, it follows that hypothesis 3 of Theorem 6.1 holds.

Remarks 6.18 and 6.17 imply immediately the validity of the following corollary of Theorem 6.1. It is worthwhile pointing out that, to our knowledge, proving this corollary alone is not easier than proving the more general Theorem 6.1; the only slight simplification lies in the fact that the jumps in multiplicities on $\text{gen-reg}\, V \setminus \operatorname{reg} V$ described in Remark 6.14 would be prevented, as the multiplicity is necessarily 1 on $\operatorname{reg} V$, but this would not contribute significantly to shortening the arguments.

COROLLARY 6.1 (**stable CMC Caccioppoli sets**). *Let $n \geq 2$ and let $|\partial^* E|$ be the multiplicity 1 integral n-varifold associated to the reduced boundary $\partial^* E$ of a Caccioppoli set $E \subset \mathbb{R}^{n+1}$. Let $\mathcal{O} \subset\subset \mathbb{R}^{n+1}$ and assume that:*

(i) *$\operatorname{sing}_C |\partial^* E| \cap \mathcal{O} = \emptyset$;*
(ii) *the set E is stationary with respect to the functional $J_{\mathcal{O}}$ as in (20), i.e. the condition (21) holds (for ambient deformations ψ_t as specified in (21));*

(iii) *gen-reg* $|\partial^* E| \cap \mathcal{O}$ *is stable as an immersion with respect to the functional* J *in (17), written with* $h_0 = \lambda$ *and* $\mathcal{O}$ *instead of* $\mathcal{U}$*, for all volume-preserving variations with initial speed* $f\nu$*, where* $f \in C^1_c(\mathcal{O} \setminus (sing|\partial^* E| \setminus gen\text{-}reg|\partial^* E|))$ *and* $\int_{gen\text{-}reg\,|\partial^* E|} f d\mathcal{H}^n = 0$.

Then $\mathcal{O} \cap (sing\,|\partial^* E| \setminus sing_T\,|\partial^* E|)$ *is empty for* $n \leq 6$*, closed and discrete for* $n = 7$ *and for* $n \geq 8$ *it is a closed set of Hausdorff dimension at most* $n-7$*. Moreover* $\mathcal{O} \cap sing_T\,|\partial^* E| \subset gen\text{-}reg\,|\partial^* E|$ *and* $sing_T|\partial^* E| \cap \mathcal{O}$ *is locally contained in a smooth submanifold of dimension* $(n-1)$.

Remark 6.19. The regularity conclusion in the preceding corollary is sharp, as shown by the examples constructed in [**Irv17**]. Very recent remarkable work by Delgadino–Maggi [**DelMag17**] classifies Caccioppoli sets in $\mathbb{R}^{n+1}$ with *finite volume* that are stationary with respect to the perimeter for volume-preserving ambient deformations, showing that they are unions of balls. Even in the Euclidean case, a local analogue of this regularity result does not hold under stationarity only, in view of [**Irv17**].

Remark 6.20. At first sight the stability assumption *(iii)* of Corollary 6.1 might seem unsuited to the context of Caccioppoli sets, since we are requiring variations as an immersion of gen-reg $|\partial^* E|$ and, in doing so, we may exit the class of Caccioppoli sets. We wish to point out however that assumption *(iii)* can be rephrased as requiring non-negativity at $t = 0$ of the second variation of the perimeter measure computed along a deformation within the class of Caccioppoli sets that enclose the same volume and that are close to the initial one with respect to the L^1_{loc}-topology. In particular is satisfied under the area-minimizing assumption in Gonzales–Massari–Tamanini [**GMT83**].

Remark 6.21. The notion of stability with respect to the L^1_{loc}-topology on Caccioppoli sets, as discussed in Remark 6.20, leads to the natural question of what can be said in the case when both stationarity and stability hold with respect to the L^1_{loc}-topology (rather than assuming stationarity for ambient deformations, as in Corollary 6.1). This is discussed in the next subsection. Under such variational assumptions a stronger result can be obtained. In fact, subject to this stationarity assumption, a weaker stability assumption, namely, stability of the ambedded part, suffices. See Corollary 6.2.

6.3.1. *Stationarity among* L^1_{loc}*-close Caccioppoli sets.* In Corollary 6.1 we required the stationarity condition for deformations induced by ambient vector fields and the stability of gen-reg $|\partial^* E|$ as an immersion. Depending on the application of the regularity theory, there might be more or less suited stationarity and stability conditions. Much effort has been devoted to the case in which the Caccioppoli set is minimizing for the isoperimetric problem ([**GMT83**], [**Mor03**]): this assumption can be viewed as sitting at one end of the spectrum, where we are allowed to compare with any other Caccioppoli set and we have a minimization property. A slightly weaker notion, of similar flavour, is that of locally minimizing, where the Caccioppoli set minimizes the perimeter measure among all Caccioppoli sets that are close to it in the sense of the L^1_{loc}-topology. At the other end of the spectrum, we might require that stationarity and stability hold for volume-preserving deformations induced by ambient vector fields.

We give, in this subsection, a corollary of our main result that is in between these two ends. In this corollary, stationarity is required with respect to the L^1_{loc}-topology; somewhat surprisingly, under such a stationarity assumption (clearly stronger than the one in Corollary 6.1), we only need a very minimal stability requirement, namely only stability of the smoothly embedded part of $\partial^* E$ for volume-preserving deformations (which are therefore induced by ambient vector fields). Beyond these variational hypotheses, no further condition, structural or otherwise, is needed and we in fact obtain a stronger conclusion than in Corollary 6.1. We here prove this in the Euclidean case; the routine extension to the case of an analytic ambient metric will be included in [**BelWic-1**]. We conjecture that the same result should hold in a smooth Riemannian manifold.

Definition 6.3. Let E be a Caccioppoli set and $\mathcal{O}$ be a bounded open set. A one-sided one-parameter family of deformations of E in $\mathcal{O}$ is a collection $\{E_t\}_{t\in[0,\varepsilon)}$ of Caccioppoli sets, for some $\varepsilon > 0$, such that the curve

$$t \in [0, \varepsilon) \to \chi_{E_t}$$

is continuous in the L^1_{loc}-topology and such that $E_t = E$ in $\mathbb{R}^{n+1} \setminus \mathcal{O}$ for every $t \in [0, \varepsilon)$ and $E_0 = E$.

A one-sided one-parameter volume-preserving family of deformations of E in $\mathcal{O}$ is a one-sided one-parameter family of deformations of E in $\mathcal{O}$ with the additional constraint that $|E \cap \mathcal{O}\,| = |E_t \cap \mathcal{O}\,|$ for every $t \in [0, \varepsilon)$.

Definition 6.4. Let E be a Caccioppoli set in $\mathbb{R}^{n+1}$ and $\mathcal{O}$ be a bounded open set. We say that E is *stationary in $\mathcal{O}$ for the perimeter measure among Caccioppoli sets that enclose the same volume* when the following condition holds:

$$\left.\frac{d}{dt}\right|_{t=0^+} \|D\chi_{E_t} \llcorner \mathcal{O}\| \geq 0$$

for any choice of one-sided one-parameter volume-preserving family of deformations of E in $\mathcal{O}$ as in Definition 6.3 such that the map $t \to \|D\chi_{E_t} \llcorner \mathcal{O}\|$ is differentiable from the right at $t = 0$.

Definition 6.4 imposes a condition that is stronger than stationarity under ambient volume-preserving deformations (i.e. those deformations that are induced by C^1_c ambient vector fields). More generally, for any one-parameter volume-preserving deformation $t \in (-\varepsilon, \varepsilon) \to E_t$, continuous with respect to the L^1_{loc}-topology on E_t and such that $t \to \|D\chi_{E_t} \llcorner \mathcal{O}\|$ is differentiable $t = 0$, then the requirement in Definition 6.4 immediately implies the stationarity condition $\frac{d}{dt}\big|_{t=0} \|D\chi_{E_t} \llcorner \mathcal{O}\| = 0$. The purpose of the requirement in Definition 6.4 is to impose a notion of stationarity in cases where the structure of the Caccioppoli set naturally gives rise only to one-sided deformations.[7]

Corollary 6.2. *Let E be a Caccioppoli set in $\mathbb{R}^{n+1}$ such that E is stationary in a bounded open set $\mathcal{O}$ for the perimeter measure among Caccioppoli sets with the same enclosed volume, in the sense of Definition 6.4. Moreover assume that the smoothly embedded part $\text{reg}\,|\partial^* E|$ is weakly stable, i.e. stable for the perimeter measure under volume-preserving ambient deformations (in the sense of hypothesis **(V)**, but with $\phi \in C^1_c(\mathcal{O} \setminus \text{sing}\,|\partial^* E|)$).*

Then there exists a closed set Σ such that $\dim_{\mathcal{H}} \Sigma \leq n - 7$ and $\partial^* E \cap (\mathcal{O} \setminus \Sigma)$ is a smoothly immersed CMC hypersurface (possibly with several connected components) and with the property that at every point $p \in \overline{\partial^* E} \cap (\mathcal{O} \setminus \Sigma)$ at which $\overline{\partial^* E}$ is not locally embedded there exists a neighbourhood $B^{n+1}_\rho(p)$ in which $\overline{\partial^* E}$ is the union of exactly two smooth complete CMC hypersurfaces in $B^{n+1}_\rho(p)$ that intersect tangentially. Moreover $\text{sing}_T |\partial^* E|$ is a finite union of submanifolds of dimensions between 0 and $n - 2$.

Corollary 6.2 generalizes the work of Gonzales–Massari–Tamanini ([**GMT83**]) that established regularity of boundaries that *minimize* area subject to the fixed enclosed volume constraint.

This result is proved by by reducing it to Corollary 6.1. See [**BelWic18**, Sect. 9]. In particular, it is shown that the stationarity condition in Corollary 6.2 rules out classical singularities and moreover forces gen-reg V to have a more restrictive structure where tangential CMC sheets do not touch along a submanifold of dimension $n - 1$ (whence by analyticity the touching set has locally finite $n - 2$ dimensional Hausdorff measure). This is the reason why the stability requirement can be weakened to only involve $\text{reg}\, V$.

7. Outline of the proof of the regularity and compactness theorems for CMC varifolds

Let H_0 be a non-negative constant and $p > n$. We will denote by $\mathcal{S}_{H_0}$ the class of integral n-varifolds satisfying assumptions 1, 2, 3, 4, 5 of Theorem 6.2 with $\mathcal{U} = B^{n+1}_2(0)$ and $\||H_V|\|_{L^p(\|V\|)} \leq H_0$.

Although many intermediate steps will be required, we can summarise the strategy of the proof of Theorem 6.2 with the following three main theorems, which will all be proved by simultaneous induction.

Theorem 7.1 (Sheeting Theorem). *Let q be a positive integer. There exists $\varepsilon = \varepsilon(n,p,q,H_0) \in (0,1)$ such that if $V \in \mathcal{S}_{H_0}$, $(\omega_n 2^n)^{-1} \|V\|(B^{n+1}_2(0)) <$*

---

[7] Roughly speaking, and as made clear in [**BelWic18**, Sect. 9], when the reduced boundary has a touching singularity or a classical singularity, then the L^1_{loc}-topology allows to "break the singularity apart" in one direction only. From this perspective, a Caccioppoli set with a touching singularity or a classical singularity should be thought of as sitting "at the boundary" of the space of Caccioppoli sets: its deformations are therefore naturally one-sided, and the stationarity condition must be formulated as an inequality.

$q+1/2$, $q-1/2 \le \omega_n^{-1}\|V\|\,(B_1 \times \mathbb{R}^n(0)) < q+1/2$ and

$$\int_{B_1^n(0)\times\mathbb{R}} |x^{n+1}|^2\, d\|V\|(X) + \frac{1}{\varepsilon}\left(\int_{B_1^n(0)\times\mathbb{R}} |h|^p\, d\|V\|(X)\right)^{\frac{1}{p}} < \varepsilon \quad \textit{then}$$

$$V \llcorner \left(B^n_{1/2}(0)\times\mathbb{R}\right) = \sum_{j=1}^{q} |\mathrm{graph}\, u_j|$$

where $u_j \in C^{1,\alpha}\left(B^n_{1/2}(0);\mathbb{R}\right)$ and $u_1 \le u_2 \le \cdots \le u_q$, with

$$\|u_j\|^2_{C^{1,\alpha}(B^n_{1/2}(0))} \le C\left(\int_{B_1^n(0)\times\mathbb{R}} |x^{n+1}|^2\, d\|V\|(X) + \left(\int_{B_1^n(0)\times\mathbb{R}} |H_V|^p\, d\|V\|(X)\right)^{\frac{1}{p}}\right)$$

for some fixed constants $\alpha = \alpha(n,p,q,H_0) \in (0,1/2)$, $C = C(n,p,q,H_0) \in (0,\infty)$ and each $j = 1,2,\ldots,q$.

Theorem 7.2 (Minimum Distance Theorem). *Let $\delta \in (0,1/2)$ and let $\mathbf{C} \in \mathrm{IV}_n(\mathbb{R}^{n+1})$ be a stationary cone in $\mathbb{R}^{n+1}$ such that $\mathrm{spt}\,\|\mathbf{C}\|$ consists of three or more n-dimensional half-hyperplanes meeting along a common $(n-1)$-dimensional subspace. There exists $\varepsilon = \varepsilon(\mathbf{C},\delta,n,p,H_0) \in (0,1)$ such that if $V \in \mathcal{S}_{H_0}$ and $(\omega_n 2^n)^{-1}\|V\|(B_2^{n+1}(0)) \le \Theta\left(\|\mathbf{C}\|,0\right) + \delta$ then*

$$\mathrm{dist}_{\mathcal{H}}\left(\mathrm{spt}\,\|V\| \cap B_1^{n+1}(0), \mathrm{spt}\,\|\mathbf{C}\| \cap B_1^{n+1}(0)\right) > \varepsilon\,.$$

Theorem 7.3 (Higher Regularity Theorem). *Let q be a positive integer and let $V \in \mathcal{S}_{H_0}$ be such that*

$$V \llcorner \left(B^n_{1/2}(0)\times\mathbb{R}\right) = \sum_{j=1}^{q} |\mathrm{graph}\, u_j|$$

where $u_j \in C^{1,\alpha}\left(B^n_{1/2}(0);\mathbb{R}\right)$ for some $\alpha \in (0,1/2)$, and $u_1 \le u_2 \le \cdots \le u_q$. Then

$$\mathrm{spt}\,\|V\| \cap \left(B^n_{1/2}(0)\times\mathbb{R}\right) = \cup_{j=1}^{\tilde{q}} \mathrm{graph}\,\tilde{u}_j$$

for some $\tilde{q} \le q$ and distinct functions $\tilde{u}_j : B_{1/2}(0) \to \mathbb{R}$ with $\tilde{u}_1 \le \tilde{u}_2 \le \cdots \le \tilde{u}_{\tilde{q}}$ where:

(i) *$\tilde{u}_j \in C^2(B^n_{1/2}(0);\mathbb{R})$ and solves the CMC equation on $B^n_{1/2}(0)$ (and hence by elliptic regularity $\tilde{u}_j \in C^\infty(B^n_{1/2}(0);\mathbb{R})$) for each $j \in \{1,2,\ldots,\tilde{q}\}$, and the usual interior derivative estimates hold).*

(ii) *if $\tilde{q} \ge 2$, the graphs of $\tilde{u}_j$ touch at most in pairs, i.e. if there exist $x \in B^n_{1/2}(0)$ and $i \in \{1,2,\ldots \tilde{q}-1\}$ such that $\tilde{u}_i(x) = \tilde{u}_{i+1}(x)$ then $D\tilde{u}_i(x) = D\tilde{u}_{i+1}(x)$ and $\tilde{u}_j(x) \ne \tilde{u}_i(x)$ for all $j \in \{1,2,\ldots \tilde{q}\} \setminus \{i,i+1\}$.*

Thus the Higher Regularity Theorem says that the touching singularities of V are always two-fold and they are in gen-reg V.

7.1. The induction scheme. The proofs of the above three theorems are all carried out simultaneously by an induction argument. Thus, let $q \geq 2$ be an integer, and assume by induction the following:

INDUCTION HYPOTHESES.

(H1) Theorem 7.1 holds with any $q' \in \{1, \ldots, q-1\}$ in place of q.
(H2) Theorem 7.2 holds whenever $\Theta(\|\mathbf{C}\|, 0) \in \{3/2, \ldots, q - 1/2, q\}$.
(H3) Theorem 7.3 holds with any $q' \in \{1, \ldots, q-1\}$ in place of q.

Completion of induction is achieved by carrying out, assuming (H1), (H2), (H3), the following four steps in the order they are listed:

(i) Prove Theorem 7.1;
(ii) prove Theorem 7.2 in case $\Theta(\|\mathbf{C}\|, 0) = q + 1/2$;
(iii) prove Theorem 7.2 in case $\Theta(\|\mathbf{C}\|, 0) = q + 1$;
(iv) prove Theorem 7.3.

The base case $q = 1$ of Theorem 7.1 is a direct consequence of Allard's regularity theorem [**All72**], and the case $\Theta(\|\mathbf{C}\|, 0) = 3/2$ of Theorem 7.2 follows from a theorem of Simon [**Sim93**, Theorem 4].[8] We wish to point out that, within step (i), there is a large "substep" that is still part of the inductive scheme and is needed to develop the necessary "linear theory" for the Sheeting Theorem.

The case $\Theta(\|\mathbf{C}\|, 0) = 2$ of Theorem 7.2 follows by taking $q = 1$ in the argument for step (iii) above, and using the case $q = 1$ of Theorem 7.1 and the case $\Theta(\|\mathbf{C}\|, 0) = 3/2$ of Theorem 7.2 in place of the induction hypotheses (H1), (H2) respectively. In the case $q = 1$ Theorem 7.3 is void and just needs to be replaced with the consequence of Allard's regularity theorem and the CMC assumption to obtain C^2 regularity and the validity of the CMC equation (i.e. it becomes the standard higher regularity for the base case $q = 1$ of Theorem 7.1).

The three theorems above will be combined with the following elementary proposition, used at a number of places in the induction argument (precisely at a number of places in the proof of the sheeting theorem and in the proof of the minimum distance theorem), in order to complete the proof of Theorem 6.1.

PROPOSITION 7.1. *Let $V \in \mathcal{S}_H(\Omega)$, where Ω is an open subset of $B_2^{n+1}(0)$, and for $2 \leq q \in \mathbb{N}$ let $S_q = \{Z : \Theta(\|V\|, Z) \geq q\}$. Assume that (H1), (H2), (H3) are satisfied and assume further that $S_q \cap \Omega = \emptyset$: then*

(i) $(\text{sing}\, V \setminus \text{sing}_T V) \cap \Omega = \emptyset$ if $n \leq 6$, $(\text{sing}\, V \setminus \text{sing}_T V) \cap \Omega$ is discrete if $n = 7$ and $\dim_{\mathcal{H}}((\text{sing}\, V \setminus \text{sing}_T V) \cap \Omega) \leq n - 7$ for $n \geq 8$.

(ii) $\text{sing}_T V \cap \Omega$ is locally contained in a smooth submanifold of dimension $n - 1$.

[8] Incidentally, neither of these requires the stability hypothesis, and they both hold for stationary integral varifolds of arbitrary co-dimension.

Proof of Proposition 7.1. *Statement (i).* This follows using a standard tangent cone analysis, by means of Federer stratification theorem and Simons' stability result, see [**Wic14a**, Sect. 6 Remarks 2 and 3].

Statement (ii). Given two C^2 functions $u_1 \le u_2$ both satisfying the CMC equation (with the same modulus for the mean curvature) and such that the set $T = \{u_1 = u_2\}$ has vanishing $\mathcal{H}^n$-measure and u_1 and u_2 are tangential at all points in T, then the mean curvature vector of graphu_1 points downwards and the mean curvature vector of graphu_2 points upwards. In particular the CMC equations read

$$\operatorname{div}\left(\frac{Du_1}{\sqrt{1+|Du_1|^2}}\right) = -h\,,\ \operatorname{div}\left(\frac{Du_2}{\sqrt{1+|Du_2|^2}}\right) = h$$

for $h > 0$. This follows from the Hopf boundary point lemma applied to the difference $u_1 - u_2$. We want to show first that the set $\{D(u_1 - u_2) = 0\}$ is contained in a submanifold of dimension $\le n-1$. By the condition of opposite signs for the mean curvature, we get that for any $x \in T$ there exists $i \in \{1, \dots, n\}$ such that $D^2_{ii}(u_1 - u_2)$ is non-zero at x and therefore the implicit function theorem gives that $D_i(u_1 - u_2) = 0$ is locally around x a smooth submanifold of dimension $\le n-1$. In particular we have that $\text{sing}_T V$ is contained, locally around any point, in a smooth submanifold of dimension $\le n-1$. □

In the proof of our Sheeting Theorem 7.1 we will need to make use of the following adaptation of [**SchSim81**, Theorem 2]. It is through the application of this result at various places in the argument that the stability assumption predominantly enters our proof.

