Existence of immersed spheres minimizing curvature functionals in compact 3-manifolds

Ernst Kuwert, Andrea Mondino, Johannes Schygulla

Abstract

We study curvature functionals for immersed 2-spheres in a compact, three-dimensional Riemannian manifold M. Under the assumption that the sectional curvature K^M is strictly positive, we prove the existence of a smooth immersion $f : \mathbb{S}^2 \to M$ minimizing the L^2 integral of the second fundamental form. Assuming instead that $K^M \leq 2$ and that there is some point $\overline{x} \in M$ with scalar curvature $R^M(\overline{x}) > 6$, we obtain a smooth minimizer $f : \mathbb{S}^2 \to M$ for the functional $\int \frac{1}{4}|H|^2 + 1$, where H is the mean curvature.

Key Words: L^2 second fundamental form, Willmore functional, direct methods in the calculus of variations, geometric measure theory, elliptic regularity theory.

AMS subject classification:
53C21, 53C42, 58E99, 35J60

1 Introduction

Let M be a three-dimensional, compact Riemannian manifold with metric h. For any immersed closed surface $f : \Sigma \hookrightarrow M$ with induced metric $g = f^*h$ and second fundamental form A, we consider the functional

$$E(f) = \frac{1}{2} \int_\Sigma |A|^2 \, d\mu_g. \tag{1}$$

We denote by H the mean curvature vector and by A° the tracefree component of A. The extrinsic curvature is related to the intrinsic curvature, i.e. the sectional curvature K_g of the induced metric and the sectional curvature K_f^M of the tangent plane in TM, by the Gauß equation

$$\frac{1}{4}|H|^2 - \frac{1}{2}|A^\circ|^2 = \frac{1}{2}\big(|H|^2 - |A|^2\big) = K_g - K_f^M. \tag{2}$$

Integrating and using the Gauß-Bonnet theorem yields the well-known identities

$$\frac{1}{4}\int_\Sigma |H|^2 \, d\mu_g + \frac{1}{2}\int_\Sigma |A^\circ|^2 \, d\mu_g = \frac{1}{2}\int_\Sigma |A|^2 \, d\mu_g = \frac{1}{2}\int_\Sigma |H|^2 \, d\mu_g + \int_\Sigma K_f^M \, d\mu_g - 2\pi\chi(\Sigma), \tag{3}$$

where $\chi(\Sigma)$ is the Euler characteristic. For $M = \mathbb{R}^3$ the functional E reduces to the classical Willmore energy given by

$$W(f) = \frac{1}{4}\int_\Sigma |H|^2 \, d\mu_g, \tag{4}$$

more precisely we have $E(f) = 2W(f) - 2\pi\chi(\Sigma)$. In [Will] Willmore proved the inequality $W(f) \geq 4\pi$ for all $f : \Sigma \to \mathbb{R}^3$, with equality only for the round spheres.

In the present paper we study the problem of minimizing $E(f)$ in the class of immersed spheres in the Riemannian manifold M. Any totally geodesic $f : \mathbb{S}^2 \to M$ is trivially a minimizer, but totally geodesic immersions do not always exist. For instance, there are no totally umbilic surfaces in the Berger spheres (except $\mathbb{S}^3$), see [ST]. For appropriate parameters, these spheres have positive sectional curvature [Dan]. We prove the following existence result.

Theorem 1.1. *Let M be a compact, 3-dimensional Riemannian manifold. On the class $[\mathbb{S}^2, M]$ of smooth immersions $f : \mathbb{S}^2 \to M$, consider the functional*

$$E : [\mathbb{S}^2, M] \to \mathbb{R},\ E(f) = \frac{1}{2} \int_{\mathbb{S}^2} |A|^2 \, d\mu_g.$$

If M has sectional curvature $K^M > 0$, then there exists a minimizer f in $[\mathbb{S}^2, M]$ for E.

We remark that our proof actually needs only the two conditions that $\inf_{[\mathbb{S}^2,M]} E(f) < 4\pi$ and that the area is bounded along some minimizing sequence. We always have $\inf_{[\mathbb{S}^2,M]} E(f) \leq 4\pi$, since the energy goes to 4π for a sequence of distance spheres shrinking to a point. Moreover, the strict inequality is necessary to rule out such a minimizing sequence. For example, if M has strictly negative sectional curvature then $E(f) > 4\pi$ for any sphere immersed into M by equation (3), and the infimum is not attained. Of course, the boundedness of the area along the minimizing sequence is also necessary, at least if we want subconvergence of the surface measures. The first condition will be settled using a local expansion around a point with strictly positive scalar curvature. The strong curvature assumption $K^M > 0$ of Theorem 1.1 is used to obtain the upper area bound. Possibly, the situation when the area actually goes to infinity (in the case when K^M is not strictly positive) can be studied using results of Hutchinson [Hu1] on curvature varifolds, see also [MonVar].

In asymptotically flat 3-manifolds M, spheres which are critical points of related curvature functionals have been constructed recently by Mondino [Mon1, Mon2] and Lamm, Metzger & Schulze [LMS], see also [LM]. They obtain the solutions as perturbations of round spheres using implicit function type arguments. In [SiL] L. Simon proved the existence of an embedded torus in $\mathbb{R}^n$, which minimizes the classical Willmore functional. Our approach implements his fundamental theory in the case of spheres immersed into the Riemannian manifold M. Recently, an alternative approach to Simon's theorem was developed by Rivière [Riv].

We now briefly outline the contents of the paper. In Section 2, we gain some global control in terms of area and diameter bounds. For the lower diameter bound we use the bound $\inf E < 4\pi$ mentioned above. Local area bounds are then obtained by adapting Simon's monotonicity formula [SiL]. In Section 3 we prove Theorem 1.1. First we obtain a limiting measure as a candidate for the minimizer. Adapting the arguments of [SiL] to the Riemannian situation, we establish $C^{1,\alpha} \cap W^{2,2}$ regularity away from a finite set of bad points where the curvature significantly concentrates. If a closed surface in $\mathbb{R}^3$ has Willmore energy below 8π, as is the case in [SiL], then the area ratio is bounded below two by the monotonicity formula. Unfortunately, this involves a global argument which does not generalize immediately to our situation in M. We rule out the formation of branch points using the global bound $\inf E < 4\pi$. This step involves a degree argument for the Gauß map, which does not extend to higher codimension. Eventually, we exclude all bad points and finally prove smoothness. The fact that the limiting measure comes from an immersed sphere is proved using a compactness result of Breuning [Breu].

In the final section 4 we discuss the following variant of Theorem 1.1.

Theorem 1.2. *For a closed, three-dimensional Riemannian manifold M, consider on the class of immersions $f : \mathbb{S}^2 \to M$ the functional*

$$W_1(f) = \int_{\mathbb{S}^2} \Big(\frac{1}{4}|H|^2 + 1\Big)\, d\mu_g.$$

If M has sectional curvature $K^M \leq 2$ and moreover the scalar curvature $R^M(\overline{x}) > 6$ for some point $\overline{x} \in M$, then there exists a smooth minimizer f in $[\mathbb{S}^2, M]$ for W_1.

We remark that the curvature conditions in Theorem 1.2 can be fulfilled, for instance they hold for a round sphere $\mathbb{S}^3(R)$ if $\frac{1}{\sqrt{2}} \leq R < 1$. One motivation to study the functional is that if we transform the classical Willmore integral from $\mathbb{R}^3$ to $\mathbb{S}^3$ using stereographic projection, then we obtain the functional $W_1(f)$. Moreover, minimal surfaces are obvious critical points of W_1. The existence of minimizers for the functional $\int |H|^2 \, d\mu_g$, possibly with branch points, was proposed in [SiProc]. In our theorem, the assumption $K^M \leq 2$ in Theorem 1.2 is mainly used to rule out the branch points.

Acknowledgments

The second author would like to thank his supervisor Prof. A. Malchiodi for proposing to study the Willmore functional in Riemannian manifolds, and for his constant support. All authors acknowledge the support by the DFG Collaborative Research Center SFB/Transregio 71 and of M.U.R.S.T, within the project B-IDEAS "Analysis and Beyond", making possible our cooperation with mutual visits at Freiburg and Trieste.

2 Global bounds for the minimizing sequence

Here we collect some basic information for minimizing sequences of the functional E: global and local upper area bounds and a lower diameter bound. The first observation, following directly from (3), is

Proposition 2.1. *Let M be a compact Riemannian 3-manifold with sectional curvature $K^M > 0$. Then, for any immersed, closed surface $f : \Sigma \hookrightarrow M$ the total area $\mu_g(\Sigma)$ is bounded by*

$$\mu_g(\Sigma) \leq C\left(E(f) + 2\pi\chi(\Sigma)\right) \quad \text{with } C = \frac{1}{\min_M K^M} < \infty. \tag{5}$$

We next apply Simon's monotonicity formula in $\mathbb{R}^m$ to show a local, quadratic area bound.

Lemma 2.2. *Let $f : \Sigma \hookrightarrow M$ be a closed immersed surface in a compact 3-manifold, with*

$$W(f) + \mu_g(\Sigma) \leq \Lambda \quad \text{for some } \Lambda < \infty.$$

Then for any $x \in M$, $\rho > 0$ we have an estimate

$$\mu_g\big(\{p \in \Sigma : f(p) \in B_\rho(x)\}\big) \leq C\rho^2, \quad \text{where } C = C(\Lambda, M).$$

Proof. By Nash's theorem, there is an isometric embedding $I : M \hookrightarrow \mathbb{R}^m$ for some $m \in \mathbb{N}$. The second fundamental forms of f, $I \circ f$ and I are related by the formula

$$A^{I\circ f}(\cdot,\cdot) = DI|_f \circ A^f(\cdot,\cdot) \oplus (A^I \circ f)(Df, Df).$$

Taking the trace and squaring yields for an orthonormal basis $v_i = Df \cdot e_i$

$$|H^{I\circ f}|^2 = |H^f|^2 + \Big|\sum_{i=1}^2 A^I \circ f(v_i, v_i)\Big|^2 \leq |H^f|^2 + 2|A^I|^2 \circ f.$$

Integrating we see that $W(I \circ f) \leq W(f) + C\,\mu_g(\Sigma)$ where $C = \frac{1}{2}\max|A^I|^2$. Thus for any $x \in M$, we get from Simon's monotonicity formula, see (1.3) in [SiL],

$$\mu_g\big(\{p \in \Sigma : f(p) \in B^M_\rho(x)\}\big) \leq \mu_g\big(\{p \in \Sigma : I(f(p)) \in B^{\mathbb{R}^m}_\rho\big(I(x)\big)\}\big) \leq C\rho^2,$$

with constant C depending on $W(f)$, $\mu_g(\Sigma)$ and on $\max|A^I|$. ■

Next we state an asymptotic expansion for the energy E on geodesic spheres around a point $x \in M$, which follows from the well-known expansion of the metric in exponential coordinates. Since $|A|^2 = |A^\circ|^2 + \frac{1}{2}|H|^2$, we may combine Proposition 3.1 in [Mon1] with Lemma 3.5 and Proposition 3.8 in [Mon2] to get the result. Note that for $M = \mathbb{R}^3$ we always have $E(f) \geq 4\pi$, with equality only for round spheres, by [Will].

Lemma 2.3. *Let M be a 3-dimensional Riemannian manifold. Then for geodesic spheres $S_\rho(x) = \{y \in M : \mathrm{dist}(y,x) = \rho\}$ around $x \in M$ we have the expansion*

$$E\big(S_\rho(x)\big) = 4\pi - \frac{2\pi}{3} R^M(x)\rho^2 + \mathcal{O}(\rho^3) \qquad \text{as } \rho \searrow 0.$$

In particular, if the scalar curvature $R^M(\overline{x}) > 0$ for some $\overline{x} \in M$, then $\inf_{f\in[\mathbb{S}^2,M]} E(f) < 4\pi$.