Theorem 7.4 (Schoen-Simon Sheeting Theorem with codimension 7 singular set). *If in Theorem 7.1 we assume, in place of the hypotheses 1, 2, 3 (of Theorem 6.2), that*

(i) $\text{sing}\,V \setminus \text{sing}_T\,V = \emptyset$ in case $n < 6$, $\text{sing}V \setminus \text{sing}_T\,V$ is discrete in case $n = 7$ or $\dim_{\mathcal{H}}(\text{sing}V \setminus \text{sing}_T V) \le n-7$ in case $n \ge 8$

and that

(ii) $\text{sing}_T\,V \subset \text{gen-reg}\,V$,

and keep all other assumptions, then the conclusion of Theorem 7.1 holds.

Proof of Theorem 7.4. Let $W = |\text{reg}\,V|$, i.e. the multiplicity 1 varifold associated with $\text{reg}\,V$. Then of course $\text{spt}\,\|W\| = \text{spt}\,\|V\|$. It suffices to show that under the hypotheses of Theorem 7.4 that

$$W \llcorner \left(B^n_{1/2}(0) \times \mathbb{R}\right) = \sum_{j=1}^{\tilde{q}} |\text{graph}\,\tilde{u}_j|$$

where $\widetilde{q} \le q$, $\tilde{u}_j \in C^{1,\alpha}\left(B^n_{1/2}(0); \mathbb{R}\right)$ and $\tilde{u}_1 \le \tilde{u}_2 \le \cdots \le \tilde{u}_{\tilde{q}}$, with

$$\|\tilde{u}_j\|^2_{C^{1,\alpha}(B^n_{1/2}(0))} \le C\left(\int_{B^n_1(0)\times\mathbb{R}} |x^{n+1}|^2\, d\|V\|(X) + |h|\right)$$

for some fixed constants $\alpha = \alpha(n,p,q,H_0) \in (0,1/2)$, $C = C(n,p,q,H_0) \in (0,\infty)$ and each $j = 1,2,\ldots,\tilde{q}$.

Since $\text{gen-reg}\, W = \text{gen-reg}\, V$ is a smooth CMC immersion, the first variation formula $\int_W \text{div}_W X d\mathcal{H}^n \llcorner W = -\int_W \vec{H}|_W \cdot X d\mathcal{H}^n \llcorner W$ holds for every $X \in C^1_c(B^{n+1}_2(0) \setminus (\text{sing} V \setminus \text{sing}_T V); \mathbb{R}^{n+1})$, where $\vec{H} = h\hat{\nu}$ (for a constant $h \in \mathbb{R}$) is the (classical) mean curvature of $\text{reg}\, V$. By (i) and (ii) $\text{sing} V \setminus \text{gen-reg}\, V$ $(= \text{sing} V \setminus \text{sing}_T V)$ has codimension 7 or higher, therefore a standard cutoff argument (which requires only $\mathcal{H}^{n-1}(\text{sing} V \setminus \text{sing}_T V) = 0$ and the Euclidean volume growth for W) allows to check that the first variation formula $\int_W \text{div}_W X d\mathcal{H}^n \llcorner W = -\int_W \vec{H}|_W \cdot X d\mathcal{H}^n \llcorner W$ holds for every $X \in C^1_c(B^{n+1}_2(0); \mathbb{R}^{n+1})$, where $\vec{H}$ is the (classical) mean curvature of $\text{reg}\, V$.

The proof of the above can then be obtained by following the arguments in [**SchSim81**] very closely. We cannot obtain Theorem 7.4 directly from [**SchSim81**] because the singular set would include also $\text{sing}_T V$, which can however have dimension $(n-1)$ and this is not allowed by the assumptions in [**SchSim81**]. However, the proof in [**SchSim81**] carries over to our setting. The approximate graph decomposition constructed in [**SchSim81**] will cover $\text{sing}_T V$ as well. In particular the (smooth) ordered graphs of the approximate decomposition will be weakly ordered, i.e. they can touch tangentially but cannot cross. We will now describe how to adapt the arguments from [**SchSim81**].

The first step is to obtain the analogue of [**SchSim81**, Lemma 1] (which, in our case, has $\mu_1 = 0$) for an arbitrary $\varphi \in C^1_c(B^n_1 \times \mathbb{R})$. It is enough to obtain this for $\varphi \geq 0$, since only this case will be used later. For the proof of the inequality we need, following [**SchSim81**, proof of Lemma 1], to use the strong stability inequality with the test function $\varphi(1-(\nu \cdot \nu_0)^2)^{1/2}$. In order to justify its use, we note that ν is defined up to sign on $\text{gen-reg}\, W \cap \text{sing}_T V$ and thus $\varphi(1-(\nu \cdot \nu_0)^2)^{1/2}$ is well-defined on $\text{gen-reg}\, W$, moreover it is locally Lipschitz on $\text{gen-reg}\, W$ because $|\nabla (1-(\nu \cdot \nu_0)^2)^{1/2}| \leq |A|$; therefore it admits an ambient non-negative Lipschitz extension by Kirszbraun theorem; the strong stability inequality extends from C^1 functions compactly supported away from $\text{sing} V \setminus \text{sing}_T V$ to C^1 functions with arbitrary compact support in view of the assumption that $\text{sing} V \setminus \text{sing}_T V$ has codimension 7 (standard cutoff argument, as in [**SchSim81**]); moreover the strong stability inequality extends from C^1 to Lipschitz compactly supported functions, by a simple approximation. The use of the strong stability inequality with the test function $\varphi(1-(\nu \cdot \nu_0)^2)^{1/2}$ leads to [**SchSim81**, (2.1)] and the rest of [**SchSim81**, proof of Lemma 1] can be carried through with routine modifications considering the abstract immersed hypersurface and pulling-back the functions on it.

The arguments in [**SchSim81**, Sects. 3, 4] can now be followed to conclude the proof: we can construct a partial graph decomposition (where the partial CMC graphs are allowed to touch) and then show that the “excess” decays, so that all of spt $\|W\|$ must be covered by the graph decomposition in the end. We stress that the arguments for the excess decay require (compare [**SchSim81**, proof of Lemma 3]):

(i) the use of (the analogue of) [**SchSim81**, Lemma 1] with ambient non-negative compactly supported test functions (of the form $\zeta[\log\left(2^{1/\varepsilon}\lambda^1\overline{\varphi}_0\right)]_+$, with notations as [**SchSim81**, p. 763]);

(ii) the use of (the analogue of) [**SchSim81**, Lemma 1] with a special non-ambient test function, that is compactly supported on a single sheet G_i of a (previously constructed) partial graph decomposition (of the form $\zeta\psi_i$, with the notations of [**SchSim81**, p. 763]).

The first case has been covered in the previous discussion. In the second case, we can use Lemma 1 straight from [**SchSim81**], since we are dealing with a smooth CMC graph on some connected open set Ω and [**SchSim81**, Lemma 1] only needs to assume the validity of the strong stability inequality on G_i, which is true for arbitrary CMC graphs, as explained in [**BelWic18**, Appendix B]; so the second case is also covered.

The arguments described so far lead to the decay result at the origin (for the fine excess E) given by [**SchSim81**, Lemma 4]. We can ensure that the L^2-excess $\hat{E}$ is uniformly small in $B_2(0)$ and in $B_1(X)$ for any choice of $X \in B_{\frac{1}{2}}(0)$. The functional J that we are addressing does not satisfy assumption [**SchSim81**, (1.6)]; on the other hand, we observe that, upon pushing-forward the given varifold V by a translation T_X (so that an arbitrary point X becomes 0) and restricting to the unit ball, we have that the translated varifold $(T_X)_\# V \llcorner B_1(0)$ fulfils the same assumptions for the same functional J (indeed, the notions of being CMC and stable are independent of the choice of coordinates). Therefore we obtain the decay result at an arbitrary $X \in B_{\frac{1}{2}}(0)$ and can complete the proof following [**SchSim81**, p. 775]. □

7.2. Additional difficulties in the CMC case. As mentioned before, the main regularity result, Theorem 6.1, is first reduced to Theorem 6.2 where "strong stability" (i.e. stability with respect to J for uncontrained deformations) of the C^2 immersed part of the varifold V can be assumed. Subsequently, the proof of Theorem 6.2 is divided into three main steps, the Sheeting Theorem (Theorem 7.1), the Minimum Distance Theorem (Theorem 7.2) and the Higher Regularity Theorem (Theorem 7.3), all proved simultaneously by induction. The Higher Regularity Theorem in the case of zero mean curvature (treated in [**Wic14a**]) is an immediate consequence of the Hopf boundary point lemma and the standard elliptic regularity theory. Much of the additional effort needed in [**BelWic18**] for the CMC case goes into the proof of the Higher Regularity Theorem, but let us first discuss the Sheeting Theorem and the Minimum Distance Theorem.

The proofs of the inductive steps of the Sheeting Theorem and the Minimum Distance Theorem follow closely the corresponding argument in [**Wic14a**], but with two key new aspects. One is that they make essential inductive use of the Higher Regularity Theorem. The other is that the conclusion of the Sheeting Theorem yields, initially, a weaker Hölder exponent for the gradient (of the functions defining the sheets) than in [**Wic14a**]. This exponent needs to be improved (as we do in the inductive step for the Higher Regularity Theorem) by independent arguments. The reason for this initially weaker conclusion is that the key excess-decay result needed for the Sheeting Theorem in the present context is obtained for an excess $\hat{E}$ that has, as is usual when the mean curvature H_V is non-zero, e.g. as in [**All72**], an extra lower order additive term (in addition to the L^2 height term) involving H_V. In contrast to the multiplicity 1 setting of [**All72**] however, establishing excess-decay in the higher-multiplicity setting of [**BelWic18**] requires a priori estimates for the varifold that make crucial use of the monotonicity formula. Consequently, the best possible choice for the lower order term in $\hat{E}$ is of the order $\sqrt{\|H_V\|_{L^p(\|V\|)}}$; see the definition of $\hat{E}$ in Theorem 6.1. This limitation arises precisely from the "error term" in the monotonicity formula when $H_V \neq 0$. Hence the excess-decay result we establish will initially only prove the Sheeting Theorem with $C^{1,\alpha}$ sheets for a value of $\alpha < \frac{1}{2}(1 - \frac{n}{p})$. Although we can improve this Hölder exponent by a second run of the argument with the additional knowledge that H_V is constant in the graph region, the best value of α we can get at this stage is still $< \frac{1}{2}$. For this reason one can view the above Sheeting Theorem for the CMC case as serving more the purpose of removing topological complexity of the varifold in the interior than giving regularity; the latter is delegated to the Higher Regularity Theorem.

In [**Wic14a**], since $H_V = 0$, the value of α is irrelevant and higher regularity of the sheets is immediate. This is because by the Hopf boundary point lemma, the distinct sheets making up the support of the varifold are disjoint, and hence the functions defining the individual sheets satisfy separately the minimal surface equation weakly.

In the CMC case, the sheets do not separate in this manner, and our hypotheses in fact allow an a priori optimally large set T of points where the sheets may touch each other; indeed, the only a priori control we have on T is that $\mathcal{H}^n(T) = 0$ (which follows from the structural hypothesis (3) of Theorems 6.1 and 6.2, a sharp condition).

Thus starting from just knowing $C^{1,\alpha}$ regularity, for some $\alpha < 1/2$, of the distinct sheets of the support of the varifold which are allowed to touch on a set T of measure zero, we need to prove their C^2 regularity. The intermediate steps necessary to carry this out involve first improving the Hölder exponent $\alpha \in (0, 1/2)$ in the Sheeting Theorem to some $\alpha \geq \frac{1}{2}$, and then using this improved Hölder regularity of the derivatives to prove $W^{2,2}$ estimates. This then leads to C^2 (in fact $C^{2,\alpha}$) regularity. These steps are detailed in [**BelWic18**, Sect. 7].

We remark that the stronger hypothesis $\mathcal{H}^{n-1}(\operatorname{spt}\|V\| \setminus \operatorname{reg}_1 V) = 0$ would lead to a substantially simpler proof of the Higher Regularity Theorem. This is because then $\mathcal{H}^{n-1}(T) = 0$ and hence by a straightforward cutoff function argument T can be shown to be removable for the PDE (the CMC equation) satisfied, in the complement of T, by the functions defining the sheets. This stronger hypothesis however is undesirable from the point of view of applications; for instance, it is not implied by the general structure theory of Caccioppoli sets, nor does it permit a compactness theorem for the hypersurfaces as pointed out before. In the general case, we still of course show removability of T (for $C^{1,\alpha}$ functions solving the the CMC equation away from T) but the proof is considerably more involved.

We refer the reader to [**Wic14a**] for the proof of the Sheeting and Minimum Distance theorems in the case $H_V = 0$ (see the survey article [**Wic**] for a detailed outline), and to [**BelWic18**] for the necessary modifications in the CMC case. We next describe, in detail, the inductive step of the Higher Regularity Theorem.

8. Proof of the Higher Regularity Theorem

The last part simultaneous proof by induction is the completion of induction for the Higher Regularity Theorem, i.e. step (iv) listed in Sect. 7.1.

In view of this, by assumption we have that V decomposes as q sheets:

$$V \llcorner \left(B^n_{1/2}(0) \times \mathbb{R}\right) = \sum_{j=1}^{q} |\mathrm{graph}\, u_j|,$$

with $u_j \in C^{1,\alpha}\left(B^n_{1/2}(0); \mathbb{R}\right)$ and $u_1 \leq u_2 \leq \cdots \leq u_q$, $\alpha \in (0, \frac{1}{2})$. Whenever $\mathrm{spt}\, \|V\|$ is embedded at $(x, u_j(x))$ then the CMC equation is valid (in weak form) for u_j in a neighbourhood of x, implying C^2 regularity for u_i in that neighbourhood (and even C^∞ by bootstrapping). Moreover keeping in mind that for V as in assumptions of Theorem 7.3 the density θ is everywhere integer-valued, we can inductively assume that $\mathrm{spt}\, \|V\| \cap \{\theta \leq q-1\} \subset \mathrm{gen\text{-}reg}\, V$. Note that we cannot hope that the u_j's are separately smooth and CMC, since the assumptions allow the possibility presented in example 6.14: in such a situation, we would have $u_1 \leq u_2 \leq u_3$ with u_1 and u_3 smooth and CMC but u_2 is only $C^{1,1}$ and it coincides partly with u_1 and partly wth u_3. This is why in order to express the higher regularity result, we need to find a representation of $\mathrm{spt}\, \|V\|$ by neglecting multiplicities.

8.1. Absence of ℓ-fold touching singularities for $\ell \geq 3$. Our first claim is that, for $H \neq 0$ and $X \notin \mathrm{reg}\, V$, whenever there exist (under the assumptions of Theorem 7.3) sheets touching at X then *exactly two sheets* touch at X when we discard multiplicities and moreover *the mean curvature vector points upwards on the top sheet and downwards on the bottom sheet* on the embedded parts. In particular three-fold, four-fold etc. touching singularities are ruled out.

LEMMA 8.1. *Under the assumptions of Theorem 7.3 and assuming the validity of (H1), (H2), (H3), let $X = (x, X^{n+1}) \in \mathrm{spt}\, \|V\|$ be a point of* *density q where spt$\|V\|$ is not embedded. Assume that $H \neq 0$. Then*

(1) *spt*$\|V\| = \operatorname{graph} u_1 \cup \operatorname{graph} u_q$ *and* $X = (x, X^{n+1}) \in \mathit{sing}_T\, V$*;*

(2) *in a neighbourhood of X and away from sing$V$, we have that* $\vec{H} \cdot \hat{e}^{n+1} > 0$ *on* $\operatorname{graph} u_q$ *and* $\vec{H} \cdot \hat{e}^{n+1} < 0$ *on* $\operatorname{graph} u_1$*.*

Remark 8.1. The same argument shows that if $H = 0$ then under the assumptions of Theorem 7.3 any two graphs u_i, u_{i+1} either coincide identically (i.e. $u_i \equiv u_{i+1}$) or have empty intersection (i.e. $u_i < u_{i+1}$). Then the minimal surface PDE very directly implies that each u_j is smooth.

Proof. Any two graphs touching must touch tangentially. Let $\pi : B^n_{1/2}(0) \times \mathbb{R} \to B^n_{1/2}(0)$ be the orthogonal projection, and let $C = \pi\{y \in \operatorname{spt}\|V\| : \theta(y) = q\}$ be the closed set in $B^n_{1/2}(0)$ above which the density of the varifold is q, i.e. the set of $x \in B^n_{1/2}(0)$ such that $u_1(x) = \cdots = u_q(x)$. If $C = B^n_{1/2}(0)$ then V is embedded in the cylinder $B^n_{1/2}(0) \times \mathbb{R}$ contrary to the assumption.

Therefore $C \subsetneq B^n_{1/2}(0)$. Denote by B the ball $B^n_{1/2}(0)$ and consider the open set $(B \setminus C) \times \mathbb{R}$: here we have the validity of the inductive assumption, as the density is $\leq q-1$. Consider $x \in B$ such that $\operatorname{spt}\|V\|$ is embedded at $(x, u_j(x))$ for every j. The set U of such points x is open and dense (by Allard's theorem). Then we can "neglect multiplicities" and define $\tilde{q}(x)$ as the number of points without multiplicity above x. By embededness $\tilde{q}$ is locally constant on U, i.e. constant on each connected component of U. This allows to define $\tilde{u}_1 < \cdots < \tilde{u}_{\tilde{q}}$ on U, with $\tilde{q}$ constant on each connected component of U and $\operatorname{spt}\|V \llcorner (U \times \mathbb{R})\|$ described by the union of the graphs of $\tilde{u}_j$. We will however show that $\tilde{q}$ actually extends to a locally constant function on $B \setminus C$. Indeed, clearly $U \subset B \setminus C$ and on $(B \setminus C) \times \mathbb{R}$ we can use the inductive assumptions, thanks to which we know that $\operatorname{spt}\|V \llcorner (B \setminus C) \times \mathbb{R}\| \subset \text{gen-reg}\, V$: by definition of gen-reg this means that at any point y in $(B \setminus C) \times \mathbb{R}$ where V is not embedded the structure of $\operatorname{spt}\|V\|$ is, in a neighbourhood, exactly the union of two C^2 graphs touching tangentially on a set of 0-measure. Therefore if $\pi(y)$ is on the (topological) boundary of any two connected components of U, the number $\tilde{q}$ cannot change when we pass from one connected component to the other. This allows to extend uniquely the definition of the functions $\tilde{u}_1 \leq \cdots \leq \tilde{u}_{\tilde{q}}$ to $B \setminus C$ so that the union of their graphs is equal to $\operatorname{spt}\|V \llcorner (B \setminus C) \times \mathbb{R}\|$ and moreover each $\tilde{u}_j$ is C^2 on $B \setminus C$ and satisfies the CMC PDE. In particular $\tilde{q}$ is locally constant on $B \setminus C$. The advantage in passing to the functions $\tilde{u}_1 \leq \cdots \leq \tilde{u}_{\tilde{q}}$ is that these are smooth and CMC on $B \setminus C$. Note that necessarily we have $u_1 = \tilde{u}_1$ and $u_q = \tilde{u}_{\tilde{q}}$.