At several points in this paper we work in local normal coordinates. The following lemma collects the relevant inequalities between the Riemannian and the coordinate quantities.

Lemma 2.4. *Let $h_{1,2}$ be Riemannian metrics on a manifold M, with norms satisfying*

$$(1+\varepsilon)^{-1}\|\cdot\|_1 \leq \|\cdot\|_2 \leq (1+\varepsilon)\|\cdot\|_1 \qquad \text{for some } \varepsilon \in (0,1].$$

For any smooth immersed surface $f : \Sigma \to M$, the following inequalities hold with universal $C < \infty$:

(i) $\mathrm{dist}_1(x,y) \leq (1+\varepsilon)\mathrm{dist}_2(x,y)$ *for all* $x, y \in M$*;*

(ii) $B^{h_1}_\sigma(x) \subset B^{h_2}_\rho(x)$*, whenever* $(1+\varepsilon)\sigma \leq \rho$*;*

(iii) $\mu_{g_1} \leq (1+C\varepsilon)\mu_{g_2}$*, where* $g_{1,2} = f^*(h_{1,2})$*;*

(iv) $\|A_1\|_1^2 \leq \big(1 + C(\varepsilon+\delta)\big)\|A_2\|_2^2 + C\delta^{-1}\|\Gamma\|^2_{h_1} \circ f$ *for any* $\delta \in (0,1]$*, where* $\Gamma := D^{h_1} - D^{h_2}$ *and* D^{h_i} *is the covariant derivative with respect to the metric* h_i*.*

Proof. The statements (i) and (ii) are obvious. To compare the Jacobians of f with respect to $h_{1,2}$, we use $\|\cdot\|_{g_1} \leq (1+\varepsilon)\|\cdot\|_{g_2}$ and compute for $v, w \in T_p\Sigma$ with $g_2(v,w) = 0$

$$\|v \wedge w\|^2_{g_1} = \|v\|^2_{g_1}\|w\|^2_{g_1} - g_1(v,w)^2 \leq (1+\varepsilon)^4\|v\|^2_{g_2}\|w\|^2_{g_2} = (1+\varepsilon)^4\|v\wedge w\|^2_{g_2}.$$

This proves the inequality (iii). Next we compare the norms for a bilinear map $B : T_p\Sigma \times T_p\Sigma \to T_{f(p)}M$. Choose a basis v_α of $T_p\Sigma$ such that $g_1(v_\alpha, v_\beta) = \delta_{\alpha\beta}$ and $g_2(v_\alpha, v_\beta) = \lambda_\alpha\delta_{\alpha\beta}$. Then

$$\lambda_\alpha = \|v_\alpha\|_{g_2} \leq (1+\varepsilon)\|v_\alpha\|_{g_1} = 1+\varepsilon,$$

and putting $w_\alpha = v_\alpha/\lambda_\alpha$ we obtain

$$\|B\|_1^2 = \sum_{\alpha,\beta=1}^{2} \lambda_\alpha^2\lambda_\beta^2\|B(w_\alpha,w_\beta)\|^2_{h_1} \leq (1+C\varepsilon)\sum_{\alpha,\beta=1}^{2}\|B(w_\alpha,w_\beta)\|^2_{h_2} = (1+C\varepsilon)\|B\|_2^2.$$

Now denote by $P^\perp_{1,2} : T_{f(p)}M \to (T_pf)^{\perp_{h_{1,2}}}$ the orthogonal projections onto the normal spaces with respect to $h_{1,2}$. Then we have for any $\delta > 0$ the estimate

$$\begin{aligned}
\big\|A_1\big\|_1^2 &= \big\|P_1^\perp D^{h_1}(Df)\big\|_1^2 \\
&\leq \big\|P_2^\perp D^{h_1}(Df)\big\|_1^2 \\
&\leq \big\|P_2^\perp\big(D^{h_2}(Df) + \Gamma\circ f(Df,Df)\big)\big\|_1^2 \\
&\leq (1+\delta)\big\|P_2^\perp D^{h_2}(Df)\big\|_1^2 + C\delta^{-1}\|\Gamma\|^2_{h_1}\circ f \\
&\leq (1+\delta)(1+C\varepsilon)\|A_2\|_2^2 + C\delta^{-1}\|\Gamma\|^2_{h_1}\circ f.
\end{aligned}$$

This proves the inequality (iv). ■

The lower diameter bound follows by combining Proposition 2.1, Lemma 2.3 and the following fact.

Proposition 2.5. *Let M be a compact Riemannian 3-manifold. Assume there is a minimizing sequence $f_k \in [\mathbb{S}^2, M]$ for $E(f)$ with $\mathrm{diam}\, f_k(\mathbb{S}^2) \to 0$ and $\mu_{g_k}(\Sigma) \leq C$. Then*

$$\inf_{f\in[\mathbb{S}^2,M]} E(f) = 4\pi.$$

Proof. After passing to a subsequence, we may assume that the $f_k(\mathbb{S}^2)$ converge to a point $x_0 \in M$. For given $\varepsilon \in (0,1]$ we choose $\rho > 0$, such that in Riemann normal coordinates $x \in B_\rho(0) \subset \mathbb{R}^3$

$$\frac{1}{1+\varepsilon}|\cdot| \leq \|\cdot\|_h \leq (1+\varepsilon)|\cdot| \quad \text{and} \quad |\Gamma_{ij}^k(x)| \leq \varepsilon.$$

We have $f_k(\mathbb{S}^2) \subset B_\rho(x_0)$ for large k. Denoting by A^e, g_k^e the quantities with respect to the coordinate metric, we get from Willmore's inequality and Lemma 2.4

$$4\pi \leq \frac{1}{2}\int_\Sigma |A^e_{f_k}|_e^2 \, d\mu_{g_k^e} \leq (1+C\varepsilon)(1+\delta)\frac{1}{2}\int_\Sigma |A_{f_k}|^2 \, d\mu_{g_k} + C(\delta)\varepsilon^2 \, \mu_{g_k}(\Sigma).$$

Since $\mu_{g_k}(\Sigma) \leq C$ by assumption, we may let first $k \to \infty$, then $\varepsilon \searrow 0$ and finally $\delta \searrow 0$ to obtain

$$\liminf_{k\to\infty} E(f_k) \geq 4\pi.$$

As the upper bound follows from Lemma 2.3, the lemma is proved. ■

Lemma 2.6. *Let $f : \Sigma \to M$ be a closed immersed surface in a compact 3-manifold, and put $\Sigma_\rho(x_0) = f^{-1}(B_\rho(x_0))$ for $x_0 \in M$ and $\rho > 0$. There exist constants $\rho_0 > 0$ and $C < \infty$ depending only on M, such that for $x_0 \in f(\Sigma)$ we have*

$$\frac{\mu_g(\Sigma_\sigma(x_0))}{\sigma^2} \leq C\Big(\frac{\mu_g(\Sigma_\rho(x_0))}{\rho^2} + \int_{\Sigma_\rho(x_0)} |H|^2 \, d\mu_g\Big) \quad \text{whenever } 0 < \sigma \leq \rho \leq \rho_0.$$

Proof. Again, we use an isometric embedding $I : M \to \mathbb{R}^m$. For $x_0 \in M$ we put

$$\Sigma_\rho^{\mathbb{R}^m}(x_0) = (I \circ f)^{-1}\big(B_\rho^{\mathbb{R}^m}(I(x_0))\big).$$

Choosing $\rho_0 > 0$ appropriately, we have $I(B_\rho(x_0)) \subset \big(B_\rho^{\mathbb{R}^m}(I(x_0)) \cap I(M)\big) \subset I(B_{2\rho}(x_0))$ and hence

$$\Sigma_\rho(x_0) \subset \Sigma_\rho^{\mathbb{R}^m}(x_0) \subset \Sigma_{2\rho}(x_0).$$

Now from [SiL], we obtain for $0 < \sigma \leq \rho/2 \leq \rho_0$ the estimate

$$\begin{aligned}
\frac{\mu_g(\Sigma_\sigma(x_0))}{\sigma^2} &\leq \frac{\mu_g(\Sigma_\sigma^{\mathbb{R}^m}(x_0))}{\sigma^2} \\
&\leq C\Big(\frac{\mu_g(\Sigma_{\rho/2}^{\mathbb{R}^m}(x_0))}{\rho^2} + \int_{\Sigma_{\rho/2}^{\mathbb{R}^m}(x_0)} |H^{I\circ f}|^2 \, d\mu_g\Big) \\
&\leq C\frac{\mu_g(\Sigma_\rho(x_0))}{\rho^2} + C\int_{\Sigma_\rho(x_0)} |H^f|^2 \, d\mu_g + C\max|A^I|^2 \mu_g(\Sigma_\rho(x_0)) \\
&\leq C(1+\rho_0^2\max|A^I|^2)\frac{\mu_g(\Sigma_\rho(x_0))}{\rho^2} + C\int_{\Sigma_\rho(x_0)} |H^f|^2 \, d\mu_g.
\end{aligned}$$

This settles the inequality, if $\rho \geq 2\sigma$. As the claim is trivial for $\rho \in [\sigma, 2\sigma]$, the lemma is proved. ■

Lemma 2.7. *Let M be a Riemannian 3-manifold, and $f : \Sigma \hookrightarrow M$ a closed immersed surface with*

$$W(f) + \mu_g(\Sigma) \leq \Lambda \quad \text{for some } \Lambda < \infty.$$

For any $\eta > 0$ there exist $\rho_0 = \rho_0(M,\eta) > 0$ and $C = C(M,\Lambda) < \infty$, such that for any $x_0 \in M$, $x \in B_{\rho_0}(x_0)$ and $0 < \rho \leq \rho_0$ the following inequalities hold, where $B^e, g^e, \ldots$ are defined with respect to normal coordinates centered at x_0:

$$B_\sigma(x) \subset B_\rho^e(x),\ B_\sigma^e(x) \subset B_\rho(x) \quad \text{if } (1+\eta)\sigma \leq \rho; \tag{6}$$

(7) $$\frac{1}{1+\eta}\mu_{g^e}(\Sigma_\rho(x)) \le \mu_g(\Sigma_\rho(x)) \le (1+\eta)\mu_{g^e}(\Sigma_\rho(x));$$

(8) $$\frac{1}{1+\eta}\int_{\Sigma_\rho(x)} |A_e|_e^2\, d\mu_{g^e} - C\rho^2 \le \int_{\Sigma_\rho(x)} |A|^2\, d\mu_g \le (1+\eta)\int_{\Sigma_\rho(x)} |A_e|_e^2\, d\mu_{g^e} + C\rho^2.$$

Proof. We can assume that the assumption of Lemma 2.4 is satisfied on $B_{2\rho_0}(x_0)$ with $\varepsilon = C(M)\rho_0^2$. The first two statements follow directly from that lemma. For (8) we choose $\delta = \rho_0^2$ in Lemma 2.4. Using $\|\Gamma\|_e \le C\rho_0$, the statement follows by combining with Lemma 2.2. ∎

3 Proof of Theorem 1.1

For proving existence of a minimizer for the functional $E : [\mathbb{S}^2, M] \to \mathbb{R}$, $E(f) = \frac{1}{2}\int_{\mathbb{S}^2} |A|^2\, d\mu_g$, we use the direct method in the calculus of variations. Let $f_k : \mathbb{S}^2 \hookrightarrow M$ be a minimizing sequence of immersed closed surfaces for the functional E. Denote by μ_k the Radon measure on M given by

(9) $$\mu_k(E) = \mu_{g_k}\left(f_k^{-1}(E)\right) = \int_E \theta_{f_k}(y)\, d\mathcal{H}^2(y),$$

where θ_{f_k} is the multiplicity and g_k is the induced metric.

By Proposition 2.1 we can assume

(10) $$\mu_k \to \mu \quad \text{weakly as Radon measures.}$$

Using this convergence and the monotonicity formula Lemma 2.6, it follows as in [SiL] that

(11) $$\operatorname{spt}\mu_k \to \operatorname{spt}\mu \quad \text{in the Hausdorff distance sense.}$$

This Hausdorff convergence, together with Lemma 2.3 and Proposition 2.5, implies that

(12) $$\operatorname{diam}_h(\operatorname{spt}\mu) \ge \liminf_k(\operatorname{diam}_h \operatorname{spt}\mu_k) > 0.$$

When working in normal coordinates, we denote the Euclidean coordinate quantities with an index $''e''$, for example μ_k^e, H_k^e, A_k^e, ..., while the Riemannian quantities won't have any index.

In order to prove regularity, we would like to apply Simon's Graphical Decomposition Lemma proved in [SiL]. The most important assumption is that the L^2-norm of the second fundamental form is locally small, which we will need simultaneously for infinitely many k. Therefore we define the so called bad points with respect to a given $\varepsilon > 0$ in the following way: Define the Radon measures α_k on M by

$$\alpha_k = \mu_k \llcorner |A_k|^2.$$

Since $\alpha_k(M) \le C$, there exists a Radon measure α on M such that (after passing to a subsequence) $\alpha_k \to \alpha$ weakly as Radon measures. It follows that $\operatorname{spt}\alpha \subset \operatorname{spt}\mu$ and $\alpha(M) \le C$.

Definition 3.1. *We define the bad points with respect to $\varepsilon > 0$ by*

$$\mathcal{B}_\varepsilon = \left\{\xi \in \operatorname{spt}\mu \,\middle|\, \alpha(\{\xi\}) > \varepsilon^2\right\}.$$

Remark 3.2. *Since $\alpha(M) \le C$, there exist only finitely many bad points. Moreover if $\xi_0 \in \operatorname{spt}\mu \setminus \mathcal{B}_\varepsilon$, there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ such that $\alpha(B_{\rho_0}(\xi_0)) < 2\varepsilon^2$, and since $\alpha_k \to \alpha$ weakly we get*

(13) $$\int_{B_{\rho_0}(\xi_0)} |A_k|^2\, d\mu_k \le 2\varepsilon^2 \quad \textit{for } k \textit{ sufficiently large.}$$

From now on fix a point $\xi_0 \in \operatorname{spt}\mu \setminus \mathcal{B}_\varepsilon$ and choose normal coordinates around that point. In the following we will work in these fixed coordinates. Using the estimates in normal coordinates in Lemma 2.7 as well as Lemma 2.2, the next Lemma is easily derived.

Lemma 3.3. *For $\varepsilon \le \varepsilon_0$ there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ and a $\beta = \beta(M) > 0$, such that for all $\xi \in \operatorname{spt}\mu \cap B^e_{\frac{\rho_0}{2}}(\xi_0)$, for all $\rho \le \frac{\rho_0}{4}$ and infinitely many k*

$$i) \quad \mu^e_k(\overline{B^e_\rho(\xi)}) \le \beta\rho^2,$$
$$ii) \quad \int_{B^e_\rho(\xi)} |A^e_k|^2 d\mu^e_k \le 3\varepsilon^2.$$

Thanks to Lemma 3.3 we are in position to apply the Graphical Decomposition Lemma of Leon Simon (Lemma 2.1 in [SiL]).

Lemma 3.4. *For $\varepsilon \le \varepsilon_0$ there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ such that for all $\xi \in \operatorname{spt}\mu \cap B^e_{\frac{\rho_0}{2}}(\xi_0)$, all $\rho \le \frac{\rho_0}{4}$ and for infinitely many k the following holds: There exist 2-dimensional planes L_l, $1 \le l \le M_k$, containing ξ; functions $u^l_k \in C^\infty(\overline{\Omega^l_k}, L^\perp_l)$, where $\Omega^l_k = (B^e_\lambda(\xi) \cap L^l) \setminus \bigcup_m d^l_{k,m}$ with $\lambda > \frac{\rho}{4}$ and where the sets $d^l_{k,m} \subset L^l$ are pairwise disjoint closed discs disjoint from $\partial B^e_\lambda(\xi)$; sets $P^{k,l}_j \subset M$, $1 \le j \le N_{k,l}$, which are diffeomorphic to discs and disjoint from $\operatorname{graph} u^l_k$; and open, connected sets $U^l_k \subset f^{-1}_k(B^e_{\frac{\rho}{4}}(\xi))$, such that*

$$(i) \quad D^l_k := \operatorname{graph} u^l_k \cup \bigcup_{j=1}^{N_{k,l}} P^{k,l}_j \quad \textit{is a topological disc},$$

$$(ii) \quad f_k(U^l_k) = D^l_k \cap B^e_{\frac{\rho}{4}}(\xi) = \Big(\operatorname{graph} u^l_k \cup \bigcup_{j=1}^{N_{k,l}} P^{k,l}_j\Big) \cap B^e_{\frac{\rho}{4}}(\xi),$$

$$(iii) \quad f^{-1}_k(B^e_{\frac{\rho}{4}}(\xi)) \ \textit{is the disjoint union of the sets}\ U^l_k.$$

Moreover the following estimates hold:

$$M_k \le c = c(M), \tag{14}$$

$$\sum_{l,m} \operatorname{diam} d^l_{k,m} + \sum_{l,j} \operatorname{diam} P^{k,l}_j \le c \left(\int_{B^e_\rho(\xi)} |A^e_k|^2 \, d\mu^e_k\right)^{\frac{1}{4}} \rho \le c\varepsilon^{\frac{1}{2}}\rho, \tag{15}$$

$$||u^l_k||_{L^\infty(\Omega^l_k)} \le c\varepsilon^{\frac{1}{6}}\rho + \delta_k, \quad ||Du^l_k||_{L^\infty(\Omega^l_k)} \le c\varepsilon^{\frac{1}{6}} + \delta_k, \quad \textit{where } \delta_k \to 0. \tag{16}$$

Next we prove a lower 2-density bound for the minimizing sequence f_k away from the bad points, which we will need later.

Proposition 3.5. *For $\varepsilon \le \varepsilon_0$ there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ and a constant $C = C(M) > 0$, such that for all $\xi \in \operatorname{spt}\mu \cap B_{\rho_0}(\xi_0)$ and all $\rho \le \rho_0$*

$$\frac{\mu(B_\rho(\xi))}{\rho^2} \ge C.$$

Proof. Let $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ as in Remark 3.2 and $\xi \in B_{\frac{\rho_0}{2}}(\xi_0)$. It follows that $B_{\frac{\rho_0}{2}}(\xi) \subset B_{\rho_0}(\xi_0)$. Choose according to the Hausdorff distance sense convergence a sequence $\xi_k \in \operatorname{spt}\mu_k$ such that $\xi_k \to \xi$. Therefore for given $\rho \le \rho_0$ and k sufficiently large it follows that $B_{\frac{\rho}{4}}(\xi_k) \subset B_{\frac{\rho}{2}}(\xi) \subset B_{\rho_0}(\xi_0)$. Since the norm of the mean curvature can be estimated by the norm of the second fundamental form, we get from (13) for k sufficiently large

$$\int_{B_{\frac{\rho}{4}}(\xi_k)} |H_k|^2 \, d\mu_k \leq c \int_{B_{\rho_0}(\xi_0)} |A_k|^2 \, d\mu_k \leq c\varepsilon^2.$$

By letting $\sigma \to 0$ in Lemma 2.6, it follows that

$$1 \leq C \left(\frac{\mu_k(B_{\frac{\rho}{4}}(\xi_k))}{\rho^2} + \varepsilon^2 \right).$$

Choosing $\varepsilon_0^2 \leq \frac{1}{2C}$ we get for k sufficiently large $\frac{\mu_k(B_{\frac{\rho}{2}}(\xi))}{\rho^2} \geq C > 0$, and the rest of the Proposition follows from the weak convergence $\mu_k \to \mu$. ■

In the next step we estimate the L^2-norm of the second fundamental form on small balls around the "good points". This estimate will help us to show that the candidate minimizer μ is actually the measure associated to $C^{1,\alpha} \cap W^{2,2}$-graphs in a neighborhood around the good points.

Lemma 3.6. *For $\varepsilon \leq \varepsilon_0$ there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ such that for all $\xi \in \operatorname{spt} \mu \cap B^e_{\frac{\rho_0}{2}}(\xi_0)$ and all $\rho \leq \frac{\rho_0}{4}$*

$$\liminf_{k \to \infty} \int_{B^e_{\frac{\rho}{8}}(\xi)} |A^e_k|^2 \, d\mu^e_k \leq c\rho^\alpha,$$

where $c < \infty$ and $\alpha \in (0,1)$ only depend on the manifold M.

PROOF. Let $\varepsilon \leq \varepsilon_0$ such that Lemma 3.3 and Lemma 3.4 hold. Let $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ as before and apply the Graphical Decomposition Lemma for $\rho \leq \frac{\rho_0}{4}$ given by Lemma 3.4 to infinitely many k. For these k (surface index), $l \in \{1, \ldots, M_k\}$ (slice index) and $\gamma \in \left(\frac{\rho}{16}, \frac{3\rho}{32}\right)$ define the set

$$C^l_\gamma(\xi) = \left\{ x + y \,\middle|\, x \in B^e_\gamma(\xi) \cap L_l, y \in L_l^\perp \right\}.$$

From the estimates for the diameters of the pimples $P^{k,l}_j$ and the C^1-estimates for the graph functions u^l_k, it follows that

$$D^l_k \cap C^l_\gamma(\xi) = D^l_k \cap C^l_\gamma(\xi) \cap \overline{B^e_{\frac{\rho}{4}}(\xi)} \quad \text{for } \varepsilon \leq \varepsilon_0 \text{ and } \delta_k \text{ sufficiently small.} \tag{17}$$

Next define the set A^l_k by

$$A^l_k(\xi) = \left\{ \gamma \in \left(\frac{\rho}{16}, \frac{3\rho}{32} \right) \,\middle|\, \partial C^l_\gamma(\xi) \cap \bigcup_j P^{k,l}_j = \emptyset \right\}.$$

For $\varepsilon \leq \varepsilon_0$ it follows that

$$\mathcal{L}^1(A^l_k(\xi)) \geq \frac{\rho}{32} - \sum_j \operatorname{diam} P^{k,l}_j \geq \frac{\rho}{32} - c\varepsilon^{\frac{1}{2}}\rho \geq \frac{\rho}{64}.$$

From Lemma 5.2 it follows that there exists a set $T_l \subset \left(\frac{\rho}{16}, \frac{3\rho}{32}\right)$ with $\mathcal{L}^1(T_l) \geq \frac{\rho}{64}$, such that for all $\gamma \in T_l$

$$\partial C^l_\gamma(\xi) \cap \bigcup_j P^{k,l}_j = \emptyset \quad \text{for infinitely many } k.$$

Now let $\gamma \in T_l$ be arbitrary (it will be chosen later). We apply the Extension Lemma 5.1 given in the Appendix to get a function $w_k^l \in C^\infty\left(\overline{B_\gamma^e(\xi)} \cap L_l, L_l^\perp\right)$ such that

$$\begin{aligned} w_k^l = u_k^l \quad &, \quad \frac{\partial w_k^l}{\partial \nu} = \frac{\partial u_k^l}{\partial \nu} \quad \text{on } \partial B_\gamma^e(\xi) \cap L_l, \\ \|w_k^l\|_{L^\infty(B_\gamma^e(\xi)\cap L_l)} &\leq c\varepsilon^{\frac{1}{6}}\gamma + \delta_k, \quad \text{where } \delta_k \to 0, \\ \|\operatorname{D} w_k^l\|_{L^\infty(B_\gamma^e(\xi)\cap L_l)} &\leq c\varepsilon^{\frac{1}{6}} + \delta_k, \quad \text{where } \delta_k \to 0, \\ \int_{B_\gamma^e(\xi)\cap L_l} |\operatorname{D}^2 w_k^l|^2 &\leq c\gamma \int_{\operatorname{graph} u_k^l|_{\partial B_\gamma^e(\xi)\cap L_l}} |A_k^e|^2\, d\mathcal{H}_e^1, \end{aligned}$$

where $d\mathcal{H}_e^1$ is the 1-dimensional Euclidean Hausdorff measure.