Take any connected component A of $B \setminus C$ on which $\tilde{q} \geq 2$ (such a component exists or else V would be embedded everywhere). We will show that there exists $m \in \{1, \ldots \tilde{q}-1\}$ such that $\tilde{u}_1|_A = \cdots \tilde{u}_m|_A$ and $\tilde{u}_{m+1}|_A = \tilde{u}_{\tilde{q}}|_A$ (m might depend on which connected component has been chosen). For $y \in A$ consider open balls centred at x and contained in A and pick the supremum R of the radii of such balls: then $B_R(y) \subset A$ and there exists $p \in \partial B_R(y) \cap C$. On the ball B_R, by the inductive assumptions, each $\tilde{u}_j$ is smooth and solves the CMC equation

$$\operatorname{div}\left(\frac{D\tilde{u}_j}{\sqrt{1+|D\tilde{u}_j|^2}}\right) = |h| \text{ or } -|h|$$

classically, with the sign on the right-hand side depending on whether the mean curvature vector $\vec{H}$ points upwards or downwards, i.e. whether respectively $\vec{H} \cdot \hat{e}^{n+1} > 0$ or $\vec{H} \cdot \hat{e}^{n+1} < 0$ on $|\text{graph}(\tilde{u}_j|_{B_R(y)})|$. For any couple $\tilde{u}_i, \tilde{u}_j$ (with $i \neq j$) we then consider the possibility that the mean curvatures on these two graphs point in the same direction. Subtracting the PDE for $\tilde{u}_{j+1}$ from the one for $\tilde{u}_j$ we find that $\tilde{u}_j - \tilde{u}_{j+1}$ is 0 on C and solves an elliptic PDE on $B_R(y)$, namely (24) with the right-hand side replaced by 0. Unless $\tilde{u}_j \equiv \tilde{u}_{j+1}$ on $B_R(y)$ we can apply Hopf boundary point lemma, yielding that the derivative of $\tilde{u}_j - \tilde{u}_{j+1}$ along the outward normal to the boundary of the ball at ξ must be strictly positive: this contradicts the vanishing of the gradient of $\tilde{u}_j - \tilde{u}_{j+1}$ on C (tangentiality of the graphs, equivalently absence of classical singularities).

Summing up, we have by now that $\text{spt}\,\|V\|$ in the cylinder $B^n_{1/2}(0) \times \mathbb{R}$ is represented by the union of the graphs of $u_1 = \tilde{u}_1$ and $u_q = \tilde{u}_{\tilde{q}}$ (since every other $\tilde{u}_j$ coincides with one of these two on $B \setminus C$ and on C clearly all graphs coincide). Moreover there exists a non-empty open set A on which $\tilde{u}_1 < \tilde{u}_q$. We can conclude that actually $u_1 = u_q$ only on the set C; on $B^n_{1/2}(0) \setminus C$ we have $u_1 < u_q$. Indeed [**BelWic18**, Lemma A2] forces $V = q_1(x)|\text{graph}(u_1)| + (q - q_1(x))|\text{graph}(u_q)|$, with q_1 integer-valued and locally constant on the embedded part. The set $\{u_1 = u_q\}$ is therefore exactly the set C.

The preceding argument settles part 1 of the lemma and moreover shows that, for any two sheets touching, the mean curvature vectors must point in opposite directions. However, if the top sheet $\tilde{u}_{j+1}$ has mean curvature vector pointing downwards and the bottom sheet $\tilde{u}_j$ has mean curvature pointing upwards, then the CMC equation is solved by $\tilde{u}_j$ with right hand side $|h|$ and by $\tilde{u}_{j+1}$ with right hand side $-|h|$. So the same argument yields that $\tilde{u}_j - \tilde{u}_{j+1}$ is a subsolution of the same elliptic PDE of which it was solution before, more precisely $\tilde{u}_j - \tilde{u}_{j+1}$ solves the PDE (24) with the right-hand side replaced by $-2|h| \int \zeta$. Hopf lemma applies again, so part 2 is proved as well. □

8.2. Associated PDEs. In order to complete the proof of Theorem 7.3 we need to prove C^2 regularity of each sheet across the touching points, i.e. we need to show that these touching points belong to gen-reg V. Note that we only know that the set of points where sheets touch (tangentially) is a set of $\mathcal{H}^n$-measure zero, but such an estimate is too weak to allow any capacity argument for the extension of the CMC equation to each sheet. In other words, with our knowledge so far there could still be (for a single sheet) a singular part of the generalized mean curvature that is concentrated on the touching set (with the mean curvature of the other sheet that annihilates the first). In the following subsections we will rule out this possibility. By Lemma 8.1 we only need to consider the case of two distinct sheets, so for the sequel we will work in the following situation.

HYPOTHESES 8.1. *We consider two sheets $u_1 \leq u_2$, each of them is $C^{1,\alpha}(B_1^n(0))$ for $\alpha < \frac{1}{2}$, they can touch tangentially (the touching singularities $sing_T$ are contained in this touching set) and they cannot cross. Away from the touching set $\{(x, u_1(x)) : u_1(x) = u_2(x)\}$ we have C^2-regularity and constant mean curvature in the classical sense.*

REMARK 8.2. By Lemma 8.1 on any open set $B_\rho^n \times \mathbb{R}$ where there are no singular points, the two sheets have mean curvature vectors pointing in opposite directions, precisely the top sheet u_2 must have mean curvature pointing upwards and the bottom sheet u_1 must have mean curvature pointing downwards. The varifold V is given by integration on the union of the two sheets endowed with an integer multiplicity. The multiplicity is $q \in \mathbb{N}$ on the touching set, it is denoted by $q_1(x)$ on u_1 (on the set where $u_1 < u_2$) and, by [**BelWic18**, Lemma A2], it is $q - q_1(x)$ on u_2 (on the set where $u_1 < u_2$). Moreover $q_1(x)$ is locally constant on the C^2 embedded portions.

The two auxiliary functions that will play a key role in this section are the average $u_a := \frac{u_1+u_2}{2}$ and the semi-difference $v := \frac{u_2-u_1}{2}$ of the two sheets.

PDE for the difference. Consider $v = \frac{u_2-u_1}{2}$, well-defined on $B_1^n(0)$ and assume without loss of generality that $(0,0) \in B_1^n(0) \times \mathbb{R}$ is a touching singularity of the varifold. Then $v(0) = 0$ and $Dv(0) = 0$. The set

$$T = \{x \in B_1^n : v(x) = 0, Dv(x) = 0\} \tag{22}$$

is the projection on B_1^n of $\{(x, u_1(x)) : u_1(x) = u_2(x)\}$ (i.e. where the two sheets agree - we know that they must be tangential there).

We are assuming that on the embedded C^2 parts the two sheets are CMC with mean curvatures having the same absolute value $H > 0$. Moreover for u_1 we have $\vec{H} = -H\hat{\nu}$ and for u_2 we have $\vec{H} = H\hat{\nu}$, with $\hat{\nu} \cdot \vec{e}^{n+1} > 0$ (i.e. normal vectors $\hat{\nu}$ to the sheets pointing upwards).

Then consider the (possibly several) connected components of $B_1^n \setminus T$, these are open sets above which u_1 and u_2 are C^2 and classically CMC. Then we can neglect multiplicities on the two sheets (since on such a connected component each sheet is counted with constant multiplicity) and deduce the equations

$$\begin{aligned} \int_{B_1^n} \frac{D_i u_1}{\sqrt{1+|Du_1|^2}} D_i \zeta &= H \int_{B_1^n} \zeta \\ \int_{B_1^n} \frac{D_i u_2}{\sqrt{1+|Du_2|^2}} D_i \zeta &= -H \int_{B_1^n} \zeta \end{aligned} \tag{23}$$

valid for all test functions ζ compactly supported in $B_1^n \setminus T$. Writing

$$F^i(p) = \frac{p_i}{\sqrt{1+|p|^2}} \quad \text{for } p \in \mathbb{R}^n,$$

setting $a_{ij}(p,q) = \int_0^1 \frac{\partial F^i}{\partial p_j}(p-(1-t)q)dt$ and taking the difference of the two equations in (23), we find

$$\int_{B_1^n} a_{ij}(Du_a, Dv) D_j v D_i \zeta = -H \int_{B_1^n} \zeta$$

for all test functions ζ compactly supported in $B_1^n \setminus T$.

Since $\frac{\partial F^i}{\partial p_j} = \frac{\delta_{ij}}{\sqrt{1+|p|^2}} - \frac{p_i p_j}{(1+|p|^2)^{\frac{3}{2}}}$, $a_{ij}(p,q) = \delta_{ij} + b_{ij}(p,q)$ with $|b_{ij}(p,q)| \leq C(|p|^2+|q|^2)$, which will allow us to view the PDE obtained for v, namely,

$$\int_{B_1^n} (\delta_{ij} + b_{ij}(Du_a, Dv)) D_j v D_i \zeta = -H \int_{B_1^n} \zeta \tag{24}$$

as a perturbation of (the weak form of) $\Delta v = H$. Recall that we only use test functions ζ compactly supported in $B_1^n \setminus T$. By assumption T has zero $\mathcal{H}^n$-measure and $v=0$, $Dv=0$ on T.

PDE for the average. Consider u_1 with multiplicity $q_1(x)$ and u_2 with multiplicity $q_2(x) = q - q_1(x)$, with $q_j(x)$ constant on any open set disjoint from T. The first variation gives

$$\int \left(\frac{q_1(x) D_i u_1}{\sqrt{1+|Du_1|^2}} + \frac{q_2(x) D_i u_2}{\sqrt{1+|Du_2|^2}} \right) D_i \zeta = H \int (q_1(x) - q_2(x)) \zeta. \tag{25}$$

Writing $F^i(p) = \frac{p_i}{\sqrt{1+|p|^2}}$ for $p \in \mathbb{R}^n$, seting

$$a_{ij}(p,q) = 2\int_0^1 (q_1(x) + t(q_2(x) - q_1(x))) \frac{\partial F^i}{\partial p_j}((2t-1)p+q)dt \quad \text{and}$$

$$b_{ij}(p,q) = a_{ij}(p,q) + (q_2(x)-q_1(x))\delta_{ij} \int_0^1 \frac{(2t-1)}{\sqrt{1+|(2t-1)Du_a + Dv|^2}} dt,$$

we obtain from the above that

$$\begin{aligned} \int_\Omega b_{ij}(Du_a, Dv) D_j u_a D_i \zeta &= H \int_\Omega (q_2(x) - q_1(x))\zeta \\ &\quad - \int_\Omega (q_2(x)-q_1(x)) \left(\int_0^1 \frac{1}{\sqrt{1+|(2t-1)Du_a+Dv|^2}} dt \right) D_i v D_i \zeta \end{aligned}$$

8.3. Higher Hölder regularity for the gradients. Next we need to improve exponent α of the $C^{1,\alpha}$ regularity obtained in the Sheeting Theorem 7.1 for $\alpha < \frac{1}{2}$:

PROPOSITION 8.1. *Under the assumption that* $V \in \mathcal{S}_H$ *and*

$$V \llcorner \left(B^n_{1/2}(0) \times \mathbb{R}\right) = \sum_{j=1}^q |\mathrm{graph}\, u_j|,$$

with $u_j \in C^{1,\alpha}\left(B^n_{1/2}(0);\mathbb{R}\right)$ *and* $u_1 \leq u_2 \leq \cdots \leq u_q$ *for some* $\alpha(0,\frac{1}{2})$, *we have that* $u_j \in C^{1,\alpha}\left(B^n_{1/2}(0);\mathbb{R}\right)$ *for any* $\alpha \in (0,1)$.

In particular we will need $\alpha \geq \frac{1}{2}$ at a later stage in this section. Note that it suffices, in view of Lemma 8.1, to focus on two sheets u_1 and u_2, for which we may assume the knowledge of $C^{1,\alpha}$ regularity for some $\alpha < \frac{1}{2}$: the aim is to conclude that then actually the Hölder exponent can be improved. Roughly speaking the strategy for the proof of Proposition 8.1 is to show a De Giorgi type decay for each separate sheet by obtaining such a decay for the semi-difference v and for the average u_a at all touching points.

Consider the semi-difference v on $B_1 = B^n_1(0)$ (without loss of generality we are considering graphs on $B_1(0)$, which can always be obtained by homothetic rescaling) and let $K := \{v = 0, Dv = 0\}$ and assume without loss of generality that $0 \in K$. We saw in (24) that v satisfies a PDE of the form

(26)

$$\int_{B_1} (\delta_{ij} + b_{ij}(x))\, D_j v D_i \zeta = -H \int_{B_1} \zeta \quad \text{for } \zeta \in C^1_c(B_1 \setminus K) \quad \text{with } H > 0.$$

In the case of (24), the coefficients $b_{ij}(x)$ are, more precisely, the $C^{0,\alpha}$ functions $b_{ij}(Du_a(x), Dv(x))$. In order to produce a De Giorgi type decay for v we will use a contradiction argument/blow up method for which the following a priori estimate is needed.

PROPOSITION 8.2 (Schauder estimate for the semi-difference). *There exist* $\beta, H_0 > 0$ *such that if* $v \in C^{1,\alpha}(B_1)$ *satisfies the PDE (26) on* $B_1 \setminus K$ *with* $\sup_{B_1} |b_{ij}| + [b_{ij}]_{\alpha,B_1} \leq \beta$ *and* $H \leq H_0$ *then*

$$[Dv]_{\alpha,B_{1/2}} \leq C(\|v\|_{L^2(B_1)} + H),$$

where C *depends only on* n, H_0, β.

The key difficulty here is that the PDE is satisfied only away from the closed set K, of which no regularity properties are known. To prove Proposition 8.2 we adapt the scaling argument due to L. Simon ([**Sim97**]).

LEMMA 8.2. $\forall \delta > 0\, \exists\, C > 0$, $C = C(\delta, H_0, \beta, \delta)$ *such that*

$$[Dv]_{\alpha,B_{1/2}} \leq \delta [Dv]_{\alpha,B_1} + C(|v|_{0,B_1} + |Dv|_{0,B_1} + H).$$

PROOF OF LEMMA 8.2. The proof proceeds by contradiction: assume that we can find $\delta > 0$ such that for every $k \in \mathbb{N}$ there is a v_k satisfying the same assumptions and such that the reverse inequality holds, i.e. for every k we have (with $0 \leq H_k \leq H$)

$$[Dv_k]_{\alpha,B_{1/2}} \geq \delta[Dv_k]_{\alpha,B_1} + k(|v_k|_0 + |Dv_k|_0 + H_k). \tag{27}$$

Choose $x_k, y_k \in B_{1/2}$ such that

$$\frac{|Dv_k(x_k) - Dv_k(y_k)|}{|x_k - y_k|^\alpha} > \frac{1}{2}[Dv_k]_{\alpha,B_{1/2}}. \tag{28}$$

Denote $\rho_k = |x_k - y_k|$. Observe that in view of (28) and (27) we have $\frac{1}{2}[Dv_k]_{\alpha,B_{1/2}} \leq \frac{2|Dv_k|_0}{\rho_k^\alpha} \leq \frac{2[Dv_k]_{\alpha,B_{1/2}}}{k\rho_k^\alpha}$. Therefore we must have $\rho_k \to 0$ as $k \to \infty$.

We distinguish two cases: (I) $B_{\rho_k}(x_k) \subset B_1 \setminus K_k$ for all k, (II) there exists a subsequence (not relabelled) such that there is a point z_k in the set $K_k \cap B_{\rho_k}(x_k)$.

If we are in case (I) consider the rescaling of v_k as follows:

$$w_k(x) := \frac{v_k(x_k + \rho_k x) - v_k(x_k) - \rho_k \sum_{i=1}^n D_i v_k(x_k) x_i}{\rho_k^{1+\alpha}[Dv_k]_{\alpha,B_1}}.$$

If we are in case (II) consider the rescaling of v_k as follows:

$$w_k(x) := \frac{v_k(x_k + \rho_k x)}{\rho_k^{1+\alpha}[Dv_k]_{\alpha,B_1}}$$

and denote $p_k = \frac{z_k - x_k}{\rho_k}$. Then we have $w_k = 0$ and $Dw_k = 0$ at the point $p_k \in B_1(0)$ (the point p_k plays the role of an "anchor point" so there is no need to subtract the first jet, which we must do in case (I) in order to make the origin the "anchor point").

In either case we consider the sequence w_k thus obtained and we will now prove that it converges (up to extraction of a subsequence) in C^1 on any compact set to a $C^{1,\alpha}$ function w defined on $\mathbb{R}^n$ and such that w is harmonic on the open set $\{w \neq 0\}$. Then we will show that such a w is harmonic in the whole of $\mathbb{R}^n$.

Case (I): by definition we have $w_k(0) = 0$, $Dw_k(0) = 0$ and let $\zeta_k := \frac{y_k - x_k}{\rho_k}$ so that $|\zeta_k| = 1$. By (28) and (27)

$$|Dw_k(\zeta_k) - Dw_k(0)| > \frac{\delta}{2},$$

and moreover we know that $[Dw_k]_{\alpha,B_{\frac{1}{2\rho_k}}(0)} \leq 1$ by rescaling properties. Since $Dw_k(0) = 0$ then for $0 < \sigma < \frac{1}{2\rho_k}$ and $x \in B_\sigma(0)$ we have $|Dw_k(x)| = |Dw_k(x) - Dw_k(0)| \leq [Dw_k]_{\alpha,B_{\frac{1}{2\rho_k}}(0)}|x - 0|^\alpha \leq \sigma^\alpha$. Similarly $|w_k(x)| \leq \sigma^{1+\alpha}$.

We now want to send $k \to \infty$. We have just seen that on any compact set a tail of the sequence satisfies the requirements of Ascoli-Arzelà's theorem: this yields a $C^{1,\alpha}$ function w on $\mathbb{R}^n$ to which a subsequence of w_k converges in C^1 on any compact set (a diagonal argument is needed here, by using an exhausting sequence of compact sets $\overline{B}_{\frac{1}{2\rho_k}}$). We may assume, up to extracting a further subsequence that we do not relabel, that $\zeta_k \to \zeta$ as $k \to \infty$ with $|\zeta| = 1$. The function w satisfies (by the C^1 convergence)

$$[Dw]_{\alpha,\mathbb{R}^n} \leq 1 \quad \text{and} \quad |Dw(\zeta) - Dw(0)| > \frac{\delta}{2}. \tag{29}$$

Case (II): this time, with the notation $p_k = \frac{z_k - x_k}{\rho_k}$, we have $w_k(p_k) = 0$, $Dw_k(p_k) = 0$. We have still $\zeta_k := \frac{y_k - x_k}{\rho_k}$ so that $|\zeta_k| = 1$. By (28) and (27)

$$|Dw_k(\zeta_k) - Dw_k(0)| > \frac{\delta}{2},$$

and moreover we know that $[Dw_k]_{\alpha,B_{\frac{1}{2\rho_k}}(0)} \leq 1$ by rescaling properties. For $0 < \sigma < \frac{1}{2\rho_k}$ and $x \in B_\sigma(0)$ we have $|Dw_k(x)| = |Dw_k(x) - Dw_k(p_k)| \leq [Dw_k]_{\alpha,B_{\frac{1}{2\rho_k}}(0)}|x - p_k|^\alpha \leq (\sigma+1)^\alpha$. Similarly $|w_k(x)| \leq (\sigma+1)^{1+\alpha}$. As before we can extract a converging subsequence using Ascoli-Arzelà's theorem and get as above w of class $C^{1,\alpha}$ on $\mathbb{R}^n$ with (29) valid.

Since $v_k \geq 0$ for all k, in case (II) we will have $w \geq 0$ and the set $\{w = 0\}$ is the set of z such that there exists a sequence $z_k \to z$ with $x_k + \rho_k z_k \in K_k$. We will consider the open set $\{w > 0\}$ and will show that w is harmonic there.

In case (I), on the other hand, let Z_k be the closed set such that $x_k + \rho_k Z_k = K_k$; then w_k takes on the set Z_k the same value as the affine function $\frac{-v_k(x_k) - \rho_k \sum_{i=1}^n D_i v_k(x_k) \cdot_i}{\rho_k^{1+\alpha}[Dv_k]_{\alpha,B_1}}$; here $\cdot$ denotes the variable and $\cdot_i$ its i-th coordinate. Therefore the limsup Z (as $k \to \infty$) of the sets Z_k will be such that the value of w on Z coincides with the value taken by a certain affine function, which is the C^1-limit of those exhibited above. We will consider the open set $\mathbb{R}^n \setminus Z$ and will conclude that w is harmonic there.

Observe that both in case (I) and case (II) we can obtain a PDE for w on the open set $\mathbb{R}^n \setminus Z$ by suitable rescaling the (weak) PDEs for v_k and passing to the limit by the C^1 convergence: indeed the open set on which we are focusing comes from dilations of the open sets on which the PDEs for v_k are valid. The limiting process follows the lines of the blow up argument in L. Simon's [**Sim97**] and will be omitted.

We now have w that is harmonic away from a closed set where it coincides with an affine function. By subtracting this affine function we obtain both in cases (I) and (II) that we have a function $w \in C^{1,\alpha}$ that is harmonic wherever it is non-zero. Then w is harmonic in the whole of $\mathbb{R}^n$. To see this we show first of all that $|D^2 w|$ (which is well-defined on $\mathbb{R}^n \setminus Z$ since w is smooth there) is locally L^2-summable on $\mathbb{R}^n \setminus Z$.