Observe that, with an analogous argument as above using the estimates on w_k^l, we get

$$\operatorname{graph} w_k^l \subset \overline{B_{\frac{\rho}{4}}^e(\xi)} \quad \text{for } \varepsilon \leq \varepsilon_0 \text{ and } \delta_k \text{ sufficiently small.} \tag{18}$$

By exchanging for each l the disc $D_k^l \cap C_\gamma^l(\xi)$ with the disc $\operatorname{graph} w_k^l$, we get a new immersed surface $\tilde{\Sigma}_k$, which can be parametrized on $\mathbb{S}^2$ by a $C^{1,1}$-immersion $\tilde{f}_k : \mathbb{S}^2 \hookrightarrow M$. To simplify notation at this point and later in the paper we just write

$$\tilde{\Sigma}_k = \left(f_k(\mathbb{S}^2) \setminus \left(\bigcup_l D_k^l \cap C_\gamma^l(\xi)\right)\right) \cup \bigcup_l \operatorname{graph} w_k^l. \tag{19}$$

More precisely we have to do the following: Choose a radius $\gamma' > \gamma$ such that the disc D_k^l has a smooth graph representation by u_k^l on the annulus $A := B_{\gamma'}(\xi) \setminus \overline{B_\gamma(\xi)} \cap L_l$. Consider the disjoint union of the disc $B_{\gamma'}(\xi) \cap L_l$ and the topological disc $\Delta := \mathbb{S}^2 \setminus f_k^{-1}(D_k^l \cap \overline{C_\gamma^l(\xi)})$. Consider the diffeomorphism $\phi : A \to \mathbb{S}^2 \setminus f_k^{-1}(D_k^l \cap C_{\gamma'}^l(\xi) \setminus \overline{C_\gamma^l(\xi)})$ given by $\phi(x) = f_k^{-1}(x + u_k^l(x))$. We define the smooth 2-manifold Σ by identifying $x \in A$ and $p \in \mathbb{S}^2 \setminus f_k^{-1}(D_k^l \cap C_{\gamma'}^l(\xi) \setminus \overline{C_\gamma^l(\xi)})$ if $\phi(x) = p$. We thus get a $C^{1,1}$-immersion $\tilde{f}_k : \Sigma \to M$ by putting

$$\tilde{f}_k = \begin{cases} f_k & \text{on } \Delta \\ x + u_k^l(x) & \text{for } x \in A \\ x + w_k^l(x) & \text{for } x \in \overline{B_\gamma(\xi)} \cap L_l. \end{cases}$$

It is easy to check that Σ is orientable and has cohomology $H^1(\Sigma) = 0$, and hence Σ is diffeomorphic to $\mathbb{S}^2$. This constructs the desired $C^{1,1}$-immersion of $\mathbb{S}^2$.

From the definition of γ we have that

$$\int_{\operatorname{graph} w_k^l} |A_e|^2\, d\mathcal{H}_e^2 \leq c \int_{B_\gamma^e(\xi)\cap L_l} |\operatorname{D}^2 w_k^l|^2 \leq c\gamma \int_{\operatorname{graph} u_k^l|_{\partial B_\gamma^e(\xi)\cap L_l}} |A_k^e|^2 d\mathcal{H}_e^1 = c\gamma \int_{\partial C_\gamma^l(\xi)\cap D_k^l} |A_k^e|^2 d\mathcal{H}_e^1.$$

Until now $\gamma \in T_l \subset \left(\frac{\rho}{16}, \frac{3\rho}{32}\right)$ was arbitrary. Since $\mathcal{L}^1(T_l) \geq \frac{\rho}{64}$, it easily follows from a simple Fubini-type argument as done in [SiL] that we can choose $\gamma \in T_l$ such that for every l, k

$$\int_{\operatorname{graph} w_k^l} |A_e|^2\, d\mathcal{H}_e^2 \leq c \int_{\left(D_k^l \cap C_{\frac{3\rho}{32}}^l(\xi) \setminus C_{\frac{\rho}{16}}^l(\xi)\right) \setminus \bigcup_j P_j^{k,l}} |A_k^e|^2\, d\mathcal{H}_e^2.$$

Now notice that for $\varepsilon \le \varepsilon_0$ (this follows from the estimates for u_k^l and $\mathrm{D}\, u_k^l$)

$$B^e_{\frac{\rho}{16}}(\xi) \subset C^l_{\frac{\rho}{16}}(\xi) \quad \text{and} \quad \left(D_k^l \cap C^l_{\frac{3\rho}{32}}(\xi)\right) \setminus \bigcup_j P_j^{k,l} \subset \left(D_k^l \cap B^e_{\frac{\rho}{8}}(\xi)\right) \setminus \bigcup_j P_j^{k,l}.$$

We get that

$$\int_{\operatorname{graph} w_k^l} |A_e|^2 \, d\mathcal{H}_e^2 \le c \int_{D_k^l \cap B^e_{\frac{\rho}{8}}(\xi) \setminus B^e_{\frac{\rho}{16}}(\xi)} |A_k^e|^2 \, d\mathcal{H}_e^2.$$

By summing over l and using the uniform bound on M_k it follows that

$$\sum_{l=1}^{M_k} \int_{\operatorname{graph} w_k^l} |A_e|^2 \, d\mathcal{H}_e^2 \le c \sum_{l=1}^{M_k} \int_{D_k^l \cap B^e_{\frac{\rho}{8}}(\xi) \setminus B^e_{\frac{\rho}{16}}(\xi)} |A_k^e|^2 \, d\mathcal{H}_e^2 = c \int_{B^e_{\frac{\rho}{8}}(\xi) \setminus B^e_{\frac{\rho}{16}}(\xi)} |A_k^e|^2 \, d\mu_k^e. \tag{20}$$

Since f_k is a minimizing sequence for the functional E we get

$$E(\tilde{f}_k) \ge E(f_k) - \varepsilon_k, \quad \text{where } \varepsilon_k \to 0,$$

which implies

$$\sum_{l=1}^{M_k} \int_{\operatorname{graph} w_k^l} |A|^2 \, d\mathcal{H}^2 \ge \int_{B^e_{\frac{\rho}{16}}(\xi)} |A_k|^2 \, d\mu_k - \varepsilon_k. \tag{21}$$

Using the estimates of Lemma 2.7 we finally get that

$$\int_{B^e_{\frac{\rho}{16}}(\xi)} |A_k^e|^2 \, d\mu_k^e \le c \int_{B^e_{\frac{\rho}{8}}(\xi) \setminus B^e_{\frac{\rho}{16}}(\xi)} |A_k^e|^2 \, d\mu_k^e + \varepsilon_k + c\rho^2. \tag{22}$$

By adding c times the left hand side of this inequality to both sides ("hole-filling") we get that for all $\rho \le \frac{\rho_0}{4}$ and infinitely many k

$$\int_{B^e_{\frac{\rho}{16}}(\xi)} |A_k^e|^2 \, d\mu_k^e \le \theta \int_{B^e_{\frac{\rho}{8}}(\xi)} |A_k^e|^2 \, d\mu_k^e + \varepsilon_k + c\rho^2,$$

where $\theta = \frac{c}{c+1} \in (0,1)$ only depends on the manifold M. Now if we let $g(\rho) = \liminf_{k\to\infty} \int_{B^e_\rho(\xi)} |A_k^e|^2 \, d\mu_k^e$, we get that

$$g(\rho) \le \theta g(2\rho) + c\rho^2 \quad \text{for all } \rho \le \frac{\rho_0}{64}.$$

In view of Lemma 5.3 in the Appendix the present Lemma is proved. ■

Now we are able to show that, in a neighborhood of the good points, the limit measure μ is the Radon measure associated to $C^{1,\alpha} \cap W^{2,2}$-graphs. First we recall the setting shortly: We had that $u_k^l : \Omega_k^l \to L_l^\perp$, where the set Ω_k^l was given by

$$\Omega_k^l = (B^e_\lambda(\xi) \cap L_l) \setminus \bigcup_m d^l_{k,m},$$

where $\lambda > \frac{\rho}{4}$, and where the sets $d^l_{k,m} \subset L_l$ are pairwise disjoint closed discs which do not intersect $\partial B^e_\lambda(\xi)$.

Define the quantity $\alpha_k(\rho)$ by

$$\alpha_k(\rho) = \int_{B^e_{4\rho}(\xi)} |A_k^e|^2 \, d\mu_k^e,$$

and notice that by Lemma 3.6 and Lemma 3.3 we have

$$\liminf_{k\to\infty} \alpha_k(\rho) \le \min\left\{c\rho^\alpha, c\varepsilon^2\right\} \quad \text{for all } \rho \le \frac{\rho_0}{128}. \tag{23}$$

Moreover it follows from Lemma 3.4 that

$$\sum_m \operatorname{diam} d^l_{k,m} \le c\alpha_k(\rho)^{\frac{1}{4}}\rho. \tag{24}$$

Therefore for $\varepsilon \le \varepsilon_0$ we may apply the generalized Poincaré inequality Lemma 5.4 to the functions $f^l_j = \mathrm{D}_j\, u^l_k$ and $\delta = c\alpha_k(\rho)^{\frac{1}{4}}\rho$ in order to get a constant vector η^l_k, with $|\eta^l_k| \le c\varepsilon^{\frac{1}{6}} + \delta_k \le c$ and $\delta_k \to 0$, such that

$$\int_{\Omega^l_k} \left|\mathrm{D}\, u^l_k - \eta^l_k\right|^2 \le c\rho^2 \int_{\Omega^l_k} \left|\mathrm{D}^2\, u^l_k\right|^2 + c\alpha_k(\rho)^{\frac{1}{4}}\rho^2 \sup_{\Omega^l_k} \left|\mathrm{D}\, u^l_k\right|^2.$$

Now we have

$$\int_{\Omega^l_k} \left|\mathrm{D}^2\, u^l_k\right|^2 \le c \int_{\operatorname{graph} u^l_k} |A^e_k|^2\, d\mathcal{H}^2_e \le c \int_{B^e_{2\rho}(\xi)} |A^e_k|^2\, d\mu^e_k \le c\alpha_k(\rho).$$

Since $|\mathrm{D}\, u^l_k| \le c$ and $\alpha_k(\rho) \le 1$ for $\varepsilon \le \varepsilon_0$, it follows that

$$\int_{\Omega^l_k} \left|\mathrm{D}\, u^l_k - \eta^l_k\right|^2 \le c\alpha_k(\rho)^{\frac{1}{4}}\rho^2. \tag{25}$$

Now let $\overline{u}^l_k \in C^{1,1}(B^e_\lambda(\xi) \cap L_l, L^\perp_l)$ be an extension of u^l_k to all of $B^e_\lambda(\xi) \cap L_l$ as in Lemma 5.1, namely

$$\begin{aligned}
\overline{u}^l_k &= u^l_k \quad \text{in } B^e_\lambda(\xi) \cap L_l \setminus \bigcup_m d^l_{k,m}, \\
\overline{u}^l_k = u^l_k \quad &\text{and} \quad \frac{\partial \overline{u}^l_k}{\partial \nu} = \frac{\partial u^l_k}{\partial \nu} \quad \text{on } \bigcup_m \partial d^l_{k,m}, \\
||\overline{u}^l_k||_{L^\infty(d^l_{k,m})} &\le c\varepsilon^{\frac{1}{6}}\rho + \delta_k, \quad \text{where } \delta_k \to 0, \\
||\mathrm{D}\,\overline{u}^l_k||_{L^\infty(d^l_{k,m})} &\le c\varepsilon^{\frac{1}{6}} + \delta_k, \quad \text{where } \delta_k \to 0.
\end{aligned}$$

It follows that $||\overline{u}^l_k||_{L^\infty(B^e_\lambda(\xi)\cap L_l)} + ||\mathrm{D}\,\overline{u}^l_k||_{L^\infty(B^e_\lambda(\xi)\cap L_l)} \le c$, where c is independent of k. From the gradient estimates for the function $\overline{u}^l_k$, since $|\eta^l_k| \le c$ and because of (24) we get that

$$\int_{B^e_\lambda(\xi)\cap L_l} \left|\mathrm{D}\,\overline{u}^l_k - \eta^l_k\right|^2 \le c\alpha_k(\rho)^{\frac{1}{4}}\rho^2. \tag{26}$$

Thus, in view of (23), we conclude that

$$\liminf_{k\to\infty} \int_{B^e_\lambda(\xi)\cap L_l} \left|\mathrm{D}\,\overline{u}^l_k - \eta^l_k\right|^2 \le \min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{2}}\rho^2\right\} \quad \text{for all } \rho \le \frac{\rho_0}{128}. \tag{27}$$

Moreover, it trivially follows that $\left\|\overline{u}^l_k\right\|_{W^{1,2}(B^e_\lambda(\xi)\cap L_l)} \le c\rho^2 \le c$. Therefore it follows that the sequence $\overline{u}^l_k$ is equicontinuous and uniformly bounded in $C^1(B^e_\lambda(\xi) \cap L_l, L^\perp_l)$ and $W^{1,2}(B^e_\lambda(\xi) \cap L_l, L^\perp_l)$, and we get the existence of a function $u^l_\xi \in C^{0,1}(B^e_\lambda(\xi) \cap L_l, L^\perp_l)$ such that (after passing to a subsequence)

$$\begin{aligned}
\overline{u}^l_k &\to u^l_\xi \quad \text{in } C^0(B^e_\lambda(\xi) \cap L_l, L^\perp_l), \\
\overline{u}^l_k &\rightharpoonup u^l_\xi \quad \text{weakly in } W^{1,2}(B^e_\lambda(\xi) \cap L_l, L^\perp_l),
\end{aligned}$$

and such that the function u^l_ξ satisfies the estimates

$$\frac{1}{\rho}||u^l_\xi||_{L^\infty(B^e_\lambda(\xi)\cap L_l)} + ||\mathrm{D}\, u^l_\xi||_{L^\infty(B^e_\lambda(\xi)\cap L_l)} \le c\varepsilon^{\frac{1}{6}}.$$

Be aware that, a priori, the limit function might depend on the point ξ. Indeed, the sequence u^l_k depends on ξ since it comes from the Graphical Decomposition Lemma which is a local statement.

Observe that, up to subsequences, $\eta^l_k \to \eta^l$ with $|\eta^l| \le c\varepsilon^{\frac{1}{6}}$. Since $\mathrm{D}\,\overline{u}^l_k \rightharpoonup \mathrm{D}\,u^l_\xi$ weakly in $L^2(B^e_\lambda(\xi)\cap L_l)$ it follows that $\mathrm{D}\,\overline{u}^l_k - \eta^l_k \rightharpoonup \mathrm{D}\,u^l_\xi - \eta^l$ weakly in $L^2(B^e_\lambda(\xi)\cap L_l)$, and by lower-semicontinuity, estimate (27) implies that

$$\int_{B^e_\lambda(\xi)\cap L_l} \left|\mathrm{D}\,u^l_\xi - \eta^l\right|^2 \le \min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{2}}\rho^2\right\} \quad \text{for all } \rho \le \frac{\rho_0}{128}. \tag{28}$$

Now we can prove the graphical representation of the limit measure μ.

Lemma 3.7. *For $\varepsilon \le \varepsilon_0$ there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon)$ such that for all $\xi \in \operatorname{spt}\mu \cap B^e_{\frac{\rho_0}{2}}(\xi_0)$ and all $\rho \le \rho_0$ we have*

$$\mu\llcorner B^e_\rho(\xi) = \sum_{l=1}^M \mathcal{H}^2\llcorner\left(\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi)\right),$$

where $\mathcal{H}^2$ denotes the 2-dimensional Hausdorff measure of the Riemannian manifold M, and where each function $u^l_\xi \in C^{0,1}\left(B^e_\rho(\xi)\cap L_l, L_l^\perp\right)$ is as above, in particular

$$\frac{1}{\rho}||u^l_\xi||_{L^\infty(B^e_\rho(\xi)\cap L_l)} + \left\|\mathrm{D}\,u^l_\xi\right\|_{L^\infty(B^e_\rho(\xi)\cap L_l)} \le c\varepsilon^{\frac{1}{6}}.$$

Proof. First we claim that for all $\rho \le \frac{\rho_0}{128}$ we have

$$\mu_k\llcorner B^e_\rho(\xi) = \sum_{l=1}^M \mathcal{H}^2\llcorner\left(\operatorname{graph} \overline{u}^l_k \cap B^e_\rho(\xi)\right) + \theta_k, \tag{29}$$

where θ_k is a signed measure with $\liminf_{k\to\infty}$ of the total mass is smaller than $\min\left\{c\rho^{2+\alpha}, c\varepsilon\rho^2\right\}$, i.e. $\theta_k = \theta^1_k - \theta^2_k$ with $\liminf_{k\to\infty}\left(\theta^1_k(M) + \theta^2_k(M)\right) \le \min\left\{c\rho^{2+\alpha}, c\varepsilon\rho^2\right\}$.

To prove the claim recall that the diameter estimates in Lemma 3.4, the quadratic area decay and the monotonicity formula Lemma 2.6 yield

$$\sum_{m,l}\mathcal{L}^2(d^l_{k,m}) + \sum_{j,l}\mathcal{H}^2(P^{k,l}_j) \le c\alpha_k(\rho)^{\frac{1}{2}}\rho^2.$$

Thus Lemma 3.6 yields for $\rho \le \frac{\rho_0}{128}$

$$\liminf_{k\to\infty}\sum_{m,l}\mathcal{L}^2(d^l_{k,m}) + \liminf_{k\to\infty}\sum_{j,l}\mathcal{H}^2(P^{k,l}_j) \le \min\left\{c\rho^{2+\alpha}, c\varepsilon\rho^2\right\}.$$

The Graphical Decomposition Lemma 3.4 yields $\mu_k\llcorner B^e_\rho(\xi) = \sum_{l=1}^M \mathcal{H}^2\llcorner\left(\operatorname{graph}\overline{u}^l_k \cap B^e_\rho(\xi)\right) + \theta_k$, where

$$\theta_k = \sum_{l=1}^M \mathcal{H}^2\llcorner\left(\left(D^l_k \setminus \operatorname{graph}\overline{u}^l_k\right)\cap B^e_\rho(\xi)\right) - \sum_{l=1}^M \mathcal{H}^2\llcorner\left(\left(\operatorname{graph}\overline{u}^l_k \setminus D^l_k\right)\cap B^e_\rho(\xi)\right) = \theta^1_k - \theta^2_k.$$

We have that $\theta^1_k(M) \le \sum_{j,l}\mathcal{H}^2(P^{k,l}_j)$ and $\theta^2_k(M) \le c\sum_{m,l}\mathcal{L}^2(d^l_{k,m})$, and (29) follows.

Now by taking limits in the measure theoretic sense we claim that

$$\mu\llcorner B^e_\rho(\xi) = \sum_{l=1}^M \mathcal{H}^2\llcorner\left(\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi)\right) + \theta_\xi, \tag{30}$$

where θ_ξ is a signed measure with total mass smaller than $\min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{4}}\rho^2\right\}$. This equation holds for all $\rho \leq \frac{\rho_0}{128}$ such that

$$\mu\left(\partial B^e_\rho(\xi)\right) = \mathcal{H}^2 \llcorner \operatorname{graph} u^l_\xi \left(\partial B^e_\rho(\xi)\right) = 0 \quad \text{for all } l,$$

which holds for almost every ρ.

To prove (30) let $U \subset M$ be an open subset.

1.) Let $\rho \leq \frac{\rho_0}{128}$ such that $\mu\left(\partial B^e_\rho(\xi)\right) = 0$. Moreover assume that $\mu \llcorner B^e_\rho(\xi)\,(\partial U) = 0$. It follows that $\mu\left(\partial\left(U \cap B^e_\rho(\xi)\right)\right) = 0$ and therefore $\mu_k\left(U \cap B^e_\rho(\xi)\right) \to \mu\left(U \cap B^e_\rho(\xi)\right)$.

2.) Let $\rho \leq \frac{\rho_0}{128}$ such that $\mathcal{H}^2 \llcorner \operatorname{graph} u^l_\xi \left(\partial B^e_\rho(\xi)\right) = 0$. Assume that $\mathcal{H}^2 \llcorner (\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi))\,(\partial U) = 0$. Now in general it follows for the 2-dimensional Hausdorff measure of some $C^{0,1}$-graph u that

$$\mathcal{H}^2(\operatorname{graph} u) = \int \sqrt{\det g} = \int \sqrt{A(x,u(x)) + B_i(x,u(x))\partial_i u(x) + C_{ij}(x,u(x))\partial_i u(x)\partial_j u(x)},$$

where the coefficients A, B_i, C_{ij} just depend on the metric h and are uniformly bounded in terms of the manifold M. Especially for the coefficient A we have

$$A(x) = h_{11}(x)h_{22}(x) - h_{12}(x)^2, \tag{31}$$

where h_{ij} are the coefficients of the metric h of M. Therefore we get that the coefficient A is bounded from below by a positive constant, namely there exists a constant $c_0 > 0$ such that

$$\sup_{x \in M} A(x) \geq c_0 > 0. \tag{32}$$

Using the L^∞-bounds for the functions $\overline{u}^l_k$ and the coefficients A, B_i, C_{ij}, we get that

$$\begin{aligned}
&\left|\mathcal{H}^2 \llcorner \left(\operatorname{graph} \overline{u}^l_k \cap B^e_\rho(\xi)\right)(U) - \mathcal{H}^2 \llcorner \left(\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi)\right)(U)\right| \\
&\quad \leq c \int_{L_l} \left|\chi_{U \cap B^e_\rho(\xi)}(x, \overline{u}^l_k(x)) - \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x))\right| + \int_{L_l} \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x)) \left|\sqrt{\det \overline{g}^l_k} - \sqrt{\det g^l}\right|
\end{aligned}$$