Consider a convex smooth and monotone increasing function $\gamma_\delta : [0, \infty[\to \mathbb{R}$, for $\delta > 0$, such that

$$\gamma_\delta(t) = \begin{cases} 0 & \text{for } 0 \leq t \leq \delta \\ t - \delta & \text{for } t \geq 2\delta \end{cases},$$

and γ_δ is positive increasing. Noting that w is smooth wherever $\gamma_\delta(|\nabla w|^2) \neq 0$ we can compute

$$\Delta\left(\gamma_\delta(|\nabla w|^2)\right) = \operatorname{div}\left(\gamma'_\delta(|\nabla w|^2)\nabla(|\nabla w|^2)\right) =$$

$$= \gamma''_\delta(|\nabla w|^2)\left|\nabla(|\nabla w|^2)\right|^2 + \gamma'_\delta(|\nabla w|^2)\Delta(|\nabla w|^2))$$

(keeping in mind that the higher derivatives of w have only appeared where they are classically well-defined). Using the convexity assumption $\gamma''_\delta \geq 0$ and Bochner's formula for harmonic functions we conclude that

$$\Delta\left(\gamma_\delta(|\nabla w|^2)\right) \geq 2\gamma'_\delta(|\nabla w|^2)|D^2_{ij}w|^2.$$

Let ψ be any non-negative bump function that is compactly supported in $B_R \subset \mathbb{R}^n$ and identically 1 in $Z \cap B_{R-1}$ (R large). Then the previous inequality yields

$$2\int_{|Dw|\geq\sqrt{2\delta}} |D^2 w|^2\psi \leq 2\int_{\mathbb{R}^n} \gamma'_\delta(|\nabla w|^2)|D^2 w|^2\psi \leq \int_{\mathbb{R}^n} \Delta\left(\gamma_\delta(|\nabla w|^2)\right)\psi \leq$$

$$\leq \int_{\mathbb{R}^n} \gamma_\delta(|\nabla w|^2)\Delta\psi \leq C_R\|Dw\|_{C^0}\int \Delta\psi,$$

and sending $\delta \to 0$ (since the r.h.s. does not depend on δ) we have the desired local summability on the open set $\{Dw \neq 0\}$. At points in the set $\{Dw = 0\} \setminus Z$ the function w is C^2 (and harmonic), in particular $D^2 w = 0$ almost everywhere on this set. In view of this we conclude the L^2-summability of $D^2 w$ on any domain $B_R \cap (\mathbb{R}^n \setminus Z)$. We note that actually we would need only L^1 summability for the next step.

Now with $f = D_i w$ for $i \in \{1, \ldots, n\}$ we can use [**BelWic18**, Lemma C1] to conclude that the function that takes the value $D_l f$ on $\mathbb{R}^n \setminus \{f = 0\}$ and the value 0 on $\{f = 0\}$ is in $L^2_{\text{loc}}(\mathbb{R}^n)$ and is the distributional l-derivative of $f = D_i v$ in $\mathbb{R}^n$. In other words we have obtained that $w \in W^{2,2}_{\text{loc}}$ and so $\Delta w = 0$ on the whole of $\mathbb{R}^n$: so w is everywhere harmonic (recall that adding back the affine function in case (I) does not change the harmonicity of the blow up function w).

To conclude the proof of Lemma 8.2 we only need to recall that we have obtained a harmonic function w on $\mathbb{R}^n$ satisfying (29). Then $D_l w$ is harmonic (for any l) and with sublinear growth (first inequality in (29)), which implies by Liouville theorem that $D_l w$ is constant. But then we contradict the second inequality in (29). $\square$

Proof of Proposition 8.2. The estimate follows as in [**Sim97**] from Lemma 8.2 by means of the "simple abstract lemma" [**Sim97**, p. 398]. $\square$

Now we are in a position to prove the De Giorgi-type decay for the "modified L^2-excess" of v.

Proposition 8.3. *Let $v \in C^{1,\alpha}(B_1)$ (with $0 < \alpha < \frac{1}{2}$) solve the PDE (26) on $B_1 \setminus K$. Assume that $\sup_{B_1}|b_{ij}| + [b_{ij}]_{\alpha,B_1} \leq \beta$ and $H \leq H_0$, with*

β, H_0 as in Proposition 8.2. Let $\mu \in (0,1)$. For every $\theta \in (0, \frac{1}{4})$ there exists $\varepsilon > 0$ (depending on θ) such that if $\int_{B_1} |v|^2 + \frac{H^2}{\varepsilon} \leq \varepsilon$ then

$$\frac{1}{\theta^{n+2}} \int_{B_\theta} |v|^2 + \frac{\theta^2}{\varepsilon} H^2 \leq C\theta^{2\mu} \left(\int_{B_1} |v|^2 + \frac{H^2}{\varepsilon} \right). \tag{30}$$

Proof of Proposition 8.3. We consider a sequence of v_k satisfying the PDE with H_k on the r.h.s. and such that $\int_{B_1} |v_k|^2 + \frac{H_k^2}{\varepsilon_k} \leq \varepsilon_k \to 0$. Consider the blow-up obtained by dividing by the L^2-height excess:

$$\tilde{v}_k := \frac{v_k}{\left(\int_{B_1} |v_k|^2 + \frac{H_k^2}{\varepsilon_k} \right)^{\frac{1}{2}}}$$

and note that $\|\tilde{v}_k\|_{C^{1,\alpha}(B_{1/2})}$ is bounded by C independently of k thanks to the a priori Schauder type estimate in Proposition 8.2. This allows to find $\tilde{v} \in C^{1,\alpha}(B_{1/2})$ to which a subsequence of $\tilde{v}_k$ converges in $C^1(B_{1/2})$. Moreover $\|\tilde{v}_k\|_{L^2(B_{1/2})}$ are equibounded ≤ 1 and for any $R < 1/2$ the norms $\|\tilde{v}_k\|_{W^{1,2}(B_R)}$ are equibounded, implying the strong L^2-convergence on every compact set in $B_{1/2}$. Then $\|\tilde{v}\|_{L^2(K)} \leq 1$ independently of K on any compact set K contained in $B_{1/2}$. So $\tilde{v} \in L^2(B_{1/2})$. Moreover $\tilde{v}$ is harmonic on $B_{1/2}$ (by the rescaling properties of the PDE the function $\tilde{v}_k$ solves the PDE (24) with $H_k \leq \varepsilon_k$ on the r.h.s. and we have C^1 convergence so we can pass the PDE to the limit) and by C^1 convergence we have that $D\tilde{v}(0) = 0$ (since $Dv_k(0) = 0$ for every k as $0 \in T$). Harmonic function theory and the vanishing of the gradient at 0 guarantee the validity of the inequality

$$\frac{1}{\theta^{n+2}} \int_{B_\theta} |\tilde{v}|^2 \leq C\theta^{2\mu} \int_{B_{1/2}} |\tilde{v}|^2, \quad \text{with } C = C_n.$$

By strong L^2-convergence on B_θ we have, for k large enough (we need to ensure the validity of $\int_{B_\theta} |\tilde{v} - \tilde{v}_k|^2 < \theta^{n+2+2\mu}$) the inequality

$$\frac{1}{\theta^{n+2}} \int_{B_\theta} |\tilde{v_k}|^2 \leq \theta^{2\mu} \left(C_n \int_{B_{1/2}} |\tilde{v}|^2 + 1 \right),$$

in other words (recall that $\int_{B_{1/2}} |\tilde{v}|^2 \leq 1$), by definition of $\tilde{v}_k$,

$$\frac{1}{\theta^{n+2}} \int_{B_\theta} |v_k|^2 \leq (C_n + 1)\theta^{2\mu} \left(\int_{B_1} |v_k|^2 + \frac{H_k^2}{\varepsilon_k} \right).$$

The desired inequality (30) follows now immediately, since the second term on the l.h.s. of (30) is easily bounded. □

The decay result obtained for (26) immediately applies to the PDE (24) for the semi-difference, upon ensuring the bounds for b_{ij} and H by performing suitable homothetic dilations.

We will now focus on the average of the two sheets and prove a similar decay property. Rather than a PDE for the average, however, we will seek a

PDE for the weighted average $q_1u_1 + q_2u_2$. Comparing with the case of the semi-difference, we wish to point out that this time we will have no control over the values of the weighted average at the points corresponding to the "touching set", nor on the gradient; on the plus side, however, we will have a PDE that is valid across the touching set.

The PDE is obtained from the first variation formula (25): in a first step we have

$$\int \left(\underbrace{\frac{q_1(x)D_iu_1}{\sqrt{1+|Du_1|^2}} + \frac{q_2(x)D_iu_2}{\sqrt{1+|Du_2|^2}}} \right) D_i\zeta = H \int (q_1(x) - q_2(x))\zeta,$$

and we can rewrite the braced expression as

$$q_1(x)D_iu_1 \left(1 + \left(\frac{1}{\sqrt{1+|Du_1|^2}} - 1\right)\right) + q_2(x)D_iu_2 \left(1 + \left(\frac{1}{\sqrt{1+|Du_2|^2}} - 1\right)\right)$$

so that the equation becomes

$$\int (q_1(x)D_iu_1 + q_2(x)D_iu_2)\, D_i\zeta = H\int (q_1(x)-q_2(x))\zeta + \int O(|Du_1|^3)D_i\zeta + \int O(|Du_2|^3)D_i\zeta.$$

Rewriting $q_1(x)D_iu_1 = D_i(q_1(x)u_1) - u_1D_iq_1$ (and similarly for u_2) we get

$$\int D_i(q_1(x)u_1 + q_2(x)u_2)D_i\zeta = H\int (q_1(x)-q_2(x))\zeta + \int (u_1D_iq_1 + u_2D_iq_2)D_i\zeta + \int O(|Du_1|^3 + |Du_2|^3)D_i\zeta$$

The middle term on the right hand side is zero, since on the support of the distributions D_iq_1, D_iq_2 (the touching set of the two graphs, see remark 8.2) we have $u_1 = u_2$ and $D_iq_1 = -D_iq_2$. This gives the PDE for the weighted average $q_1u_1 + q_2u_2$:

$$\int D_i(q_1u_1 + q_2u_2)D_i\zeta = H\int (q_1 - q_2)\zeta + \int O(|Du_1|^3)D_i\zeta + \int O(|Du_2|^3)D_i\zeta. \tag{31}$$

Choice of excess for the weighted average. We let, for a reference affine function L and $\rho \in (0, 1]$ (here $\varepsilon > 0$ is a constant to be chosen in Proposition 8.4 below):

$$E_L^2(\rho) := \rho^{-n-2} \int_{B_\rho} |q_1 u_1 + q_2 u_2 - L|^2 + \rho^2 \frac{H^2}{\varepsilon} + \frac{\left(\rho^{3\alpha}\right)^2}{\varepsilon} \left([Du_1]^3_{\alpha,B_\rho}\right)^2 + \frac{\left(\rho^{3\alpha}\right)^2}{\varepsilon} \left([Du_2]^3_{\alpha,B_\rho}\right)^2 .$$

We will prove the following decay at points where $u_1 = u_2$, assuming that 0 is such a point and that *we have rotated coordinates so that* $Du_1(0) = Du_2(0) = 0$ and q is fixed.

PROPOSITION 8.4 (De Giorgi decay for the weighted average). *Let* $\mu \in (\frac{1}{2}, 1)$ *and* $\theta \in (0, \frac{1}{4})$. *Assume that* $Du_1(0) = Du_2(0) = 0$. *There exists* $\varepsilon = \varepsilon(n, \mu, \theta)$ *and* $C = C(n, \mu)$ *such that, if* $E_L^2(1) \leq \varepsilon$ *for a certain* L, *then there exists* L' *for which*

$$E_{L'}^2(\theta) \leq C\theta^{2\mu} E_L^2(1), \tag{32}$$

$$|L - L'|^2_{C^1(B_1)} \leq C E_L^2(1). \tag{33}$$

Notational remark: the excess $E_L^2(\rho)$ computed for the function u_k will be denoted in the following proof by $E_L^2(\rho, k)$.

PROOF. Note that the last two terms in the definition of the excess E are defined so that they are naturally rescaling under the geometric transformation $\frac{u(rx)}{r}$ when evaluating the Hölder seminorm at 0.

Elliptic L^2-theory for a PDE as (31) of the form $\operatorname{div}(\nabla w) = f + \operatorname{div} g$ with $f \in L^\infty(B_1)$ and $g \in L^2(B_1(0))$ give $\|w\|_{W^{1,2}(B_{1/2})} \leq C(\|w\|_{L^2(B_1)} + \|f\|_{L^\infty(B_1)} + \|g\|_{L^2(B_1)})$, see e.g. [**GilTru**, Chap. 8]. Note that we can subtract a linear function from $q_1 u_1 + q_2 u_2$ and find (since a linear function is harmonic)

$$\int D_i(q_1 u_1 + q_2 u_2 - L) D_i \zeta = H \int (q_1 - q_2)\zeta + \int O(|Du_1|^3) D_i \zeta + \int O(|Du_2|^3) D_i \zeta$$

so by the above elliptic estimate,

$$\|q_1 u_1 + q_2 u_2 - L\|_{W^{1,2}(B_{1/2})} \leq C \left(\|q_1 u_1 + q_2 u_2 - L\|_{L^2(B_1)} + Hq + \left(\int_{B_1} O\left(|Du_1|^3\right)^2 \right)^{1/2} + \left(\int_{B_1} O\left(|Du_2|^3\right)^2 \right)^{1/2} \right)$$

Using the fact that $Du_j(0) = 0$ and hence $|Du_1|(x) = |Du_j(x) - Du_j(0)| \leq [Du_j]_{\alpha,B_1} |x|^\alpha$ in this gives

$$\|q_1 u_1 + q_2 u_2 - L\|_{W^{1,2}(B_{1/2})} \leq C \left(\|q_1 u_1 + q_2 u_2 - L\|_{L^2(B_1)} + H + [Du_1]^3_{\alpha,B_1} + [Du_2]^3_{\alpha,B_1} \right) \leq K E_L(1).$$

We can thus perform a blow up argument, assuming to have a sequence $\varepsilon_k \to 0$ and u_1^k, u_2^k with $E_L^2(1,k) \le \varepsilon_k$. Define

$$\tilde{u}_k = \frac{q_1 u_1^k + q_2 u_2^k - L}{E_L(1,k)}.$$

The elliptic estimate just discussed gives that $\|\tilde{u}_k\|_{W^{1,2}(B_{1/2})} \le K$ and therefore $\tilde{u}_k \to \tilde{u} \in W^{1,2}(B_{1/2})$ strongly in $L^2(B_{1/2})$ and weakly in $W^{1,2}(B_{1/2})$. The blow up function $\tilde{u}$ is harmonic on $B_{1/2}$. Indeed the PDE (31) for $\tilde{u}_k$ is of the form

$$\int D_i \tilde{u}_k D_i \zeta = \int f_k \zeta + \int g_k D_i \zeta$$

with $|f_k| \le \varepsilon_k$ and $|g_k| \le \varepsilon_k$, for any $\zeta \in C_c^\infty(B_{1/2})$. Passing to the limit thanks to the $W^{1,2}$-weak convergence on the left-hand-side we get

$$\int_{B_{1/2}} D_i \tilde{u} D_i \zeta = 0,$$

as desired. By the harmonicity of $\tilde{u}$ we get an affine function $\tilde{L}'$ (tangent to the graph of $\tilde{u}$ above the origin) and a constant $\tilde{C} = C(n,\mu)$ such that

$$\theta^{-n-2} \int_{B_\theta} |\tilde{u} - \tilde{L}'|^2 \le \tilde{C} \theta^{2\mu} \int_{B_{1/2}} |\tilde{u}|^2.$$

Recalling the definition of $\tilde{u}_k$ and by the L^2-strong convergence to $\tilde{u}$ we can see that for k large enough it holds

$$\theta^{-n-2} \int_{B_\theta} |q_1 u_1^k + q_2 u_2^k - L - E_L(1,k)\tilde{L}'|^2 \le \left(\tilde{C} \int_{B_{1/2}} |\tilde{u}|^2 + 1 \right) \theta^{2\mu} E_L^2(1,k),$$

and we can set L' to be the affine function $L + E_L(1,k)\tilde{L}'$ and recall that $\int_{B_{1/2}} |\tilde{u}|^2 \le K$ (the constant in the elliptic estimate): this bounds the first term in the excess that appears on the left-hand-side of (32). The remaining three terms of the excess on the left-hand-side of (32) are easily bounded by using $\theta \le 1$ and $[Du_j]_{\alpha,B_\theta} \le [Du_j]_{\alpha,B_1}$ (for $j = 1,2$).

Note that $\nabla \tilde{u}$ is harmonic and so by the mean vaue theorem $|\nabla \tilde{L}'| = |\nabla \tilde{u}(0)|$ and $|\tilde{u}(0)| = |\tilde{L}'(0)| \le \|\tilde{u}\|_{L^2(B_{1/2})} \le K$, so $|\tilde{L}'|_{C^1(B_1)} \le \|\tilde{u}\|_{W^{1,2}(B_{1/2})} \le K$ (the constant in the elliptic estimate). We therefore have also obtained the desired control (33) over the tilting of the plane. $\square$

Remark 8.3 (Iteration step). We can iteratively apply Proposition 8.4. Start with the given u_1 and u_2 with coordinates set such that $u_1(0) = u_2(0) = 0$ and $Du_1(0) = Du_2(0) = 0$. If the varifold was dilated enough we can also ensure smallness of H and of $[Du_j]_{\alpha,B_1}$ in order to have $E_L^2(1) \le \varepsilon$ (at this initial step the choice of $L = 0$ will do). Now consider the homotetic dilation of a factor θ (which dilates the ball $B_\theta^{n+1}(0)$ to the unit ball $B_1^{n+1}(0)$). The corresponding transformations for the two graphs are

$\tilde{u}_1(x) = \frac{u_1(\theta x)}{\theta}$ and $\tilde{u}_2(x) = \frac{u_2(\theta x)}{\theta}$. Setting $\tilde{f}(x) = \frac{f(\theta x)}{\theta}$ we have that $\int_{B_1} \tilde{f}^2 = \int_{B_1} \frac{f(\theta x)^2}{\theta^2} = \frac{1}{\theta^{n+2}} \int_{B_\theta} f^2$. Therefore for $q_1\tilde{u}_1 + q_2\tilde{u}_2$ we find

$$\int_{B_1} |q_1\tilde{u}_1 + q_2\tilde{u}_2 - \tilde{L}'|^2 = \theta^{-n-2} \int_{B_\theta} |q_1 u_1^k + q_2 u_2^k - L'|^2,$$

where $\tilde{L}'$ is the affine function obtained by dilating L'.

Consider the PDE (31) written for $q_1\tilde{u}_1 + q_2\tilde{u}_2$, with $\tilde{u}_1$, $\tilde{u}_2$, $\tilde{H}$ and note that $\tilde{H} = \theta H$ by rescaling properties. Therefore the second term $\frac{1}{\varepsilon}\theta^2 H^2$ that appears (for $q_1u_1 + q_2u_2$) in $E^2(\theta)$ is equal to $\frac{1}{\varepsilon}\tilde{H}^2$, i.e. the second term that appears in $E^2(1)$ for $q_1\tilde{u}_1 + q_2\tilde{u}_2$.

The dilation preserves the fact that $\tilde{u}_1 = \tilde{u}_2 = D\tilde{u}_1 = D\tilde{u}_2 = 0$ at the origin. Moreover, since $D\tilde{u}_1(x) = Du_1(\theta x)$ we have

$$\begin{aligned}[D\tilde{u}_1]_{\alpha,B_1} &= \sup_{x,y\in B_1} \frac{|D\tilde{u}_1(x) - D\tilde{u}_1(y)|}{|x-y|^\alpha} = \theta^\alpha \sup_{x,y\in B_1} \frac{|Du_1(\theta x) - Du_1(\theta y)|}{|\theta x - \theta y|^\alpha} \\ &= \theta^\alpha [Du_1]_{\alpha,B_\theta}\end{aligned}$$

and therefore (for $j \in \{1,2\}$) the term $\theta^{3\alpha}[Du_j]^3_{\alpha,B_\theta}$ that appears (for $q_1u_1 + q_2u_2$) in $E^2(\theta)$ is equal to the term $[D\tilde{u}_j]^3_{\alpha,B_1}$ that appears, for $q_1\tilde{u}_1 + q_2\tilde{u}_2$, in $E^2(1)$.