Since $\overline{u}^l_k \to u^l_\xi$ uniformly and since $\mathcal{H}^2 \llcorner \operatorname{graph} u^l_\xi \left(\partial B^e_\rho(\xi)\right) = 0$, it follows that

$$\chi_{U \cap B^e_\rho(\xi)}(x, \overline{u}^l_k(x)) \to \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x)) \quad \text{for a.e. } x \in L_l.$$

The Dominated Convergence Theorem yields

$$\int_{L_l} \left|\chi_{U \cap B^e_\rho(\xi)}(x, \overline{u}^l_k(x)) - \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x))\right| \to 0.$$

Now because of (32) and the bounds for the functions $\overline{u}^l_k$ and u^l_ξ it follows that for $\varepsilon \leq \varepsilon_0$

$$\begin{aligned}
\int_{L_l} \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x)) \left|\sqrt{\det \overline{g}^l_k} - \sqrt{\det g^l}\right| &\leq \int_{L_l} \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x)) \left|\det \overline{g}^l_k - \det g^l\right| \\
&\leq \int_{L_l} \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x)) \left|A(x, \overline{u}^l_k(x)) - A(x, u^l_\xi(x))\right| \\
&\quad + \int_{L_l} \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x)) \left|B_i(x, \overline{u}^l_k(x))\partial_i \overline{u}^l_k(x) - B_i(x, u^l_\xi(x))\partial_i u^l_\xi(x)\right| \\
&\quad + \int_{L_l} \chi_{U \cap B^e_\rho(\xi)}(x, u^l_\xi(x)) \left|C_{ij}(x, \overline{u}^l_k(x))\partial_i \overline{u}^l_k(x)\partial_j \overline{u}^l_k(x) - C_{ij}(x, u^l_\xi(x))\partial_i u^l_\xi(x)\partial_j u^l_\xi(x)\right| \\
&=: (1) + (2) + (3).
\end{aligned}$$

Now (1) $\to 0$ for $k \to \infty$ because of the uniform convergence $\overline{u}_k^l \to u_\xi^l$. The second term can be estimated in view of the boundedness of the coefficients B_i and the functions $\overline{u}_k^l$ by

$$
\begin{aligned}
(2) \quad &\le c\int_{B_\rho^e(\xi)\cap L_l} \left|B_i(x,\overline{u}_k^l(x)) - B_i(x,u_\xi^l(x))\right| + c\int_{L_l} \chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x)) \left|\mathrm{D}\,\overline{u}_k^l(x) - \mathrm{D}\,u_\xi^l(x)\right| \\
&\le c\int_{B_\rho^e(\xi)\cap L_l} \left|B_i(x,\overline{u}_k^l(x)) - B_i(x,u_\xi^l(x))\right| + c\int_{L_l} \chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x)) \left|\mathrm{D}\,\overline{u}_k^l(x) - \eta_k^l\right| \\
&\quad + c\int_{L_l} \chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x)) \left|\eta_k^l - \eta^l\right| + c\int_{L_l} \chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x)) \left|\eta^l - \mathrm{D}\,u_\xi^l(x)\right|
\end{aligned}
$$

The first term goes to 0, again by the uniform convergence $\overline{u}_k^l \to u_\xi^l$. For the second term we have that $\chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x)) = 0$ if $x \notin B^e_{\left(1-c\varepsilon^{\frac{1}{6}}\right)\rho}(\xi)\cap L_l$. This follows from the L^∞-bound for the function u_ξ^l. Therefore we get that

$$\left(\int_L \chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x))\right)^{\frac{1}{2}} \le \mathcal{L}^2\left(B^e_{\left(1-c\varepsilon^{\frac{1}{6}}\right)\rho}(\xi)\cap L_l\right)^{\frac{1}{2}} \le c\rho.$$

In view of (27) we get $\liminf_{k\to\infty}\int_{L_l}\chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x))|\,\mathrm{D}\,\overline{u}_k^l(x)-\eta_k^l| \le \min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{4}}\rho^2\right\}$. With (28) we get in the same way that $\int_{L_l}\chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x))|\eta^l - \mathrm{D}\,u_\xi^l(x)| \le \min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{4}}\rho^2\right\}$. Now since $\eta_k^l \to \eta^l$ strongly, we finally get that

$$\liminf_{k\to\infty}(2) \le \min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{4}}\rho^2\right\}. \tag{33}$$

It remains to estimate the last term (3). It follows as above that

$$(3) \quad \le c\int_{B_\rho^e(\xi)\cap L_l} \left|C_{ij}(x,\overline{u}_k^l(x)) - C_{ij}(x,u_\xi^l(x))\right| + c\int_{L_l} \chi_{U\cap B_\rho^e(\xi)}(x,u_\xi^l(x)) \left|\mathrm{D}\,\overline{u}_k^l(x) - \mathrm{D}\,u_\xi^l(x)\right|.$$

The first term goes to 0 as usual, and the second term is the same as above, which yields

$$\liminf_{k\to\infty}(3) \le \min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{4}}\rho^2\right\}. \tag{34}$$

After all we have finally shown that

$$\mathcal{H}^2\llcorner\left(\operatorname{graph}\overline{u}_k^l \cap B_\rho^e(\xi)\right)(U) = \mathcal{H}^2\llcorner\left(\operatorname{graph}u_\xi^l \cap B_\rho^e(\xi)\right)(U) + \tilde{\theta}_k(U), \tag{35}$$

where $\tilde{\theta}_k$ is a signed measure such that the $\liminf_{k\to\infty}$ of the total mass is smaller that $\min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{4}}\rho^2\right\}$. After passing to a subsequence, $\tilde{\theta}_k$ converges weakly to some signed measure $\tilde{\theta}_\xi$ with total mass smaller than $\min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{4}}\rho^2\right\}$. Assume that $\tilde{\theta}_\xi(\partial U) = 0$. It follows that $\tilde{\theta}_k(U) \to \tilde{\theta}_\xi(U)$, and therefore we get that

$$\lim_{k\to\infty}\mathcal{H}^2\llcorner\left(\operatorname{graph}\overline{u}_k^l \cap B_\rho^e(\xi)\right)(U) = \mathcal{H}^2\llcorner\left(\operatorname{graph}u_\xi^l \cap B_\rho^e(\xi)\right)(U) + \tilde{\theta}_\xi(U). \tag{36}$$

3.) Since the θ_k's were signed measures such that the lim inf of the total mass $\le \min\left\{c\rho^{2+\alpha}, c\varepsilon\rho^2\right\}$, they converge in the weak sense (after passing to a subsequence) to a signed measure $\overline{\theta}_\xi$ with total mass smaller than $\min\left\{c\rho^{2+\alpha}, c\varepsilon\rho^2\right\}$. Assuming $\overline{\theta}_\xi(\partial U) = 0$, it follows that $\theta_k(U) \to \overline{\theta}_\xi(U)$.

Now by taking limits in (29) we get from 1.), 2.) and 3.) that

$$\mu\llcorner B_\rho^e(\xi)(U) = \sum_{l=1}^{M}\mathcal{H}^2\llcorner\left(\operatorname{graph}u_\xi^l \cap B_\rho^e(\xi)\right)(U) + \theta_\xi(U),$$

where $\theta_\xi = \overline{\theta}_\xi + \tilde{\theta}_\xi$ is a signed measure with total mass smaller than $\min\left\{c\rho^{2+\alpha}, c\varepsilon^{\frac{1}{4}}\rho^2\right\}$. Notice that this equation holds for all open $U \subset M$ with $\mu \llcorner B^e_\rho(\xi)(\partial U) = \mathcal{H}^2 \llcorner (\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi))(\partial U) = \overline{\theta}_\xi(\partial U) = \tilde{\theta}_\xi(\partial U) = 0$. By choosing an appropriate exhaustion this equation holds for arbitrary open sets $U \subset M$ and (30) is shown.

Next we claim that $\operatorname{spt}\mu$ is locally given by the union of the graphs of the functions u^l_ξ, i.e. for $\rho \le \frac{\rho_0}{256}$ it follows that

$$\operatorname{spt}\mu \cap B^e_\rho(\xi) = \bigcup_{l=1}^M \operatorname{graph} u^l_\xi \cap B^e_\rho(\xi). \tag{37}$$

To prove this let $\rho \le \frac{\rho_0}{128}$ such that (30) holds.

1.) Let $x \in \operatorname{spt}\mu \cap B^e_{\frac{\rho}{2}}(\xi)$. Proposition 3.5 yields $\mu \llcorner B^e_\rho(\xi)(B^e_{\frac{\rho}{2}}(x)) = \mu(B^e_{\frac{\rho}{2}}(x)) \ge c\rho^2$. We get

$$c\rho^2 \le \sum_{l=1}^M \mathcal{H}^2\left(\operatorname{graph} u^l_\xi \cap B^e_{\frac{\rho}{2}}(x)\right) + c\varepsilon^{\frac{1}{4}}\rho^2.$$

By choosing $\varepsilon \le \varepsilon_0$ we conclude that $\sum_{l=1}^M \mathcal{H}^2\left(\operatorname{graph} u^l_\xi \cap B^e_{\frac{\rho}{2}}(x)\right) > 0$ and therefore $x \in \bigcup_{l=1}^M \operatorname{graph} u^l_\xi$.

2.) Let $z \in \bigcup_{l=1}^M \operatorname{graph} u^l_\xi \cap B^e_{\frac{\rho}{2}}(\xi)$. Write $z = x + u^l_\xi(x)$ for some $l \in \{1, \dots, M\}$ and some $x \in L_l$. If $y \in B^e_{\frac{\rho}{4}}(x) \cap L_l$ we claim that $y + u^l_\xi(y) \in B^e_{\frac{\rho}{2}}(z)$, indeed for $\varepsilon \le \varepsilon_0$ we get

$$|z - y - u^l_\xi(y)| \le |x-y| + |u^l_\xi(x) - u^l_\xi(y)| \le \left(1 + c\varepsilon^{\frac{1}{6}}\right)|x-y| \le \left(1 + c\varepsilon^{\frac{1}{6}}\right)\frac{\rho}{4} \le \frac{\rho}{2}.$$

Therefore

$$\mathcal{H}^2 \llcorner \operatorname{graph} u^l_\xi(B^e_{\frac{\rho}{2}}(z)) \ge c\,\mathcal{L}^2(B^e_{\frac{\rho}{4}}(x) \cap L_l) = c\rho^2.$$

As above we obtain $\mu(B^e_{\frac{\rho}{2}}(z)) \ge c\rho^2 - c\varepsilon^{\frac{1}{4}}\rho^2 > 0$ for $\varepsilon \le \varepsilon_0$, and conclude that $z \in \operatorname{spt}\mu$.