In other words we have seen that the excess (for $q_1\tilde{u}_1 + q_2\tilde{u}_2$) $E^2_{\tilde{L}'}(1)$ is exactly equal to the excess $E^2_{L'}(\theta)$ of $q_1u_1 + q_2u_2$ at scale θ with respect to L' and so the decay (32) ensures that the smallness condition of the excess at scale 1 is still satisfied by $q_1\tilde{u}_1 + q_2\tilde{u}_2$, upon choosing θ small enough so that $C\theta^{2\mu} < 1$, which can be done since C does not depend on θ. We are thus able to apply Proposition 8.4 to $\tilde{u}_1, \tilde{u}_2$ and iterate. Rescaling back to the original picture this iteration gives, for $d \in \mathbb{N}$, an affine function $L^{(d)}$ such that

$$E^2_{L^{(d)}}(\theta^d) \leq C^d \theta^{2\mu d} E^2_L(1).$$

Moreover the second line in (33) gives the convergence of $L^{(d)}$ to a certain affine function L_∞ (both the gradients and the translations, in the original picture, converge by the geometric control provided by the inequality). Note that C is independent of θ.

Choosing $0 < \delta < \mu$ (δ much smaller than μ) and θ small enough

$$E^2_{L_\infty}(\rho) \leq \rho^{2(\mu-\delta)} E^2_{L_\infty}(1).$$

Moreover we know that $Du_1(0) = Du_2(0) = 0$: this means that L_∞ is actually the zero-function and the previous inequality shows how the excess decays to it.

An analogous iteration argument applies to $u_1 - u_2$ based on the decay established in Proposition 8.3.

Proof of Proposition 8.1. By translating and rotating appropriately and applying the above decay results, we obtain L^2 average decay of u_1 and u_2 at every touching singularity to an affine function. Moreover away from these points we have complete smoothness and elliptic estimates for u_1 and u_2: combining these two facts (see [**Wic14a**, Lemma 4.3]) with standard theory of Campanato spaces gives that $Du_1, Du_2 \in C^{0,\mu-\delta}(B_{1/2})$. Here μ can be chosen as close as we want to 1 and δ is as small as we wish. $\square$

8.4. $W^{2,2}$ estimates. The aim of this section is to conclude that $v \in W^{2,2}_{\text{loc}}(B)$. The function v solves (24) for test functions ζ compactly supported in $B^n_1 \setminus T$. By assumption $v = 0$, $Dv = 0$ on T. Moreover $v > 0$ on $B \setminus T$ and Dv is Hölder continuous on B and actually C^1 on $B \setminus T$ (the fact that T has zero $\mathcal{H}^n$-measure is not needed for this argument). With Proposition 8.5 below we will ensure a L^2-bound for $D^2 v|_{B\setminus T}$: once this is achieved, [**BelWic18**, Lemma C1] will immediately imply that $v \in W^{2,2}_{\text{loc}}(B)$.

Proposition 8.5. *We have the summability statement* $\int_{B\setminus T} |D^2 v|^2 < \infty$.

The proof occupies the rest of this section. For the purpose of obtaining the L^2-bound for $D^2 v|_{B\setminus T}$ we will use the PDE (24) with the test function $D_k(\gamma_\delta(D_k v)\zeta)$ and $\zeta \in C^\infty_c(B)$ (for a fixed k), where, for $\delta > 0$, $\gamma_\delta : \mathbb{R} \to \mathbb{R}$ is a smooth non-decreasing function such that $\gamma_\delta(t) = 0$ for $|t| < \delta/2$, $\gamma_\delta(t) = t - \delta$ for $t > \delta$, $\gamma_\delta(t) = t + \delta$ for $t < -\delta$ and $\gamma'_\delta(t) \leq 1$ for all $t \in \mathbb{R}$. Plugging in, switching the order of differentation and integrating by parts we obtain

$$\int_B (\delta_{ij} + b_{ij}(Du_a, Dv))\, D_k D_j v\, D_i(\gamma_\delta(D_k v)\zeta) + \int_B D_k\left((\delta_{ij} + b_{ij}(Du_a, Dv))\right) D_j v\, D_i(\gamma_\delta(D_k v)\zeta) = 0.$$

Standard computations, with the notation $b_{ij}(p,q)$ and $p = (p_1, \ldots, p_n)$, $q = (q_1, \ldots, q_n)$, yield

$$\begin{aligned}
&\int_{|D_k v|>\delta} (\delta_{ij} + b_{ij}(Du_a, Dv))(D_j D_k v)(D_i D_k v)\zeta \\
&\quad + \int_{|D_k v|>\delta} (D_{q_l} b_{ij})(Du_a, Dv)\, D_k D_l v\, D_j v\, D_i D_k v\, \zeta \\
&\qquad = -\int_{|D_k v|>\delta} (\delta_{ij} + b_{ij}(Du_a, Dv))(D_j D_k v)\gamma_\delta(D_k v) D_i\zeta \\
&\qquad\quad - \int_{|D_k v|>\delta} (D_{q_l} b_{ij})(Du_a, Dv)\, D_k D_l v\, D_j v\, \gamma_\delta(D_k v) D_i\zeta \\
&\qquad\quad - \int_{|D_k v|>\delta} (D_{p_l} b_{ij})(Du_a, Dv)\, D_k D_l u_a\, D_j v\, D_i D_k v\, \zeta \\
&\qquad\quad - \int_{|D_k v|>\delta} (D_{p_l} b_{ij})(Du_a, Dv)\, D_k D_l u_a\, D_j v\, \gamma_\delta(D_k v) D_i\zeta
\end{aligned} \tag{34}$$

and relabelling indexes in the second term on the left-hand-side we can rewrite the left hand side of the previous equality (34) in the form

$$\int_{|D_k v|>\delta} (\delta_{ij} + b_{ij}(Du_a, Dv))(D_j D_k v)(D_i D_k v)\zeta + \int_{|D_k v|>\delta} (D_{q_j} b_{il})(Du_a, Dv)\, D_k D_j v\, D_l v\, D_i D_k v\, \zeta. \tag{35}$$

Here notice that the matrix whose coefficients are

$$\delta_{ij} + b_{ij}(Du_a, Dv)) + \sum_l (D_l v)(D_{q_j} b_{il})(Du_a, Dv)$$

is uniformly positive definite (by the smallness of b_{ij} in C^1-norm and the Hölder continuity of Du_a and Dv with $Du_a(0) = Dv(0) = 0$; either start with a good dilation so that these quantities are small enough, or work at this stage in a small enough ball around 0). So there exists $c > 0$ (independent of δ) such that

$$c \int_{|D_k v|>\delta} |\nabla(D_k v)|^2 \zeta \leq (35) \quad \text{for all } \zeta \geq 0. \tag{36}$$

Recall that (35) is the left-hand side of (34) and so we can replace (35) by the right-hand side of (34). Therefore we will now analyse the four terms on the right-hand side of (34) and find suitable bounds. We begin with the third term: by Young's inequality, choosing $\varepsilon > 0$ small compared to c in (36), we find

$$\begin{aligned} &\left| \int_{|D_k v|>\delta} (D_{p_l} b_{ij})(Du_a, Dv)\, D_k D_l u_a\, D_j v\, D_i D_k v\, \zeta \right| \\ &\quad \leq \frac{1}{\varepsilon} \int_{|D_k v|>\delta} |(D_{p_l} b_{ij})(Du_a, Dv)|^2 |D_k D_l u_a|^2 |D_j v|^2 \zeta + \varepsilon \int_{|D_k v|>\delta} |D_i D_k v|^2 \zeta. \end{aligned} \tag{37}$$

The second term on the r.h.s. of (37) will be absorbed on the left-hand side of (36): it is not a priori clear, however, that the first term on the r.h.s. of (37) is summable with a uniform bound, independently of δ. This is obtained in the following lemma (which will also be needed later on in Sect. 8.5).

Lemma 8.3. *We have the following summability statements:*

$$\int_{B\setminus T} |D^2 u_a|^2 |D_j v|^2 \zeta < \infty \quad \textit{and} \quad \int_{B\setminus T} |D^2 v|^2 |D_j v|^2 \zeta < \infty.$$

PROOF. The summability of $\int\limits_{B\setminus T} |D^2u_a|^2|D_jv|^2$ comes from the following observations. Let $x \in B \setminus T$ and take the largest open ball $B_R(x)$ that is disjoint from T, so that $R = \mathrm{dist}(x,T)$. Then $\exists y \in T \cap \partial B_R(x)$. The function D_lu_1 satisfies (by differentiating the CMC equation)

$$\int_{B_R(x)} \frac{1}{\sqrt{1+|Du_1|^2}}\left(\delta_{ij} - \frac{D_iu_1D_ju_1}{1+|Du_1|^2}\right) D_j(D_lu_1)D_i\zeta = 0 \tag{38}$$

with uniform ellipticity on $B\setminus T$. In other words $f = D_lu_1$ satisfies $\mathrm{div}(A\nabla f) = 0$ with $A = (a_{ij})_{i,j=1}^n$, where $a_{ij} = \frac{\partial a^i}{\partial p_j}(Du_1) = \frac{1}{\sqrt{1+|Du_1|^2}}\left(\delta_{ij} - \frac{D_iu_1D_ju_1}{1+|Du_1|^2}\right)$ and $a^i(p) = \frac{p_i}{\sqrt{1+|p|^2}}$, so a_{ij} are smooth uniformly elliptic entries (we can assume that ∇u_1 is small) and hence by Schauder estimates, we have that for any constant $\lambda_1 \in \mathbb{R}$,

$$|\nabla(D_lu_1)|(x) \leq \frac{C}{R}\sup_{B_R(x)}|D_lu_1 - \lambda_1|. \tag{39}$$

Similarly, for any constant $\lambda_2 \in \mathbb{R}$,

$$|\nabla(D_lu_2)|(x) \leq \frac{C}{R}\sup_{B_R(x)}|D_lu_2 - \lambda_2|. \tag{40}$$

Taking the average $u_a = \frac{u_1+u_2}{2}$ we get

$$\nabla(D_lu_a) = \frac{1}{2}(\nabla(D_lu_1) + \nabla(D_lu_2))$$

so from (39) and (40) we infer, after choosing $\lambda_1 = D_lu_1(y)$ and $\lambda_2 = D_lu_2(y)$ and using the Hölder continuity of Du_1 and Du_2, that $|D^2u_a|(x) \leq C\mathrm{dist}(x,T)^{\alpha-1}$. Therefore

$$\int\limits_{B\setminus T} |D^2u_a|^2|D_jv|^2\zeta \leq \int\limits_{B\setminus T} \mathrm{dist}(x,T)^{2(2\alpha-1)}\zeta < \infty \tag{41}$$

since we can choose $\alpha \geq \frac{1}{2}$, as proved in Proposition 8.1. Note that the same summability behaviour can be obtained for $v = \frac{u_1-u_2}{2}$, i.e. $|D^2v|(x) \leq C\mathrm{dist}(x,T)^{\alpha-1}$ and

$$\int\limits_{B\setminus T} |D^2v|^2|D_jv|^2\zeta \leq \int\limits_{B\setminus T} \mathrm{dist}(x,T)^{2(2\alpha-1)}\zeta < \infty. \tag{42}$$

□

Lemma 8.3 establishes the desired bound on the first term on the r.h.s. of (37). In order to conclude the L^2-bound for $D^2v|_{B\setminus T}$ in Proposition 8.5, we now go back to the remaining three terms on the right hand side of (34). Let us analyse the second:

$$\left| \int\limits_{|D_k v|>\delta} (D_{q_l} b_{ij})(Du_a, Dv)\, D_k D_l v\, D_j v\, \gamma_\delta(D_k v) D_i \zeta \right| \leq \int\limits_{|D_k v|>\delta} |(D_{q_l} b_{ij})(Du_a, Dv)|\, |D_k D_l v|\, |D_j v|\, (|D_k v| + \delta) |D_i \zeta| \tag{43}$$

and we can use the summability of $|D^2 v||Dv|$ on $B \setminus T$, which is a consequence of Lemma 8.3, so the right-hand-side of (43) is finite and bounded independently of δ. We can argue similarly for the first and fourth terms on the right hand side of (34). With this we conclude (recall that there is an absorption on the left hand side to be performed, for each fixed δ, when splitting the third term on the right hand side of (34)), sending $\delta \to 0$, that

$$\int\limits_{B\setminus T} |\nabla(D_k v)|^2 \zeta < \infty.$$

We thus have $D^2 v|_{B\setminus T}$ is in L^2_{loc}. As mentioned at the beginning of this section, this implies $v \in W^{2,2}_{\text{loc}}(B)$ since we can now apply [**BelWic18**, Lemma C1] ($D^2 v$ develops no further distribution supported on T).

8.5. Higher regularity conclusions. Consider the PDE for v (24): the aim of this subsection is to extend the PDE for v across T and, as a straightforward consequence, complete the proof of (iv), i.e. the inductive step of Theorem 7.3. While it is indeed true that the PDE for v is valid for arbitrary test functions $\zeta \in C^\infty_c(B)$ (this can be proved exactly with the same argument that we are going to use in this subsection) what we will need is the extension of the PDE for v that takes into accound multiplicities as well.

Consider the function $f(x) = q_1(x)(\delta_{ij} + b_{ij}(Du_a, Dv))D_j v$ on B. Recall that q_1 is constant on each connected component of $B \setminus T$. The function f is C^1 on $B \setminus T$ and its derivative is summable on $B \setminus T$, in view of the bounds established in Lemma 8.3. Moreover it vanishes everywhere on the set T (since Dv vanishes there[9]) and it is Hölder continuous on B. Indeed for $y \in T$ and $x \in B \setminus T$ we have (recall that $0 < q_1(x) < q$)

$$|f(x) - f(y)| = |f(x)| = |q_1(x)||\delta_{ij} + b_{ij}(Du_a, Dv)||D_j v| \leq K_{b_{ij},q}|Dv(x)| =$$
$$= K_{b_{ij},q}|Dv(x) - Dv(y)| \leq K_{b_{ij},q}|x-y|^\alpha.$$

For $x, y \in B \setminus T$ and x and z in distinct connected components of $B \setminus T$ we consider the segment joining x and z, which must intersect T at a point y. Then $|f(x) - f(z)| \leq |f(x) - f(y)| + |f(z) - f(y)| \leq K_{b_{ij},q}|x-y|^\alpha + K_{b_{ij},q}|z-y|^\alpha \leq 2K_{b_{ij},q}|x-z|^\alpha$. For x, y in the same connected component of $B \setminus T$ then q_1 is constant, so the Hölder continuity follows.

[9] q_1 is not really defined on T but we just set the function to be 0 since Dv vanishes.

We then have by [**BelWic18**, Lemma C1] that the distributional derivative of f on B is an L^1 function, namely is is just the classical derivative on $B \setminus T$ extended by setting it to be 0 on T. So we have $f \in W^{1,1}(B)$. We then have, for an arbitrary $\zeta \in C_c^\infty(B)$:

$$\begin{aligned}
&\int_B q_1(x)(\delta_{ij} + b_{ij}(Du_a, Dv))D_j v D_i \zeta \\
&\quad = -\int_B D_i \left(q_1(x)(\delta_{ij} + b_{ij}(Du_a, Dv))D_j v\right) \zeta \\
&\quad = -\int_{B\setminus T} D_i \left(q_1(x)(\delta_{ij} + b_{ij}(Du_a, Dv))D_j v\right) \zeta \\
&\quad = -\int_{B\setminus T} q_1(x) D_i \left((\delta_{ij} + b_{ij}(Du_a, Dv))D_j v\right) \zeta = -\int_{B\setminus T} q_1(x) H\zeta,
\end{aligned}$$

where in the second equality we use the fact that T has measure 0, in the third equality the fact that q_1 is locally constant on $B \setminus T$ and in the last equality the fact that $D_i\left((\delta_{ij} + b_{ij}(Du_a, Dv))D_j v\right) = H$ on $B \setminus T$ (by the PDE (24), which is valid in its strong form on $B \setminus T$ since v is C^2 there). Since T has zero measure we can write the last term of the previous chain of equalities as $-\int_B q_1(x) H\zeta$: in conclusion we find a version of the PDE for v that holds on the whole of B also when we take multiplicity into account, namely:

$$\int_B q_1(x)(\delta_{ij} + b_{ij}(Du_a, Dv))D_j v D_i \zeta = -\int_B q_1(x) H\zeta. \tag{44}$$

Now consider the first variation identity (25), valid for any arbitrary $\zeta \in C_c^\infty(B)$:

$$\int_B \left(\frac{q_1(x) D_i u_1}{\sqrt{1+|Du_1|^2}} + \frac{q_2(x) D_i u_2}{\sqrt{1+|Du_2|^2}}\right) D_i\zeta = \int_B (q_1(x) - q_2(x)) H\zeta.$$

We rewrite it, using $q_2(x) = q - q_1(x)$, as

$$\begin{aligned}
\int_B q \frac{D_i u_2}{\sqrt{1+|Du_2|^2}} D_i \zeta + \int_B q_1(x) \left(\frac{D_i u_1}{\sqrt{1+|Du_1|^2}} - \frac{D_i u_2}{\sqrt{1+|Du_2|^2}}\right) D_i\zeta \\
= \int_B (2q_1(x) - q) H\zeta.
\end{aligned}$$

The term $\left(\frac{D_i u_1}{\sqrt{1+|Du_1|^2}} - \frac{D_i u_2}{\sqrt{1+|Du_2|^2}}\right)$ is exactly $2(\delta_{ij} + b_{ij}(Du_a, Dv))D_j v$, as computed pointwise when we obtained the PDE (24) for v (it is 0 on T). Using (44) we then find

$$\int_B q \frac{D_i u_2}{\sqrt{1+|Du_2|^2}} D_i\zeta = -\int qH\zeta.$$

Simplifying q we have the standard CMC equation in weak form for u_2, where the test function ζ is arbitrary and $u_2 \in C^{1,\alpha}(B)$. The difference quotients method yields $u_2 \in W^{2,2}(B)$ and differentiating the PDE proves Du_2 is $C^{1,\alpha}$. Bootstrapping yields $u_2 \in C^\infty$.

With the conlcusions obtained so far we have achieved the proof of the Higher Regularity Theorem 7.3 subject to the inductive assumptions, i.e. step (iv) of the general inductive program from Sect. 7. At this stage the entire induction process is complete, and with it the proof of Theorem 6.2 (and consequently Theorem 6.1 is proved as well).

9. Curvature estimates for weakly stable CMC hypersurfaces

Theorem 6.3 above—the Compactness Theorem—is a qualitative statement implied by the curvature estimate of Theorem 7.1—the Sheeting Theorem—which is in essence the quantitative version of the Compactness Theorem. However, note that the Compactness Theorem holds under the assumption of weak stability, i.e. stability for volume preserving deformations (which is the natural stability assumption for non-minimal CMC hypersurfaces), whereas the Sheeting Theorem stated as above assumes strong stability. In this section and the next, we shall briefly discuss how the Compactness Theorem can in turn be used to strengthen the Sheeting Theorem by weakening the strong stability assumption to weak stability.

We begin with a statement of the estimate in non-singular dimensions (i.e. in dimensions ≤ 6). In these dimensions the result holds under weaker hypotheses than in general dimensions, in complete analogy with the long known curvature estimates for strongly stable minimal hypersurfaces ([**SSY75**, **SchSim81**]).

Theorem 9.1 (Pointwise curvature estimate in low dimension). *For each $H_0 > 0$ and $\Lambda \geq 1$, there exists $C = C(H_0, \Lambda)$ such that the following holds: Let $2 \leq n \leq 6$ and let $\Sigma \subset B_R(0) \subset \mathbb{R}^{n+1}$ be a smooth immersed hypersurface with $(\overline{\Sigma} \setminus \Sigma) \cap B_R(0) = \emptyset$, $\mathcal{H}^n(\Sigma) \leq \Lambda R^n$ and with constant scalar mean curvature H such that $|H| \leq H_0 R^{-1}$. Assume that Σ is weakly stable as a CMC immersion. If $n = 6$ suppose additionally either that Σ contains no point of transverse self-intersection (or equivalently, by the maximum principle, for each point $p \in \Sigma$ where Σ is not embedded, there is $\rho > 0$ such that $\Sigma \cap B^{n+1}_\rho(p)$ is the union of at two embedded smooth CMC hypersurfaces intersecting only tangentially), or that $\Lambda = 3 - \delta$ for some $\delta \in (0, 1)$.*

Then
$$\sup_{x \in \Sigma \cap B_{R/2}(0)} |A_\Sigma|(x) \leq C R^{-1},$$
where A_Σ denotes the second fundamental form of Σ.

Remark 1. When $n = 2$, an even stronger result, namely, a curvature estimate without the bounded area assumption, is known ([**Ye96**, **EM12**]); it is a consequence of the strong Bernstein type theorems known in this dimension ([**BarDoC84**, **BDE88**, **Pal86**, **DaS87**, **LopRos89**]).

REMARK 2. The reason, in case $n = 6$, for the additional restrictions in Theorem 9.1 (that either Σ has no transverse points or $\Lambda = 3 - \delta$) is that it is not known if a pointwise curvature estimate holds for 6-dimensional immersed strongly stable minimal hypersurfaces satisfying an arbitrary mass bound; such an estimate is only known to hold if the minimal hypersurface is either embedded ([**SchSim81**]) or is immersed and satisfies a mass bound corresponding to $\Lambda = 3 - \delta$ for some $\delta \in (0, 1)$ ([**Wic08**]). See Proposition 10.1 below.

In general dimensions, we have the following:

THEOREM 9.2 (Sheeting Theorem). *Let $\Lambda, H_0 > 0$ and $n \geq 2$. Suppose that $\Sigma \subset B_R(0) \subset \mathbb{R}^{n+1}$ is an immersed hypersurface with $\mathcal{H}^n(\Sigma) \leq \Lambda R^n$, with constant scalar mean curvature H such that $|H| \leq H_0 R^{-1}$ and with $\mathcal{H}^{n-7+\alpha}((\overline{\Sigma} \setminus \Sigma) \cap B_R(0)) = 0$ for all $\alpha > 0$. Suppose that Σ contains no point of transverse self-intersection (or equivalently, by the maximum principle, for each $p \in \Sigma$ where Σ is not embedded there is $\rho > 0$ such that $\Sigma \cap B^{n+1}_\rho(p)$ is the union of exactly two embedded smooth CMC hypersurfaces intersecting only tangentially), and that Σ is weakly stable as a CMC immersion. There exists $\delta_0 = \delta_0(n, H_0, \Lambda)$ so that if additionally*

$$\Sigma \subset \{|x^{n+1}| \leq \delta_0 R\}$$

then $\overline{\Sigma} \cap \left(B^n_{R/2}(0) \times \mathbb{R}\right)$ is equal to the union of the graphs of finitely many ordered functions $u_1 \leq \cdots \leq u_k$ defined on $B^n_{R/2}(0) := B_{R/2}(0) \cap \{x^{n+1} = 0\}$ and satisfying

$$\sup_{B^n_{R/2}(0)} \left(|Du_i| + R|D^2 u_i|\right) \leq C \left(R^2 H^2 + R^{-n-2} \int_{B_R(0) \cap \Sigma} (x^{n+1})^2\right)^{\frac{1}{2}}$$

for $i = 1, \ldots, k$, where $C = C(n)$; moreover, each u_i is a smooth CMC graph.

REMARK 3. The conclusion of Theorems 9.2 clearly fails (even for strongly stable minimal hypersurfaces) for $n \geq 7$ without the flatness assumption $\Sigma \subset \{|x^{n+1}| \leq \delta_0 R\}$, as seen by the construction of Hardt–Simon [**HS85**]. We also note that singularities do occur in stable CMC hypersurfaces (with $H \neq 0$) of dimension ≥ 7, as shown in [**Irv17**] modifying earlier work of Caffarelli–Hardt–Simon ([**CHS84**]) giving a construction of singular minimal hypersurfaces.

It is interesting to note the following: Consider a CMC hypersurface Σ immersed in $\mathbb{R}^{n+1}$ with mean curvature H (possibly equal to zero). Recall that the *Morse index* $I(\Sigma)$ of Σ is the number of negative Dirichlet eigenvalues (counting with multiplicity) of the Jacobi operator $L : C^2(\Sigma) \to \mathbb{R}$ given by $L\phi = -\Delta\, \phi - |A_\Sigma|^2 \phi$. It is readily checked that if Σ is weakly stable, then

$I(\Sigma) \leq 1$.[10] On the other hand, Theorem 9.2, and hence also Theorem 9.1 above, are *false* in every dimension ≥ 2 if we replace "Σ is weakly stable" with "Σ satisfies $I(\Sigma) = 1$." This can be seen by considering rescalings of the higher-dimensional Catenoid (the unique non-flat rotationally symmetric minimal hypersurface in $\mathbb{R}^{n+1}$) which converge weakly to a hyperplane with multiplicity two, but do not have uniformly bounded curvature (and hence do not satisfy the conclusion of Theorem 9.2). In the context of Theorem 9.1 and Theorem 9.2, the crucial difference between "weakly stable" and "$I(\Sigma) = 1$" is that complete weakly stable surfaces cannot have two ends at infinity while index one surfaces can (e.g., the Catenoid). Indeed, the one-endedness of weakly stable entire minimal hypersurfaces, a result due to Chen, Cheung and Zhou ([**CCZ08**]), is a key ingredient of our proofs of Theorem 9.1 and Theorem 9.2.

10. Proof of the curvature estimates

Theorem 9.1 and Theorem 9.2 generalise, respectively, the Schoen–Simon–Yau curvature estimates ([**SSY75**]) and the Schoen–Simon Sheeting Theorem ([**SchSim81**]) for strongly stable minimal (and certain other) hypersurfaces, by relaxing strong stability to weak stability. The methods used in [**SSY75, SchSim81**] involve the use of *positive* test functions in the stability inequality, and since they never integrate to zero, it is not clear how one could directly adapt these methods to the setting of weak stability. The strategy employed in [**BCW18**] in the proofs of Theorem 9.1 and Theorem 9.2 is different; it is a geometric approach that combines the results of [**SSY75, SchSim81**] with the fact that complete weakly stable minimal hypersurfaces have only one end, a result established by Cheng–Cheung–Zhou ([**CCZ08**]) for non-singular hypersurfaces and generalized in [**BCW18**] (in a fairly straightforward manner) to allow a small singular set. (This generalization, Theorem 10.1 below, is only necessary for Theorem 9.2.)

A key difficulty in the proof of Theorem 9.2 is to correctly "localize" the one-end result in order to transfer the "flatness" from large to small scales (see Remark 4). This is handled by a careful blow-up procedure relying on the aforementioned Compactness Theorem (Theorem 6.3) for weakly stable CMC hypersurfaces and a rigidity theorem (Lemma 10.1 below) due to Simons ([**JSim68**]) for minimal hypersurfaces of spheres.

10.1. One-endedness-at-infinity of weakly stable minimal hypersurfaces. We begin by describing the one end result we shall need, given precisely in Theorem 10.1 below. In the absence of singularities this result

[10] If $I(\Sigma) \geq 2$, then there are two L^2 orthogonal functions $\phi_1, \phi_2 \in C^2_c(\Sigma) \setminus \{0\}$ and numbers $\lambda_1, \lambda_2 < 0$ such that $\Delta \phi_j + |A_\Sigma|^2 \phi_j = -\lambda_j \phi_j$ for $j = 1, 2$, whence by direct calculation using integration by parts we see that the function $\phi = c_1\phi_1 + c_2\phi_2$, where c_1, c_2 are constants, satisfies $\int_\Sigma |A_\Sigma|^2\phi^2 - |\nabla \phi|^2 = -c_1^2\lambda_1 \int_\Sigma \phi_1^2 - c_2^2\lambda_2 \int_\Sigma \phi_2^2 > 0$ provided not both c_1, c_2 are zero. This contradicts weak stability of Σ if we choose c_1, c_2 such that $\int_\Sigma \phi = 0$.

was proven by Cheng–Cheung–Zhou [**CCZ08**], generalizing the earlier work of Cao–Shen–Zhu [**CSZ97**] establishing the same result for strongly stable minimal hypersurfaces. The proof of the result allowing a small singular set, which is needed for the present purposes, follows closely the argument of [**CCZ08**] (reproduced in [**BCW18**], Theorem 7) but with a little additional care, as described below, to handle the presence of singularities.

THEOREM 10.1 ([**BCW18**], Theorem 8). *For $n \geq 3$, suppose that V is a stationary integral n-varifold in $\mathbb{R}^{n+1}$ with $\operatorname{spt}\|V\|$ connected, $\dim_{\mathcal{H}}(\operatorname{sing} V) \leq n-7$ and with $\operatorname{sing} V \subset B_1(0)$. Assume that the regular part $\operatorname{spt}\|V\| \setminus \operatorname{sing} V$ is weakly stable. Then $\operatorname{spt}\|V\|$ has exactly one end at infinity.*

PROOF (SKETCH). Let $M = \operatorname{spt}\|V\| \setminus \operatorname{sing} V$ and suppose that M has two (or more) ends at infinity. The proof proceeds as in the non-singular case ([**CCZ08**, Theorem 3.2] reproduced in [**BCW18**, Theorem 7], together with [**CSZ97**, Lemma 2]) with a few additional arguments. First we note that by [**Ilm96**, Theorem A (ii)], $\operatorname{sing} V$ does not disconnect $\operatorname{spt}\|V\|$, so M is connected. This is needed at the end of the proof in order to have that the identical vanishing of ∇u for a C^1 function u on M implies the constancy of u on M, which provides the desired contradiction for a function u constructed appropriately; indeed, the idea is to use two ends of M to construct a non-constant bounded harmonic function u on M with $\int_M |\nabla u|^2 < \infty$ (this does not require any stability condition), and use the weak stability inequality to show that $|\nabla u| = 0$.

The existence of such u is proved in the case $\operatorname{sing} V = \emptyset$ in [**CSZ97**, Lemma 2]. The completeness assumption in [**CSZ97**, Lemma 2] is not necessarily fulfilled by our M. We note, however, that completeness is used in [**CSZ97**] only to infer that each end has infinite volume ([**CSZ97**, Lemma 1]); this fact on the other hand follows directly from the monotonicity formula. Then we can follow verbatim the arguments in [**CSZ97**, Lemma 2], only with the following additional care. When we exhaust the hypersurface with domains D_i (notation as in [**CSZ97**, Lemma 2]) we should remove, from the D_i constructed as in [**CSZ97**, Lemma 2], the closure of a smooth tubular neighbourhood N_i of $\operatorname{sing} V$ (whose size shrinks to zero as $i \to \infty$). This will produce further boundary components, in addition to those as in [**CSZ97**], on which we will set boundary value 0 when solving the Dirichlet problem [**CSZ97**, (2)].

One more additional argument is needed in view of the fact that the test function $\varphi_t|\nabla u|$ constructed in [**CCZ08**, Theorem 3.2]) might fail, a priori, to be an admissible function for the stability inequality. Indeed, $\varphi_t|\nabla u|$ is not compactly supported on M and we do not have sufficient control of $|\nabla u|$ near $\operatorname{sing} V$. (If e.g. $|\nabla u|$ were bounded near $\operatorname{sing} V$, a straightforward capacity argument would suffice.) In order to overcome this difficulty, we first observe an energy growth estimate for u in balls centred on $\operatorname{sing} V$ (inequality (45) below) which is obtained as follows: Since $\Delta u = 0$ on M, we see by integrating by parts and using the Cauchy–Schwarz inequality that for any $\phi \in C^1_c(M)$,

$$\int_M |\nabla u|^2\phi^2 = -\int_M 2\phi u\nabla u\nabla\phi \leq 2\left(\int_M \phi^2|\nabla u|^2\right)^{1/2}\left(\int_M u^2|\nabla\phi|^2\right)^{1/2},$$

from which we immediately get (using also the bounds $-1 \leq u \leq 1$)

$$\int_M |\nabla u|^2\phi^2 \leq 4\int_M u^2|\nabla\phi|^2 \leq 4\int_M |\nabla\phi|^2.$$

Consequently, a standard capacity argument that only needs that the 2-capacity of $\operatorname{sing} V$ is 0 (true in view of $\mathcal{H}^{n-2}(\operatorname{sing} V) = 0$) gives that the inequality

$$\int_M |\nabla u|^2\phi^2 \leq 4\int_M |\nabla\phi|^2$$

holds for all $\phi \in C^1_c(\mathbb{R}^{n+1})$. In particular, choosing any $p \in \operatorname{spt} V$ and $\phi \in C^1_c(B_{2r}(p))$ to be a standard bump function that is identically equal to 1 on $B_r(p)$ and identically equal to 0 on the complement of $B_{2r}(p)$ with $|\nabla^{\mathbb{R}^{n+1}}\phi| \leq \frac{2}{r}$, the preceding inequality and the monotonicity formula give

$$\int_{B_r(p)\cap M} |\nabla u|^2 \leq Cr^{n-2}, \tag{45}$$

where C is independent of r.

With this we proceed as follows: Let $\delta > 0$ and let $\{B_{r_i}(x_i)\}_{i=1}^N$ be a cover of $\operatorname{sing} V$ (compact since $\operatorname{sing} V \subset B_1(0)$) with $\sum_{i=1}^N r_i^{n-4} \leq \delta$. (Such a cover exists since $\dim_{\mathcal{H}}(\operatorname{sing} V) \leq n-7$.) Defining a cutoff function $\zeta_\delta = \min_{i\in\{1,\ldots,N\}} \zeta_i$, where $\zeta_i \in C^1(\mathbb{R}^{n+1})$ with $0 \leq \zeta_i \leq 1$, $\zeta_i = 0$ in $B_{r_i}(x_i)$, $\zeta = 1$ in $\mathbb{R}^{n+1} \setminus B_{2r_i}(x_i)$ and $|D\zeta_i| \leq 2r_i^{-1}$, in view of (45) we see that

$$\begin{aligned}\int_M |\nabla\zeta_\delta|^2|\nabla u|^2 &\leq 2\sum_{i=1}^N \int_{B_{2r_i}(x_i)\cap M} |\nabla\zeta_i|^2|\nabla u|^2 \\ &\leq 8\sum_{i=1}^N r_i^{-2}\int_{B_{2r_i}(x_i)\cap M} |\nabla u|^2 \leq 8C\sum_{i=1}^N r_i^{n-4} \leq 8C\delta\end{aligned}$$

whence

$$\int_M |\nabla\zeta_\delta|^2|\nabla u|^2 \to 0 \tag{46}$$

as $\delta \to 0$. We can now adapt the arguments of [**CCZ08**, Theorem 3.2]: let $\epsilon > 0$ be arbitrary and choose $R > 0$ so that $(n-1)R^{-2}\int_\Sigma |\nabla u|^2 < \epsilon$. For every $\delta > 0$ choose a compactly supported function $\varphi_{t,\delta}$ that is constructed in the manner φ_t is constructed in [**CCZ08**, Theorem 3.2] so as to ensure $\int \zeta_\delta\varphi_{t,\delta}|\nabla u| = 0$ (one can work with the function φ_t defined above and set $a \geq 1$ so that φ_t is 1 on the singular set and $\zeta_\delta|\nabla u| = |\nabla u|$ on $B^\Sigma_{a+2R+b}(p) \setminus B^\Sigma_{a+2R}(p)$; the t for which the zero-average condition is met will depend however on δ, hence the dependence of $\varphi_{t,\delta}$ on δ). As $\delta \to 0$, since $|\varphi_{t,\delta}| \leq 1$ and $|\nabla\varphi_{t,\delta}| \leq \frac{1}{R}$ are uniformly bounded, and moreover $b(\epsilon, a)$ can be chosen independently of δ, we get for a sequence $\delta_j \to 0^+$ that $\varphi_{t,\delta_j} \to \tilde{\varphi}$ for a compactly supported Lipschitz function satisfying $\int \tilde{\varphi}|\nabla u| = 0$ and $\tilde{\varphi}$ identically 1 on $B_a^{\Sigma}(p)$.

Plugging the (admissible) test function $\zeta_\delta \varphi_{t,\delta}|\nabla u|$ in the stability inequality we get, arguing as in [**CCZ08**, Theorem 3.2] by means of Bochner's formula and Gauss equations,

$$\begin{aligned}
&\int_M |\nabla|\nabla u||^2 (\zeta_\delta \varphi_{t,\delta})^2 \\
&\leq (n-1)\int_M |\nabla u|^2 |\nabla(\zeta_\delta \varphi_{t,\delta})|^2 \\
&= (n-1)\int_M |\nabla \zeta_\delta|^2 \varphi_{t,\delta}^2 |\nabla u|^2 + 2\varphi_{t,\delta}\zeta_\delta |\nabla u|^2 \nabla\zeta_\delta \cdot \nabla \varphi_{t,\delta} + \zeta_\delta^2 |\nabla \varphi_{t,\delta}|^2 |\nabla u|^2;
\end{aligned}$$

using the Cauchy–Schwarz inequality for the middle term on the right-hand side, setting $\delta = \delta_j$ and letting $j \to \infty$ we obtain (recalling (46)) that the first and second term on the right-hand side vanish in the limit; moreover, with the choices of ϵ and R recalled above, using that $\zeta_{\delta_j} \uparrow 1$, $\varphi_{t,\delta_j} \to \tilde{\varphi}$ as $\delta_j \to 0$, and that $|\nabla \varphi_{t,\delta_j}| \leq \frac{1}{R}$, we get

$$\int_M |\nabla|\nabla u||^2 \tilde{\varphi}^2 \leq \epsilon,$$

which allows us to conclude, in view of the arbitrariness of ϵ, that $|\nabla u|$ is constant. Since $|\nabla u| \in L^2(M)$ and M has infinite volume, we conclude that $|\nabla u| = 0$ on M. This contradicts non-constancy of u. $\square$

10.2. A partial sheeting result. Next we note a consequence of the regularity and compactness theory of [**BelWic18**] (discussed in Sect. 6 and Sect. 7 above) that will be needed for the proof of Theorem 9.2. To state the result, it is convenient to use the following notation: Let R, K_0, $H_0 \in (0,\infty)$ be fixed. Denote by $\mathcal{S}_{H_0,K_0}(B_R^{n+1}(0))$ the class of all varifolds V in $B_R^{n+1}(0)$ of the type $V = \sum_{j=1}^k q_j|M_j|$ where k and q_j, $j = 1, 2, \ldots, k$ are positive integers; each M_j is a connected, immersed, smooth hypersurface (not necessarily complete) in $B_R^{n+1}(0)$; $M = \cup_{j=1}^k M_j$ is a weakly stable CMC immersion such that

- $\mathcal{H}^{n-7+\alpha}\left(\overline{M} \setminus M\right) = 0$ for all $\alpha > 0$;
- M has no transverse points; equivalently (by the strong maximum principle), at every $p \in M$ where M is not embedded, there exists $\rho > 0$ such that $M \cap B_\rho^{n+1}(p)$ is the union of exactly two embedded complete smooth CMC hypersurfaces in $B_\rho^{n+1}(p)$ that intersect only tangentially;
- $|H_M| \leq H_0$ where H_M is the mean curvature of M;
- $\|V\|(B_R^{n+1}(0)) \leq K_0$.

By Theorem 6.3, $\mathcal{S}_{H_0,K_0}(B_R^{n+1}(0))$ is a compact family in the varifold topology. Moreover, if $V_j \in \mathcal{S}_{H_0,K_0}(B_R^{n+1}(0))$ and $V_j \rightharpoonup V$ the (constant) mean curvature of V is given by $\lim_{n\to\infty} H_j$, where H_j is the (constant) mean curvature of V_j.

The result we need is the following:

THEOREM 10.2 (Sheeting away from a point for weakly stable CMC hypersurfaces). *Let V_j be a sequence of varifolds in $\mathcal{S}_{H_0,K_0}(B_R^{n+1}(0))$ with $V_j \to V \llcorner B_R^{n+1}(0)$, where $V = q|\mathbb{R}^n \times \{0\}|$ for some positive integer q. Then*

- *either, for every j large enough, we have*

$$\operatorname{spt}\|V_j\| \cap \left(B^n_{\frac{R}{2}}(0) \times \mathbb{R}\right) = \cup_{k=1}^q \operatorname{graph} u_k,$$

where $u_k \in C^{2,\alpha}\left(B^n_{\frac{R}{2}}(0);\mathbb{R}\right)$, $u_1 \le u_2 \le \cdots \le u_q$ and $\operatorname{graph} u_k$ are separately smooth CMC graphs with small gradient, and possibly with tangential intersections;

- *or there exists a point $y \in \overline{B}^n_{R/2}(0)$ and a subsequence $V_{j'}$ such that, for any $r > 0$, the following holds: for j' large enough (depending on r) $V_{j'}$ is strongly stable in $(B^n_R(0) \times \mathbb{R}) \setminus B_r^{n+1}(y)$ and moreover $V_{j'}$ converges smoothly (with sheeting and possibly with multiplicity) to V away from y in the sense that for j' large enough (depending on r), the following decomposition holds:*

$$\operatorname{spt}\|V_{j'}\| \cap \left(B^n_{\frac{9R}{10}}(0) \setminus B^n_r(y)\right) \times \mathbb{R} = \cup_{k=1}^q \operatorname{graph} u_k$$

where $u_k \in C^{2,\alpha}\left(B^n_{\frac{9R}{10}}(0) \setminus B^n_r(y);\mathbb{R}\right)$ for $k = 1,\ldots,q$, $\operatorname{graph} u_k$ are separately smooth CMC graphs (possibly with tangential intersections) with small gradients and $u_1 \le u_2 \le \cdots \le u_q$.