Now (37) implies that the functions u^l_ξ do not depend on the point ξ in the following sense: Let $\eta \in \Sigma \cap B^e_{\frac{\rho_0}{2}}(\xi_0)$. Then we have for all $\rho \le \frac{\rho_0}{256}$ that

$$\bigcup_{l=1}^M \operatorname{graph} u^l_\xi \cap \left(B^e_\rho(\xi) \cap B^e_\rho(\eta)\right) = \bigcup_{l=1}^N \operatorname{graph} u^l_\eta \cap \left(B^e_\rho(\xi) \cap B^e_\rho(\eta)\right). \tag{38}$$

In the next step choose $\rho \le \frac{\rho_0}{256}$ such that $\mu\left(\partial B^e_\rho(\xi)\right) = \mathcal{H}^2_g \llcorner \operatorname{graph} u^l_\xi\left(\partial B^e_\rho(\xi)\right) = 0$ for all l, and that therefore, from (30),

$$\mu \llcorner B^e_\rho(\xi) = \sum_{l=1}^M \mathcal{H}^2 \llcorner \left(\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi)\right) + \theta_\xi. \tag{39}$$

Let $z \in \operatorname{spt}\mu \cap B^e_\rho(\xi) = \bigcup_{l=1}^M \operatorname{graph} u^l_\xi \cap B^e_\rho(\xi)$ and let $\sigma > 0$ such that $B^e_\sigma(z) \subset B^e_\rho(\xi)$ and such that (due to (30) for the point z) $\mu \llcorner B^e_\sigma(z) = \sum_{l=1}^N \mathcal{H}^2 \llcorner \left(\operatorname{graph} u^l_z \cap B^e_\sigma(z)\right) + \theta_z$, where the total mass of θ_z is smaller than $c\sigma^{2+\alpha}$. From (38) it follows that $\theta_\xi\left(B^e_\sigma(z)\right) = \theta_z\left(B^e_\sigma(z)\right)$, hence we get a nice decay for the signed measure θ_ξ, namely

$$\lim_{\sigma \to 0} \frac{\theta_\xi\left(B^e_\sigma(z)\right)}{\sigma^2} = 0 \quad \text{for all } z \in \operatorname{spt}\mu \cap B^e_\rho(\xi) = \bigcup_{l=1}^M \operatorname{graph} u^l_\xi \cap B^e_\rho(\xi). \tag{40}$$

Now it follows as before that for all $z \in \operatorname{spt}\mu \cap B^e_\rho(\xi) = \bigcup_{l=1}^M \operatorname{graph} u^l_\xi \cap B^e_\rho(\xi)$

(41) $$\liminf_{\sigma\to 0} \frac{\sum_{l=1}^M \mathcal{H}^2 \llcorner (\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi))(B^e_\sigma(z))}{\pi\sigma^2} \geq C > 0.$$

Now for $z \in \operatorname{spt}\mu \cap B^e_\rho(\xi) = \bigcup_{l=1}^M \operatorname{graph} u^l_\xi \cap B^e_\rho(\xi)$ and $\sigma > 0$ such that $B^e_\sigma(z) \subset B^e_\rho(\xi)$ it follows from (39), (40) and (41) that

$$\frac{\mu \llcorner B^e_\rho(\xi)\,(B^e_\sigma(z))}{\sum_{l=1}^M \mathcal{H}^2 \llcorner (\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi))\,(B^e_\sigma(z))} = 1 + \frac{\theta_\xi\,(B^e_\sigma(z))}{\sum_{l=1}^M \mathcal{H}^2 \llcorner (\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi))\,(B^e_\sigma(z))}.$$

Since the right hand side converges to 1, this shows that $\mathrm{D}_{\left(\sum_{l=1}^M \mathcal{H}^2 \llcorner (\operatorname{graph} u^l_\xi \cap B^e_\rho(\xi))\right)} \left(\mu \llcorner B^e_\rho(\xi)\right)(z) = 1$ for all $z \in \operatorname{spt}\mu \cap B^e_\rho(\xi) = \bigcup_{l=1}^M \operatorname{graph} u^l_\xi \cap B^e_\rho(\xi)$. The Lemma now follows from the Theorem of Radon-Nikodym. ■

Up to now we have shown that, away from the bad points, the limit measure μ is locally given by $C^{0,1}$-graphs with small gradient bounded by $c\varepsilon^{\frac{1}{6}}$. In the next step we will show, using the power decay in Lemma 3.6, that these graphs are actually $C^{1,\alpha} \cap W^{2,2}$-graphs, and that the L^2-norm of their Hessians satisfy a similar power decay.

Proposition 3.8. *For $\varepsilon \leq \varepsilon_0$ there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ such that*

(i) $u^l_{\xi_0} \in C^{1,\alpha}(L_l \cap B^e_{\rho_0}(\xi_0)) \cap W^{2,2}(L_l \cap B^e_{\rho_0}(\xi_0))$,

(ii) $\int_{B^e_\sigma(x)\cap L_l} |\mathrm{D}^2 u^l_{\xi_0}|^2 \leq C\sigma^\alpha$ *for all $x \in B^e_{\rho_0}(\xi_0) \cap L_l$ and all $\sigma > 0$ sufficiently small.*

Proof. By applying Lemma 3.7 to $\xi = \xi_0$, we get that for $\varepsilon \leq \varepsilon_0$ there exist $\rho_0 = \rho_0(\xi_0, \varepsilon_0) > 0$, 2-dimensional planes $L_l \subset T_{\xi_0}M, l = 1, \dots, M_{\xi_0}$, and functions $u^l_{\xi_0} \in C^{0,1}(L_l \cap B^e_{\rho_0}(\xi_0))$ such that for all $\rho \leq \rho_0$

$$\mu \llcorner B^e_\rho(\xi_0) = \sum_{l=1}^{M_{\xi_0}} \mathcal{H}^2 \llcorner \left(\operatorname{graph} u^l_{\xi_0} \cap B^e_\rho(\xi_0)\right).$$

Because of the uniform bounds on the area and the Willmore energy of the immersions f_k in the induced metric g_k, it follows from Lemma 2.7 that, for ρ_0 maybe smaller, we have $\mu^e_k(B^e_{\rho_0}(\xi_0)) \leq C$ and $\int_{B^e_{\rho_0}(\xi_0)} |H^e_k|^2 d\mu^e_k \leq C$. It follows that $\mu^e_k \llcorner B^e_{\rho_0}(\xi_0)$ defines an integral, rectifiable 2-varifold with uniformly bounded first variation. By a compactness result for varifolds (see [SiGMT]), there exists an integral, rectifiable 2-varifold μ^e with weak mean curvature vector $H^e \in L^2(\mu)$, such that (after passing to a subsequence) $\mu^e_k \llcorner B^e_{\rho_0}(\xi_0) \to \mu^e$ weakly in the sense of Radon measures and such that

(42) $$\int_U |H^e|^2\, d\mu^e \leq \liminf_{k\to\infty} \int_U |H^e_k|^2\, d\mu^e_k \quad \text{for all open } U \subset B^e_{\rho_0}(\xi_0).$$

Repeating the proof of Lemma 3.7 by replacing everywhere the Hausdorff measure $\mathcal{H}^2$ of the manifold with the Euclidean Hausdorff measure $\mathcal{H}^2_e$, we obtain for all $\rho \leq \rho_0$

$$\mu^e \llcorner B^e_\rho(\xi_0) = \sum_{l=1}^{M_{\xi_0}} \mathcal{H}^2_e \llcorner \left(\operatorname{graph} u^l_{\xi_0} \cap B^e_\rho(\xi_0)\right).$$

Since the norm of the mean curvature can be bounded by the norm of the second fundamental form, it follows from Lemma 3.6 and the lower semicontinuity above that for all $\xi \in B^e_{\rho_0}(\xi_0)$ and all $\sigma > 0$ such that $B^e_\sigma(\xi) \subset B^e_{\rho_0}(\xi_0)$

$$\int_{B^e_\sigma(\xi)} |H^e|^2 d\mu^e \leq c \liminf_{k\to\infty} \int_{B^e_\sigma(\xi)} |A^e_k|^2 d\mu^e_k \leq c\sigma^\alpha.$$

By definition of the weak mean curvature and the graph representation of μ^e it follows that the functions $u^l_{\xi_0}$ are weak solutions to the weak mean curvature equation

$$\sum_{i,j=1}^{2} \partial_j \left(\sqrt{\det g_l}\; g_l^{ij} \partial_i F_l \right) = \sqrt{\det g_l}\; H^e \circ F,$$

where $F_l(x) = x + u^l_{\xi_0}(x)$ and $(g_l)_{ij} = \delta_{ij} + \partial_i u^l_{\xi_0} \cdot \partial_j u^l_{\xi_0}$.

Now first of all it follows from a standard difference quotient argument (see [GT], Theorem 8.8) that $u^l_{\xi_0} \in W^{2,2}_{loc}(L_l \cap B^e_{\rho_0}(\xi_0))$. By applying the weak mean curvature equation to a suitable test function and using the bounds on $\mathrm{D}\, u^l_{\xi_0}$ and the power decay of the Willmore energy above one gets for $x \in B^e_{\rho_0}(\xi_0) \cap L_l$ and all $\sigma > 0$ sufficiently small that

$$\int_{B^e_{\frac{\sigma}{2}}(x) \cap L_l} |\,\mathrm{D}^2 u^l_{\xi_0}|^2 \leq c \int_{B^e_{\sigma}(x) \setminus B^e_{\frac{\sigma}{2}}(x) \cap L_l} |\,\mathrm{D}^2 u^l_{\xi_0}|^2 + c\sigma^\alpha.$$

For details see [Schy]. Now again by "hole-filling" we get $\int_{B^e_{\frac{\sigma}{2}}(x) \cap L_l} |\,\mathrm{D}^2 u^l_{\xi_0}|^2 \leq \theta \int_{B^e_{\sigma}(x) \cap L_l} |\,\mathrm{D}^2 u^l_{\xi_0}|^2 + c\sigma^\alpha$ for some $\theta \in (0,1)$. Applying Lemma 5.3 we obtain (ii). Now it follows from a Lemma of Morrey (see [GT], Theorem 7.19) that

$$\mathrm{D}\, u^l_{\xi_0} \in C^{0,\alpha}(B^e_{\rho_0}(\xi_0) \cap L_l),$$

and the Lemma is proved. ∎

Therefore we have up to now shown that our limit measure μ can be written as $C^{1,\alpha} \cap W^{2,2}$-graphs away from the bad points. Now we will handle the bad points $\mathcal{B}_\varepsilon$ and prove a similar power decay as in Lemma 3.6 for balls around the bad points. From this decay it will follow that the set of bad points is actually empty. Since the bad points are discrete and since we want to prove a local decay, we assume that there is only one bad point ξ_0, and we will again work in normal coordinates around that point.

We will start with a technical but useful Lemma.

Lemma 3.9. *Consider normal coordinates centered in ξ_0 on a neighborhood $U \subset M$. For $x \in U$ let $p \in f_k^{-1}(\{x\})$ be a preimage of x and consider the tangent space $T_p f_k$. We denote with $(T_p f_k)^{\perp_e}$ the orthogonal complement in the normal coordinates, and with $\perp_e$ the projection on $(T_p f_k)^{\perp_e}$. Then for every $\epsilon > 0$ there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$, such that for $\rho < \rho_0$ and k sufficiently large*

$$\frac{|(x-\xi_0)^{\perp_e}|_e}{|x-\xi_0|_e} \leq \epsilon \quad \text{for all } x \in (\operatorname{spt}\mu_k \cap B^e_\rho(\xi_0) \setminus B^e_{\frac{\rho}{2}}(\xi_0)) \setminus \mathcal{B}_k, \tag{43}$$

where $\mathcal{B}_k \subset \operatorname{spt}\mu_k \cap B^e_{\rho_0}(\xi_0)$ with $\mu^e_k(B^e_\rho(\xi_0) \setminus B^e_{\frac{\rho}{2}}(\xi_0) \cap \mathcal{B}_k) \leq c\epsilon\rho^2$.