PROOF. This is a synthesis of [**BelWic18**, Theorem 2.1] (Theorem 6.1 above), [**BelWic18**, Theorem 3.1] (Theorem 7.1 above), [**BelWic18**, Theorem 3.3] (Theorem 7.3 above) and [**BelWic18**, Lemma 8.1]. □

10.3. Bernstein-type theorems. The one-end result of Theorem 10.1 above implies two Bernstein-type theorems for entire weakly stable minimal hypersurfaces. The first, valid in low dimensions, is the following:

PROPOSITION 10.1. *For $2 \le n \le 6$, suppose that $\Sigma^n \subset \mathbb{R}^{n+1}$ is a complete, connected, weakly stable, immersed minimal hypersurface with with $\mathcal{H}^n(\Sigma \cap B_R) \le \Lambda R^n$ for some constant $\Lambda \ge 1$ and all $R > 0$. When $n = 6$ assume either that $\Lambda = 3 - \delta$ for some $\delta > 0$ or that Σ is embedded. Then Σ is a hyperplane.*

The second Bernstein-type result, valid in all dimensions, is as follows:

PROPOSITION 10.2. *For $n \ge 3$, suppose that V is a stationary integral n-varifold in $\mathbb{R}^{n+1}$ with $\operatorname{spt}\|V\|$ a connected set, $\operatorname{sing} V \subset B_1(0)$ (i.e. $\operatorname{spt}\|V\|$ smoothly embedded in $\mathbb{R}^{n+1} \setminus B_1(0)$) and with $\dim_{\mathcal{H}}(\operatorname{sing} V) \le n-7$. Assume that the regular part $\Sigma = \operatorname{reg} V$ $(= \operatorname{spt}\|V\| \setminus \operatorname{sing} V)$ is weakly stable and that*

V satisfies area growth $\|V\|(B_R(0)) \leq \Lambda R^n$ for some constant $\Lambda \geq 1$ and all $R > 0$. Finally, assume that for some $\varepsilon > 0$, Σ satisfies

$$|A_\Sigma|(x)|x| \leq \sqrt{n-1} - \varepsilon \tag{47}$$

for $x \in \Sigma \setminus B_1$, where $|\cdot|$ denotes the length in $\mathbb{R}^{n+1}$. Then $\operatorname{spt}\|V\|$ is an affine hyperplane.

The proof of Proposition 10.2 relies on the following rigidity result which is a direct consequence of a result of J. Simons ([**JSim68**]).

LEMMA 10.1. *Suppose that $\mathcal{C}$ is a n-dimensional minimal cone in $\mathbb{R}^{n+1}$ that is smooth away from 0 and satisfies $|A_{\mathcal{C}}|(x)|x| < \sqrt{n-1}$ for all $x \in C \setminus \{0\}$. Then $\mathcal{C}$ is a flat hyperplane.*

PROOF. Note that $M := \mathcal{C} \cap \mathbb{S}^n$ is smooth. By the given curvature estimate, we have that $|A_M| < \sqrt{n-1}$. By [**JSim68**, Corollary 5.3.2], M must be totally geodesic. This proves the assertion. $\square$

REMARK 4. Observe that the Simons cone

$$\Sigma = \left\{ (x^1, x^2, \ldots, x^8) \,:\, \sum_{j=1}^{4} x_j^2 = \sum_{j=5}^{8} x_j^2 \right\} \cap \mathbb{R}^8$$

is a (strongly) stable minimal hypercone in $\mathbb{R}^8$ satisfying $|A_\Sigma|(x)|x| = \sqrt{n-1}$ for all $x \in \Sigma \setminus \{0\}$. Thus Lemma 10.1 is sharp, and so is Proposition 10.2 in the sense that we cannot take $\varepsilon = 0$ in (47).

REMARK 5. The importance of the size of the constant in a (scale invariant) curvature estimate of the form (47) seems to have been first shown by White in [**Whi87**]. This has been refined in [**MPR16, CCE16, CKM17**]. A novelty contained in the present setting is the combination of (47) with Lemma 10.1 and with the one-end result of Theorem 10.1, allowing flatness to propagate from large to small scales. Furthermore, the present setting appears to be the first use of a curvature estimate of the type (47) in a setting where a priori there could be singularities.

PROOFS OF PROPOSITION 10.1 AND PROPOSITION 10.2 (SKETCH). In the case of Proposition 10.1, let $V = |\Sigma|$ (the multiplicity 1 varifold associated with Σ) and in the case of Proposition 10.2 let V be as given. In either case, by the assumed area growth estimate a stationary tangent cone $C = \lim_{j\to\infty} \left(\frac{1}{\rho_j}\right)_\sharp V$ to V at infinity exists, where $\rho_j \to \infty^+$. Also, since V is weakly stable, it is strongly stable near infinity. Hence by the pointwise curvature estimates (which in the case of Proposition 10.1 are established in [**SSY75**] if $n \leq 5$, or in [**SchSim81**] if $n = 6$ and Σ is embedded, or in [**Wic08**] if $n = 6$ and Σ is immersed with $\Lambda = 3-\delta$; and in the case of Proposition 10.2 assumed as in (47)), it follows that the convergence $\left(\frac{1}{\rho_j}\right)_\sharp V \to C$ is smooth on compact subsets away from the origin. So $\operatorname{spt}\|C\|$ is smooth away from the origin and is strongly stable in the case of Proposition 10.1, or satisfies $|A_{\mathcal{C}}|(x)|x| \leq \sqrt{n-1} - \varepsilon$ for all $x \in \text{spt}\,\|C\| \setminus \{0\}$ in the case of Proposition 10.2. Hence in the case of Proposition 10.1 by [**JSim68**, Theorem 6.1.1] (see the proof in [**Sim83**, Appendix B] which is valid even when the cone is, as in our case, immersed and strongly stable as an immersion away from the origin), or in the case of Proposition 10.2 by Lemma 10.1, C is a plane with some positive integer multiplicity m independent of C.

The preceding argument shows that any tangent cone at infinity is supported on a hyperplane and arises as a smooth limit of rescalings of V on compact subsets away from the origin. It follows that there is $\rho_0 > 0$ such that $\partial\, B^{n+1}_{\rho}(0)$ intersects $\text{spt}\,\|V\|$ transversely for each $\rho > \rho_0$. Furthermore, the monotonicity formula implies that the multiplicity of any tangent plane at infinity is a fixed positive integer m independent of the cone. The rest of the argument uses [**Mil63**, Theorem 3.1] and Theorem 10.1 to show that outside $B^{n+1}_1(0)$ the varifold V itself is equal to the varifold associated with the smooth submanifold $\text{spt}\,\|V\| \setminus B^{n+1}_1(0)$ taken with constant multiplicity m. In case of Proposition 10.1 this means that $m = 1$. The conclusion in either case now follows from the monotonicity formula. See [**BCW18**] for details. □

Theorem 9.1 follows directly from Proposition 10.1 via a standard blow-up argument. See [**BCW18**, Sect. 3].

For the proof of Theorem 9.2 we proceed as follows.

10.4. A preliminary curvature estimate. The next step in the proof of Theorem 9.2 is to use a (less standard) blow-up argument based on Proposition 10.2 and the regularity theory of [**BelWic18**] (specifically, Theorem 6.3, Theorem 7.1 and Theorem 7.3 above) to establish the following preliminary curvature estimate:

PROPOSITION 10.3. *Let $\Lambda, H_0 > 0$ and let $\eta > 0$. There exists $\delta_0 = \delta_0(n, \Lambda, H_0, \eta) \in (0,1)$ such that if $R > 0$ and if Σ satisfies the hypotheses of Theorem 9.2, then $\overline{\Sigma} \cap B_{\frac{9}{10}R}(0)$ is smooth and there is $z \in B_{\frac{6}{10}R}(0)$ so that*

$$|A_{\Sigma}|(x)|x - z| \leq \eta$$

for all $x \in \Sigma \cap B_{\frac{9}{10}R}(0)$, where $|x - z|$ is the distance in $\mathbb{R}^{n+1}$.

PROOF. Fix $\Lambda, H_0, \eta > 0$. Clearly, it suffices to prove the result with $R = 10$ and assuming that $\eta < \sqrt{n-1}$. For $j = 1, 2, 3, \ldots$, let $\Sigma_j \subset B_{10}(0)$ be such that the hypotheses (of Proposition 10.3) are satisfied with Σ_j in place of Σ and with $\delta_0 = j^{-1}$. We show that the conclusion (of Proposition 10.3) holds with Σ_j in place of Σ and with some point $z_j \in B_{\frac{6}{10}}$ in place of z for infinitely many j. In view of the arbitrariness of the sequence $\{\Sigma_j\}$, this will establish the result.

We start by applying Theorem 10.2 to Σ_j for sufficiently large j. If the first case of the conclusion of Theorem 10.2 holds for every Σ_j large enough, then the conclusion of Proposition 10.3 is true with $z = 0$. So we may assume (by the second case of the conclusion of Theorem 10.2) that "partial sheeting" holds, namely, that there is a point $y \in B_5(0) \cap \{x^{n+1} = 0\}$ such that Σ_j are sheeting away from y, i.e. for any $r > 0$, $\Sigma_j \cap (B_9(0) \setminus B_r(y))$ is smooth for j sufficiently large and that

$$\sup_{x \in \Sigma_j \cap (B_9(0) \setminus B_r(y))} |A_{\Sigma_j}|(x) \to 0 \tag{48}$$

as $j \to \infty$. We will subsequently replace Σ_j by $\Sigma_j \cap B_9(0)$ (to avoid any irrelevant issues with the behavior of Σ_j near its boundary).

For each $z \in B_6(0)$ and each j, let

$$\delta(\Sigma_j, z) := \inf \left\{ r > 0 : \begin{array}{c} \Sigma_j^r := \Sigma_j \setminus \overline{B_r(z)} \text{ is smooth} \\ \text{and } |A_{\Sigma_j}|(x)|x - z| \leq \eta \text{ for all } x \in \Sigma_j^r \end{array} \right\}.$$

Note that $\delta(\Sigma_j, y) \to 0$ as $j \to \infty$, by the partial sheeting result discussed above.

For every j set $\delta_j := \inf_{z \in B_6(0)} \delta(\Sigma_j, z)$ and choose $z_{j,k}$ with $\delta(\Sigma_j, z_{j,k}) \to \delta_j$ as $k \to \infty$. Passing to a subsequence, we may assume that $z_{j,k} \to z_j \in B_6(0)$. We claim that $\delta(\Sigma_j, z_j) = \delta_j$. If not, there is $\epsilon > 0$ and $w \in \Sigma_j \setminus B_{\delta_j + 2\epsilon}(z_j)$ with either (i) $w \in \operatorname{sing} \Sigma_j$ or (ii) $|A_{\Sigma_j}|(w)|w - z_j| > \eta + 2\epsilon$. Note that $w \in \Sigma_j \setminus B_{\delta_j + \epsilon}(z_{j,k})$ for k sufficiently large. Thus, in case (i), we find that, by the definition of $\delta(\cdot, \cdot)$, $\delta(\Sigma_j, z_{j,k}) \geq \delta_j + \epsilon$ for all k sufficiently large. This contradicts the choice of $z_{j,k}$. Similarly, in case (ii) we have that

$$|A_{\Sigma_j}|(w)|w - z_{j,k}| > \eta + \epsilon,$$

for k sufficiently large, since $|w - z_{j,k}| \to |w - z_j|$ as $k \to \infty$. Again, this yields a contradiction, as before.

Thus, we have arranged that z_j minimizes $\delta(\Sigma_j, \cdot)$. Since $\delta(\Sigma_j, y) \to 0$, we also have that $\delta_j \to 0$ and consequently, it follows from the definition of δ_j and (48) that $z_j \to y$. We claim that $\delta_j = 0$ for all sufficiently large j. Arguing by contradiction, we assume (upon extracting a subsequence that we do not relabel) that $\delta_j > 0$ for all j. Using this, we now perform the relevant blow-up argument. Define

$$\tilde{\Sigma}_j = \delta_j^{-1}(\Sigma_j - z_j).$$

Note that the monotonicity formula implies that $\mathcal{H}^n(\tilde{\Sigma}_j \cap B_R(0)) \leq \tilde{\Lambda} R^n$ for some $\tilde{\Lambda} = \tilde{\Lambda}(\Lambda, n, H_0)$. Moreover, the choice of δ_j implies that $\tilde{\Sigma}_j \setminus B_1$ is smooth and satisfies

$$|A_{\tilde{\Sigma}_j}|(x)|x| \leq \eta \tag{49}$$

for $x \in \tilde{\Sigma}_j \setminus B_1$. Note also that $|H_{\tilde{\Sigma}_j}| \leq \delta_j \frac{H_0}{10} \to 0$. The area bounds and weak stability imply, by the compactness theorem of [**BelWic18**] (Theorem 6.3 above), that $\tilde{\Sigma}_j$ converge in the varifold sense to $\tilde{V}$, which is stationary, weakly stable, smoothly embedded outside of a co-dimension 7 singular set and which satisfies $\|\tilde{V}\|(B_R(0)) \leq \tilde{\Lambda} R^n$. Furthermore, by the curvature estimate (49), spt $\|\tilde{V}\|$ is a smooth hypersurface $\tilde{\Sigma}_\infty$ outside of $B_1(0)$ satisfying

$$|A_{\tilde{\Sigma}_\infty}|(x)|x| \leq \eta$$

and the convergence is smooth on compact sets outside $B_1(0)$. Thus, by Proposition 10.2, each connected component of spt $\|\tilde{V}\|$ is an affine hyperplane and so spt $\|\tilde{V}\|$ is made up of finitely many parallel affine hyperplanes.

Now, we again appeal to Theorem 10.2 to conclude that the convergence of $\tilde{\Sigma}_j$ to $\tilde{V}$ occurs smoothly (possibly with multiplicity) away from some fixed point $\tilde{z} \in$ spt $\|\tilde{V}\|$ (if the sheeting actually occurs everywhere, we simply set $\tilde{z} = 0$). Note that the curvature estimate (49) implies that $|\tilde{z}| \leq 1$.

Define $\hat{z}_j = z_j + \delta_j \tilde{z}$ and let $\hat{\delta}_j := \delta(\Sigma_j, \hat{z}_j)$. Since $\hat{\delta}_j \geq \delta_j/2$ (by the definition of δ_j), there is $w_j \in \Sigma_j \setminus B_{\delta_j/2}(\hat{z}_j)$ so that either (i) $w_j \in \operatorname{sing} \Sigma_j$, or (ii) $|A_{\Sigma_j}|(w_j)|w_j - \hat{z}_j| > \eta$. We note that in either case we have that

$$\liminf_{j\to\infty} \delta_j^{-1}|w_j - \hat{z}_j| = \infty. \tag{50}$$

(For if not, then defining $\tilde{w}_j = \delta_j^{-1}(w_j - z_j)$, we find that, in the scale of $\tilde{\Sigma}_j$ discussed above,

$$|\tilde{z} - \tilde{w}_j| = \delta_j^{-1}|(\hat{z}_j - z_j) - (w_j - z_j)| = \delta_j^{-1}|w_j - \hat{z}_j|$$

is bounded above (after passing to a subsequence) and bounded below by $\frac{1}{2}$ (since $w_j \notin B_{\delta_j/2}(\hat{z}_j)$), but because $\tilde{\Sigma}_j$ sheets away from $\tilde{z}$, we find that either (i) or (ii) would be a contradiction.) Finally, we define

$$\check{\Sigma}_j := |w_j - \hat{z}_j|^{-1}(\Sigma_j - \hat{z}_j)$$

and set

$$\check{w}_j := |w_j - \hat{z}_j|^{-1}(w_j - \hat{z}_j), \qquad \check{z}_j := |w_j - \hat{z}_j|^{-1}(z_j - \hat{z}_j) = -\delta_j|w_j - \hat{z}_j|^{-1}\tilde{z}.$$

Note that it follows from (48), (i) and (ii) that $w_j \to y$ and hence (since $z_j \to y$) that $|w_j - \hat{z}_j| \to 0$. We have already shown that outside of $B_{\delta_j}(z_j)$, Σ_j is smooth and satisfies (49). This implies that $\check{\Sigma}_j$ is smooth outside of $B_{\delta_j|w_j - \hat{z}_j|^{-1}}(\check{z}_j)$ and additionally satisfies

$$|A_{\check{\Sigma}_j}|(x)|x - \check{z}_j| \leq \eta.$$

By (50), and recalling that $|\tilde{z}| \leq 1$, we have that $\delta_j|\hat{z}_j - w_j|^{-1} \to 0$ and $\check{z}_j \to 0$. As before, we may take the varifold limit $\check{V}$ of $\check{\Sigma}_j$, and the curvature estimates we have just established show that this limit occurs smoothly (possibly with multiplicity) outside of $B_{1/2}(0)$. The curvature estimates pass to the limit so by Proposition 10.2 each connected component of spt $\|\check{V}\|$ is a hyperplane. Thus, since $|\check{w}_j| = 1$, we find that $\check{w}_j \notin \operatorname{sing} \check{\Sigma}_j$ and $|A_{\check{\Sigma}_j}|(\check{w}_j) \to 0$, so

$$|A_{\check{\Sigma}_j}|(\check{w}_j)|\check{w}_j| \to 0.$$

Rescaling to the original scale, this contradicts both (i) and (ii) above (concerning w_j). This contradiction establishes that $\delta_j = 0$ for j sufficiently large.

We have now shown that Σ_j is smooth away from z_j and satisfies

$$|A_{\Sigma_j}|(x)|x - z_j| \le \eta$$

for all $x \in \Sigma_j \setminus \{z_j\}$. If $z_j \notin \Sigma_j$ there is nothing further to show. Else, arguing as in the beginning of the proof of Proposition 10.1, we see that for fixed j and any small ball $B_\rho(z_j)$, the hypersurface Σ_j is strongly stable either in $B_\rho(z_j)$ or in $B_9 \setminus \overline{B_\rho(z_j)}$; it follows from this fact that for each fixed j, there is $\rho > 0$ such that Σ_j is strongly stable in the annulus $B_\rho(z_j) \setminus \{z_j\}$, and hence in the ball $B_\rho(z_j)$. Moreover, by the curvature estimates and Lemma 10.1, any tangent cone Σ_j at z_j must be supported on a hyperplane. Thus, Σ_j is smooth at z_j by Theorems 7.1 and 7.3 ([**BelWic18**, Theorems 3.1 and 3.3]). This completes the proof of Proposition 10.3. □

10.5. Proof of the Sheeting Theorem. Theorem 9.2, the Sheeting Theorem, now follows fairly directly from Proposition 10.2 as we describe below.

By scaling, it suffices to prove Theorem 9.2 assuming $R = 10$. Let $\Lambda, H_0 > 0$. For $j = 1, 2, 3, \ldots$, let Σ_j be an immersed hypersurface of $B_{10}(0)$ such that the hypotheses of Theorem 9.2 are satisfied with $\delta_0 = j^{-1}$ and with Σ_j in place of Σ. In particular, $\mathcal{H}^n(\Sigma_j) \le \Lambda 10^n$ and $|H_{\Sigma_j}| \le H_0/10$. We claim that $\overline{\Sigma}_j \cap B_9(0)$ is smoothly immersed for all sufficiently large j and that

$$\limsup_{j\to\infty} \sup_{X \in \Sigma_j \cap B_9(0)} |A_{\Sigma_j}| < \infty.$$

In view of the arbitrariness of Σ_j, this implies that there are constants $\delta_0 = \delta_0(n, \Lambda, H_0) \in (0, 1)$ and $C = C(n, \Lambda, H_0) > 0$ such that if the hypotheses of Theorem 9.2 (with $R = 10$) are satisfied, then $\sup_{\Sigma \cap B_9(0)} |A_\Sigma| < C$; consequently, since Σ has constant mean curvature and is contained in the δ_0 neighborhood of the hyperplane $\{x^{n+1} = 0\}$, standard arguments (including elliptic estimates) then imply the conclusions of Theorem 9.2.

To establish the claim, let $\lambda_j = \sup_{\Sigma_j \cap B_9(0)} |A_{\Sigma_j}|$ and assume for a contradiction (upon extracting a subsequence that we do not relabel) that $\lambda_j \to \infty$ as $j \to \infty$. By applying Proposition 10.3 repeatedly with $\eta = \eta_j$ for a sequence $\eta_j \to 0^+$, we may find a further subsequence of the hypersurfaces (not relabeled) so that $\overline{\Sigma_j} \cap B_9(0)$ is a smooth immersion and that there is $z_j \in B_6(0)$ so that Σ_j satisfies the curvature estimate

$$|A_{\Sigma_j}|(x)|x - z_j| \le \eta_j \tag{51}$$

for all $x \in B_9(0)$. Note that since $z_j \in B_6(0)$, it follows from (51) that $|A_{\Sigma_j}|$ is uniformly bounded in the annulus $B_9(0) \setminus \overline{B}_8(0)$ and therefore the maximum of $|A_{\Sigma_j}|$ in $\Sigma_j \cap B_9(0)$ is achieved at a point $y_j \in B_8(0)$. We set

$$\tilde{\Sigma}_j = \lambda_j(\Sigma_j - y_j).$$

By construction we have $|A_{\tilde{\Sigma}_j}|(0) = 1$ and that $|A_{\tilde{\Sigma}_j}|$ is uniformly bounded on compact subsets of $\mathbb{R}^{n+1}$, so $\tilde{\Sigma}_j$ converges smoothly on compact subsets of $\mathbb{R}^{n+1}$ to a non-flat smooth hypersurface $\tilde{\Sigma}_\infty$. On the other hand, the estimate (51) is scale invariant, so for $\tilde{z}_j = \lambda_j(z_j - y_j)$, we see that

$$|A_{\tilde{\Sigma}_j}|(x)|x - \tilde{z}_j| \leq \eta_j.$$

Taking $x = 0$ in this, we find that $\tilde{z}_j \to 0$, since $\eta_j \to 0$. Hence, passing this inequality to the limit, we find that

$$|A_{\tilde{\Sigma}_\infty}|(x)|x| = 0,$$

contrary to the fact that $\tilde{\Sigma}_\infty$ is non-flat.

Acknowledgement

This article is an expanded version of the themes covered in the author's lectures at the 2017 Current Developments in Mathematics conference at Harvard University. It is a pleasure to thank the organizers Professors David Jerison, Mark Kisin, Paul Seidel, Richard Stanley, Horng-Tzer Yau and Shing-Tung Yau for their invitation to deliver the lectures and for their hospitality during the event.

References

[All72] W. K. Allard. *On the first variation of a varifold*, Ann. of Math. (2) 95 (1972), 417-491. MR 0307015

[Alm65] F. J. Almgren Jr. *Theory of varifolds* Mimeographed notes, Princeton University (1965).

[Alm66] F. J. Almgren, Jr. *Some interior regularity theorems for minimal surfaces and an extension of Bernstein's theorem.* Ann. of Math. **84** (1966), 277–292. MR 0200816

[Alm83] F. J. Almgren, Jr. *Almgren's big regularity paper: Q-valued functions minimizing Dirichlet's integral and the regularity of area minimizing rectifiable currents up to codimension two* World Scientific. MR 1777737

[AmbCab00] L. Ambrosio, X. Cabŕe. *Entire solutions of semilinear elliptic equations in* $\mathbb{R}^3$ *and a conjecture of De Giorgi.* Jour. Amer. Math. Soc., 13(4) (2000) 725–739. MR 1775735

[BarDoC84] J. L. Barbosa, M. do Carmo. *Stability of Hypersurfaces of Constant Mean Curvature* Math. Zeit. 185 (1984) 3 339-353. MR 0731682

[BDE88] J. L. Barbosa, M. do Carmo, J. Eschenburg. *Stability of Hypersurfaces of Constant Mean Curvature in Riemannian Manifolds* Math. Z. 197, 123-138 (1988). MR 0917854

[BelWic18] C. Bellettini, N. Wickramasekera. *Stable CMC integral varifolds of codimension* 1*: regularity and compactness*, arXiv 2018.

[BelWic-1] C. Bellettini, N. Wickramasekera. *Stable prescribed-mean-curvature integral varifolds of codimension* 1*: regularity and compactness*, in preparation.

[BCW18] C. Bellettini, O. Chodosh, N. Wickramasekera. *Curvature estimates and sheeting theorems for weakly stable CMC hypersurfaces*, arXiv 2018. MR 3959653

[Bir17] G. D. Birkhoff. *Dynamical systems with two degrees of freedom*, Trans. Amer. Math. Soc. 18 (1917), 199–300. MR 1501070

[Bra78] K. Brakke. *The motion of a surface by its mean curvature*, Princeton University Press (1978). MR 0485012

[Caf98] L. A. Caffarelli. *The obstacle problem revisited*, J. Fourier Anal. Appl. 4 (1998), 383-402. MR 1658612

[CHS84] L. A. Caffarelli, R. Hardt, L. Simon. *Minimal surfaces with isolated singularities* Manuscripta Math. 48 (1984) no. 1-3, 1-18. MR 0753722

[CSZ97] H.-D. Cao, Y. Shen, and S. Zhu. *The structure of stable minimal hypersurfaces in* $\mathbb{R}^{n+1}$, Math. Res. Lett. 4 (1997), no. 5, 637–644. MR 1484695

[ChaLio16] G. R. Chambers, Y. Liokumovich. *Existence of minimal hypersurfaces in complete manifolds of finite volume.* arXiv.

[CCZ08] X. Cheng, L.-F. Cheung, D. Zhou, *The structure of weakly stable constant mean curvature hypersurfaces*, Tohoku Math. J. (2) 60 (2008), no. 1, 101–121. MR 2419038

[CCE16] A. Carlotto, O. Chodosh and M. Eichmair. *Effective versions of the positive mass theorem*, Invent. Math. 206 (2016), no. 3, 975–1016. MR 3573977

[CKM17] O. Chodosh, D. Ketover, D. Maximo, *Minimal hypersurfaces with bounded index*, Invent. Math. 209 (2017), no. 3, 617–664. MR 3681392

[ChoMan] O. Chodosh, C. Mantoulidis. *Minimal surfaces and the Allen–Cahn equation on 3 manifolds: index, multiplicity and curvature estimates.* arXiv. (2018)

[ColDel03] T. H. Colding, C. De Lellis. *The min-max construction of minimal surfaces*, Surveys in Differential Geometry VIII (2003), 75-107. MR 2039986

[ColMin08-1] T. H. Colding, W. P. Minicozzi II. *Width and mean curvature flow* Geom. Topol. 12 (2008), no. 5, 2517-2535. MR 2460870

[ColMin08-2] T. H. Colding, W. P. Minicozzi II. *Width and finite extinction time of Ricci flow* Geom. Topol. 12 (2008), no. 5, 2537-2586. MR 2460871

[DaS87] A. M. Da Silveira. *Stability of complete noncompact surfaces with constant mean curvature*, Math. Ann. 277 (1987), no. 4, 629–638. MR 0901709

[DeG54] E. De Giorgi *Su una teoria generale della misura* $(r-1)$*-dimensionale in uno spazio ad* r *dimensioni* Ann. Mat. Pura Appl. (4) 36 (1954), 191-213. MR 0062214

[DeG55] E. De Giorgi. *Nuovi teoremi relativi alle misure* $(r-1)$*-dimensionali in uno spazio ad* r *dimensioni* Ricerche Mat. 4 (1955), 95-113. MR 0074499

[DeG61] E. De Giorgi. *Frontiere orientate di misura minima.* Sem. Mat. Scoula Norm. Sup. Pisa (1961), 1–56. MR 0179651

[DeLSpa14] C. De Lellis, E. N. Spadaro. *Regularity of area minimizing currents I: gradient Lp estimates.* Geom. Funct. Anal. 24 (2014), 1831–1884. MR 3283929

[DeLSpa16a] C. De Lellis, E. N. Spadaro. *Regularity of area minimizing currents II: center manifold.* Ann. of Math. 183 (2016), 499–575. MR 3450482

[DeLSpa16b] C. De Lellis, E. N. Spadaro. *Regularity of area minimizing currents III: blow-up.* Ann. of Math. 183 (2016), 577–617. MR 3450483

[DelTas13] C. De Lellis, D. Tasnady. *The existence of embedded minimal hypersurfaces* J. Differential Geom. Volume 95, Number 3 (2013), 355-388. MR 3128988

[DelMag17] M. Delgadino, F. Maggi. *Alexandrov's theorem revisited*, arXiv 2017. MR 3921314

[Dug84] J. Duggan $W^{2,p}$ *regularity for varifolds with mean curvature* Miniconference on nonlinear analysis (Canberra, 1984), Proc. Centre Math. Anal. Austral. Nat. Univ. (8) Austral. Nat. Univ. 101-106. MR 0799216

[EM12] M. Eichmair, J. Metzger. *On large volume preserving stable CMC surfaces in initial data sets*, J. Differential Geom. 91 (2012), no. 1, 81–102. MR 2944962

[EvaGar15] L. C. Evans, R. F. Gariepy *Measure theory and fine properties of functions*, Textbooks in Mathematics, CRC Press, Boca Raton, FL, 2015, xiv+299. MR 3409135

[FSV13] A. Farina, Y. Sire, E. Valdinoci. *Stable solutions of elliptic equations on Riemannian manifolds*, Jour. Geom. Anal., 23 (2013), 1158–1172. MR 3078347

[Fed70] H. Federer. *The singular sets of area minimizing rectifiable currents with codimension one and area minimizing flat chains modulo two with arbitrary codimension.* Bull. Amer. Math. Soc. **76** (1970), 767–771. MR 0260981

[FedFle60] H. Federer, W. Fleming. *Normal and integral currents*, Ann. of Math. 72 (1960), 458–520. MR 0123260

[Fle62] W. H. Fleming. *On the oriented Plateau problem.* Rend. Circ. Mat. Palermo (2) **11** (1962), 69–90. MR 0157263

[Gas] P. Gaspar. *The second inner variation of energy and the Morse index of limit interfaces.* arXiv (2017).

[GasGua1] P. Gaspa, M. Guaraco. *The Allen–Cahn equation on closed manifolds.* arXiv (2016). MR 3814054

[GasGua2] P. Gaspar, M. Guaraco. *The Weyl Law for the phase transition spectrum and density of limit interfaces.* arXiv (2018). MR 3945835

[GilTru] D. Gilbarg, N. Trudinger. *Elliptic Partial Differential Equations of Second Order*, Springer. MR 0473443

[Giu84] E. Giusti. *Minimal surfaces and functions of bounded variations*, Monographs in Mathematics, 80, Birkhäuser Verlag, Basel, 1984, xii+240. MR 0775682

[GMT83] E. Gonzalez, U. Massari, I. Tamanini. *On the regularity of boundaries of sets minimizing perimeter with a volume constraint* Indiana Univ. Math. J. 32 (1983), no. 1, 25-37. MR 0684753

[Gua15] M. Guaraco. *Min-max for phase transitions and the existence of embedded minimal hypersurfaces* arXiv 2015. MR 3743704

[HarSim79] R. Hardt, L. Simon. *Boundary regularity and embedded solutions for the oriented Plateau problem.* Ann. of Math. 110 (1979), 436-486. MR 0554379

[HS85] R. Hardt, L. Simon. *Area minimizing hypersurfaces with isolated singularities*, J. Reine Angew. Math. 362 (1985), 102–129. MR 0809969

[Hie] F. Hiesmayr. *Spectrum and index of two sided Allen–Cahn minimal hyperasurfaces.* Comm. PDE. (to appear). arXiv (2017). MR 3924215

[HutTon00] J. E. Hutchinson, Y. Tonegawa. *Convergence of phase interfaces in the van der Waals–Cahn–Hilliard theory* Calc. Var. 10 (2000), no. 1, 49-84. MR 1803974

[Ilm93] T. Ilmanen. *Convergence of the Allen–Cahn equation to Brakke's motion by mean curvature flow.* Jour. Diff. Geom. **38** (1993), 417–461. MR 1237490

[Ilm96] T. Ilmanen. *A strong maximum principle for singular minimal hypersurfaces*, Calc. Var. Partial Differential Equations 4 (1996), 443–467. MR 1402732

[IMN18] *Density of minimal hypersurfaces for generic metrics.* Ann. of Math. (2) 187 (2018), no. 3, 963–972. MR 3779962

[Irv17] C. Irving. *CMC hypersurfaces asymptotic to singular minimal cones*, Summer research project, University of Cambridge, 2017.

[Krum] B. Krummel. *Existence and regularity of multivalued solutions to elliptic equations and systems.* to appear in Comm. Anal. Geom., ArXiv.

[Krum14] B. Krummel. *Regularity of minimal hypersurfaces with a common free boundary.* Calc. Var. and PDE, 51 (2014), 525-537. MR 3268861

[KolMen15] S. Kolasiński, U. Menne. *Decay rates for the quadratic and super-quadratic tilt-excess of integral varifolds.* Nonlinear Diff. Equations and Appl. NoDEA 24 (2015). MR 3625810

[Le11] N. Le. *On the second inner variation of the Allen–Cahn functional and its applications.* Indiana Univ. Math. Jour. 60 (2011), 1843–1856. MR 3008253

[Le15] N. Le. *On the second inner variation of Allen–Cahn type energies and applications to local minimizers.* Jour. des Mat. Pures et Appl. 103 (2015), 1317–1345. MR 3343700

[LopRos89] F. J. López, A. Ros. *Complete minimal surfaces with index one and stable constant mean curvature surfaces*, Comment. Math. Helv. 64 (1989), no. 1, 34–43. MR 0982560

[Mag12] F. Maggi. *Sets of finite perimeter and geometric variational problems*, Cambridge Studies in Advanced Mathematics, 135, An introduction to geometric measure theory, Cambridge University Press, Cambridge, 2012, xx+454. MR 2976521

[MarNev14a] F. C. Marques, A. Neves. *Min-max theory and the Willmore conjecture*, Ann. of Math.179 (2014), 683–782. MR 3152944

[MarNev14b] F. C. Marques, A. Neves. *Existence of infinitely many minimal hypersurfaces in positive Ricci curvature.* Invent. Math. 209 (2017), no. 2, 577–616. MR 3674223

[MarNev16] F. C. Marques, A. Neves. *Morse index and multiplicity of min-max minimal hypersurfaces.* Cambridge Jour. Math. 4(4) (2016), 463–511. MR 3572636

[MPR16] W. H. Meeks, III, J. Pérez, A. Ros. *Local removable singularity theorems for minimal laminations*, J. Differential Geom. 103 (2016), no. 2, 319–362. MR 3504952

[Men13] U. Menne. *Second order rectifiability of integral varifolds of locally bounded first variation*, Jour. Geom. Anal. 23 (2013), 709–763. MR 3023856

[Mil63] J. Milnor. *Morse theory*, Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51, Princeton University Press, Princeton, N.J., 1963. MR 0163331

[Mir67] M. Miranda. *Sulle singolarit della frontiere minimali.* Rend. Sem. Mat. Univ. Padova. 38 (1967) 181–188. MR 0221351

[Mod87] L. Modica. *The gradient theory of phase transitions and the minimal interface criterion.* Arch. Rat. Mech. Anal. 98 (1987), 123–142. MR 0866718

[Mor03] F. Morgan. *Regularity of isoperimetric hypersurfaces in Riemannian manifolds* Trans. Amer. Math. Soc. 355 (2003), 5041-5052. MR 1997594

[Mos77] M. Moschen. *Principio di Massimo Forte per le Frontiere di Misura Minima.* Ann. Univ. Ferrara - Sez. VII -Sc. Mat. 23 (1977) 165–168. MR 0482508

[Pal86] B. W. Palmer. *Surfaces of constant mean curvature in space forms*, ProQuest LLC, Ann Arbor, MI, 1986, Thesis (Ph.D.)–Stanford University. MR 2635021

[PigRiv17] A. Pigati, T. Rivière. *The regularity of parametrized integer 2-rectifiable stationary varifolds*, arXiv 2017.

[Pit77] J. Pitts *Existence and regularity of minimal surfaces on Riemannian manifolds* Princeton University press (1977). MR 0626027

[Rei60] R. Reifenberg. *Solution of the Plateau problem for m-dimensional surfaces of varying topological type.* Acta Math. 104 (1960), 1–92. MR 0114145

[Riv15] T. Rivière. *A viscosity method in the min-max theory of minimal surfaces*, arXiv '15. MR 3735867

[Sam69] H. Samelson. *Orientability of hypersurfaces in R^n*, Proc. Amer. Math. Soc., 22, 1969, 301-302. MR 0245026

[SchSim81] R. Schoen, L. Simon. *Regularity of stable minimal hypersurfaces* Comm. Pure Appl. Math. 34 (1981), 741–797. MR 0634285

[Sch04] R. Schätzle. *Quadratic tilt-excess decay and strong maximum principle for varifolds.* Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), 3(1), (2004), 171–231. MR 2064971

[SSY75] R. Schoen, L. Simon, S.T. Yau. *Curvature estimates for minimal hypersurfaces.* Acta Math. **134** (1975), 276–288. MR 0423263

[Sim83] L. Simon. *Lectures on Geometric Measure Theory* Proceedings of the Centre for Mathematical Analysis 3, Canberra, (1984), VII+272. MR 0756417

[Sim87] L. Simon. *A strict maximum principle for area minimizing hypersurfaces.* Jour. Diff. Geom. 26 (1987), 327-335. MR 0906394

[Sim93] L. Simon. *Cylindrical tangent cones and the singular set of minimal submanifolds* Jour. Diff. Geom. 38 (1993), 585-652. MR 1243788

[Sim97] L. Simon. *Schauder estimates by scaling* Calc. Var. 5, 391-407 (1997) MR 1459795

[JSim68] J. Simons. *Minimal varieties in Riemannian mabifolds.* Annals of Math., 88(1968), 62–105. MR 0233295

[SimWic07] L. Simon, N. Wickrmasekera. *Stable branched minimal immersions with prescribed boundary.* Jour. Diff. Geom. 75 (2007), 143–173. MR 2282727

[Son18] A. Song. *Existence of infinitely many minimal hypersurfaces in closed manifolds.* arXiv.

[SW89] B. Solomon, B. White. *A strong maximum principle for varifolds that are stationary with respect to even parametric elliptic functionals.* Indiana Jour. Math. 38 (1989), 683–691. MR 1017330

[Ste88] P. Sternberg. *The effect of a singular perturbation on non convex variational problems.* Arch. Rat. Mech. Anal. 101 (1988), 209–260. MR 0930124

[Ton05] Y. Tonegawa. *On stable critical points for a singular perturbation problem* Comm. Anal. Geom. 13 (2005), 439-459. MR 2154826

[TonWic12] Y. Tonegawa, N. Wickramasekera. *Stable phase interfaces in the Van der Waals–Cahn–Hilliard theory* J. Reine Angew. Math. 668 (2012), 191-210. MR 2948876

[WanWei] K. Wang, J. Wei. *Finite Morse index implies finite ends.* Comm. Pure Appl. Math. (to appear), arXiv (2017). MR 3935478

[Whi87] B. White. *Curvature estimates and compactness theorems in 3-manifolds for surfaces that are stationary for parametric elliptic functionals*, Invent. Math. 88 (1987), no. 2, 243–256. MR 0880951

[Wic08] N. Wickramasekera. *A regularity and compactness theory for immersed stable minimal hypersurfaces of multiplicity at most 2*, J. Differential Geom. 80 (2008), no. 1, 79–173. MR 2434260

[Wic14a] N. Wickramasekera. *A general regularity theory for stable codimension 1 integral varifolds* Ann. of Math. 179 (2014), 843-1007 MR 3171756

[Wic14b] N. Wickramasekera. *A sharp strong maximum principle and a sharp unique continuation theorem for singular minimal hypersurfaces* Calc. Var. & PDE. 51 (2014), 799–812. MR 3268871

[Wic14c] N. Wickramasekera. *Regularity of stable minimal hypersurfaces: Recent advances in the theory and applications.* Surveys in Diff. Geom. 19 (2014), 231–301. MR 3381502

[Wic] N. Wickramasekera. *Stratification of stable codimension 1 rectifiable currents near points of density ¡ 3.* to appear.

[Ye96] R. Ye. *Constant mean curvature foliation: singularity structure and curvature estimate*, Pacific J. Math. 174 (1996), no. 2, 569–587. MR 1405602

[ZhoZhu17] X. Zhou, J. Zhu. *Min-max theory for constant mean curvature hypersurfaces*, arXiv 2017.