Proof. By Nash's Embedding Theorem we can assume that $M \subset \mathbb{R}^p$ is isometrically embedded for some p. Therefore the sequence $\{f_k\}_{k\in\mathbb{N}}$ can also be seen as a sequence of immersions in $\mathbb{R}^p$. Then Proposition 2.1 and the uniform bound on the Willmore energy $W(f_k)$ yield $\int |H_{\mathbb{S}^2 \hookrightarrow \mathbb{R}^p}|^2 d\mathcal{H}^2_{\mathbb{R}^p} \leq C$. By (3.32) in [SiL] there exists a $\rho_0 > 0$ such that for $\rho < \frac{\rho_0}{4}$ and k sufficiently large

$$\frac{|(x-\xi_0)^{\perp_{\mathbb{R}^p}}|_{\mathbb{R}^p}}{|x-\xi_0|_{\mathbb{R}^p}} \leq \frac{\varepsilon}{2} \quad \text{for all } x \in (f_k(\mathbb{S}^2) \cap B^{\mathbb{R}^p}_{2\rho}(\xi_0) \setminus B^{\mathbb{R}^p}_{\frac{\rho}{4}}(\xi_0)) \setminus \mathcal{B}_k,$$

where $\mathcal{B}_k \subset f_k(\mathbb{S}^2) \cap B^{\mathbb{R}^p}_{\frac{\rho_0}{2}}(\xi_0)$ with $\mathcal{H}^2_{\mathbb{R}^p}(f_k(\mathbb{S}^2) \cap B^{\mathbb{R}^p}_{2\rho}(\xi_0) \setminus B^{\mathbb{R}^p}_{\frac{\rho}{4}}(\xi_0) \cap \mathcal{B}_k) \leq c\varepsilon\rho^2$. Now it's easy to see that

$$\frac{|(x-\xi_0)^{\perp_e}|_e}{|x-\xi_0|_e} \leq \frac{|(x-\xi_0)^{\perp_{\mathbb{R}^p}}|_{\mathbb{R}^p}}{|x-\xi_0|_{\mathbb{R}^p}} + R(\rho),$$

where $R(\rho) \to 0$ as $\rho \to 0$. Therefore, by choosing ρ_0 sufficiently small such that for $\rho < \rho_0$ we have $R(\rho) < \varepsilon/2$, $M \cap (B^e_\rho(\xi_0) \setminus B^e_{\frac{\rho}{2}}(\xi_0)) \subset M \cap (B^{\mathbb{R}^p}_{2\rho}(\xi_0) \setminus B^{\mathbb{R}^p}_{\frac{\rho}{4}}(\xi_0))$ and $M \cap B^{\mathbb{R}^p}_{\frac{\rho_0}{2}}(\xi_0) \subset M \cap B^e_{\rho_0}(\xi_0)$, we obtain the result. ■

Now remember Definition 3.1 of the bad points. It follows that there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ such that for $\rho < \rho_0$ and k sufficiently large

$$\int_{B_{\frac{3}{2}\rho}(\xi_0) \setminus B_{\frac{\rho}{4}}(\xi_0)} |A_k|^2 \, d\mu_k < \frac{\varepsilon^2}{2}.$$

By choosing ρ_0 smaller if necessary it follows from Lemma 2.7 that

$$\int_{B^e_\rho(\xi_0) \setminus B^e_{\frac{\rho}{2}}(\xi_0)} |A^e_k|^2 \, d\mu^e_k \leq \varepsilon^2. \tag{44}$$

Moreover we get for $\rho < \rho_0$ and k sufficiently large that

$$\operatorname{spt} \mu_k \cap \partial B^e_{\frac{3}{4}\rho}(\xi_0) \neq \emptyset. \tag{45}$$

To prove this let $\xi_k \in \operatorname{spt} \mu_k$ such that $\xi_k \to \xi_0$. Thus $\operatorname{spt} \mu_k \cap B^e_{\frac{3}{4}\rho}(\xi_0) \neq \emptyset$ for k sufficiently large. Now suppose that $\operatorname{spt} \mu_k \cap \partial B^e_{\frac{3}{4}\rho}(\xi_0) = \emptyset$. Since $\operatorname{spt} \mu_k$ is connected, we get that $\operatorname{spt} \mu_k \subset B^e_{\frac{3}{4}\rho}(\xi_0)$. It follows that

$$\operatorname{diam}_h(\operatorname{spt} \mu_k) \leq c \operatorname{diam}_e(\operatorname{spt} \mu_k) \leq c\rho < c\rho_0,$$

and therefore, by choosing ρ_0 smaller if necessary, we get a contradiction to the lower diameter bound given in (12).

Let $z \in \operatorname{spt} \mu_k \cap \partial B^e_{\frac{3}{4}\rho}(\xi_0)$. Recalling Lemma 2.2, we may apply the Graphical Decomposition Lemma to get that

$$\mu^e_k \llcorner \overline{B^e_{\frac{\rho}{32}}(z)} = \sum_{l=1}^{M_k(z)} \mathcal{H}^2_e \llcorner \Big(\Big(\operatorname{graph} u^l_k \cup \bigcup_j P^{k,l}_j \Big) \cap \overline{B^e_{\frac{\rho}{32}}(z)} \Big),$$

where $\Omega^l_k = (B^e_\lambda(\pi_{L^l_k}(z)) \cap L^l_k) \setminus \bigcup_m d^l_{k,m}$ with $\lambda > \frac{\rho}{16}$, where L^l_k is a 2-dim. plane, and where the sets $d^l_{k,m} \subset L^l_k$ are pairwise disjoint closed discs. We have the following estimates:

$$M_k(z) \leq c = c(M), \tag{46}$$

$$\sum_{l,m} \operatorname{diam} d^l_{k,m} + \sum_{l,j} \operatorname{diam} P^{k,l}_j \leq c\varepsilon^{\frac{1}{2}} \rho, \tag{47}$$

$$\frac{1}{\rho} ||u^l_k||_{L^\infty(\Omega^l_k)} + ||Du^l_k||_{L^\infty(\Omega^l_k)} \leq c\varepsilon^{\frac{1}{6}}. \tag{48}$$

Remark 3.10. *Notice that $z \in \operatorname{spt} \mu_k \cap \partial B^e_{\frac{3}{4}\rho}(\xi_0)$ was arbitrary. Cover $B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0) \setminus B^e_{(\frac{3}{4}-\frac{1}{128})\rho}(\xi_0)$ by finitely many balls $B^e_{\frac{\rho}{64}}$ with center on $\partial B^e_{\frac{3}{4}\rho}(\xi_0)$ and where the number does not depend on ρ, namely*

$$B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0) \setminus B^e_{(\frac{3}{4}-\frac{1}{128})\rho}(\xi_0) \subset \bigcup_{i=1}^I B^e_{\frac{\rho}{64}}(y_i),$$

where $y_i \in \partial B^e_{\frac{3}{4}\rho}(\xi_0)$ and I is a universal constant. From this it follows that there exist points $\{z_k^1, \dots, z_k^{J_k}\} \subset$ $\operatorname{spt}\mu_k \cap \partial B^e_{\frac{3}{4}\rho}(\xi_0)$ *with $J_k \le I$, such that*

$$\operatorname{spt}\mu_k \cap B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0) \setminus B^e_{(\frac{3}{4}-\frac{1}{128})\rho}(\xi_0) \subset \bigcup_{i=1}^{J_k} B^e_{\frac{\rho}{32}}(z_i^k). \tag{49}$$

Now denote by

$$\left\{\Sigma_k^p \middle| 1 \le p \le P_k\right\} \tag{50}$$

the images via f_k of the connected components of $f_k^{-1}(B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0) \setminus B^e_{(\frac{3}{4}-\frac{1}{128})\rho}(\xi_0))$. From the above inclusion, the universal bound on J_k, the graphical decomposition from above and the universal bound on $M_k(z_i^k)$ we get that

$$P_k \le c, \tag{51}$$

where c is a universal constant independent on k and ρ.

In the next step we show that

$$\operatorname{dist}\left(\xi_0, L_k^l\right) \le c\varepsilon^{\frac{1}{6}}\rho \quad \text{for all } l \in \{1, \dots, M_k(z)\}. \tag{52}$$

To prove this notice that Proposition 3.5 and Lemma 2.7 imply

$$\mu_k^e(B^e_{\frac{\rho}{32}}(z)) \ge c\rho^2. \tag{53}$$

Moreover notice that

$$(\operatorname{graph} u_k^l \cap B^e_{\frac{\rho}{32}}(z)) \setminus \mathcal{B}_k \neq \emptyset,$$

where $\mathcal{B}_k$ was defined in Lemma 3.9. This follows from the graphical decomposition above, the diameter estimates for the sets $P_j^{k,l}$, the area estimate concerning the set $\mathcal{B}_k$ and (53).

Let $y \in (\operatorname{graph} u_k^l \cap B^e_{\frac{\rho}{32}}(z)) \setminus \mathcal{B}_k \subset (\operatorname{spt}\mu_k \cap B^e_\rho(\xi_0) \setminus B^e_{\frac{\rho}{2}}(\xi_0)) \setminus \mathcal{B}_k$. It follows that

$$|\xi_0 - \pi_{T_y f_k}(\xi_0)| \le \varepsilon|y - \xi_0| \le \varepsilon\left(|y - z| + |z - \xi_0|\right) \le c\varepsilon\rho.$$

Define the perturbed plane $\tilde{L}_k^l$ by $\tilde{L}_k^l = L_k^l + (y - \pi_{L_k^l}(y))$. Thus $\operatorname{dist}(\tilde{L}_k^l, L_k^l) = |y - \pi_{L_k^l}(y)| \le c\varepsilon^{\frac{1}{6}}\rho$ (since $y \in \operatorname{graph} u_k^l \cap B^e_{\frac{\rho}{32}}(z)$). Now Pythagoras yields $|y - \pi_{\tilde{L}_k^l}(\pi_{T_y f_k}(\xi_0))|^2 \le |y - \pi_{T_y f_k}(\xi_0)|^2 \le |y - \xi_0|^2 \le c\rho^2$. Since $T_y f_k$ can be parametrized in terms of $\mathrm{D}\, u_k^l(y)$ over $\tilde{L}_k^l$, we get that

$$|\pi_{T_y f_k}(\xi_0) - \pi_{\tilde{L}_k^l}(\pi_{T_y f_k}(\xi_0))| \le \|\mathrm{D}\, u_k^l\|_{L^\infty}|y - \pi_{\tilde{L}_k^l}(\pi_{T_y f_k}(\xi_0))| \le c\varepsilon^{\frac{1}{6}}\rho.$$

Therefore by triangle inequality we finally get (52).

Since $\operatorname{dist}(\xi_0, L_k^l) \le c\varepsilon^{\frac{1}{6}}\rho$, we may assume (after translation) that $\xi_0 \in L_k^l$ for all $l \in \{1, \dots, M_k(z)\}$ and k without changing the estimates for the functions u_k^l. Moreover we again have that $L_k^l \to L^l$ with $\xi_0 \in L^l$. Therefore for k sufficiently large we may assume that L_k^l is a fixed 2-dim. plane L^l.

Now we have that either the point z lies in one of the graphs or can be connected to one of the graphs. Without loss of generality we may assume that this graph corresponds to the function u_k^1. Subsequently we will work only with this function u_k^1, which is defined on some part of the plane L_1 with some discs $d_{k,m}^1$ removed. We will therefore drop the index 1. Define the set

$$T_k(z) = \left\{\tau \in \left(\frac{\rho}{64}, \frac{\rho}{\sqrt{2}\cdot 32}\right) \,\middle|\, \partial B^e_\tau(\pi_L(z)) \cap \bigcup_m d_{k,m} = \emptyset\right\}.$$

It follows from the diameter estimates and the selection principle in [SiL] that for $\varepsilon \le \varepsilon_0$ there exists a $\tau \in \left(\frac{\rho}{64}, \frac{\rho}{\sqrt{2}\cdot 32}\right)$ such that $\tau \in T_k(z)$ for infinitely many k.

Since $\xi_0 \in L$, it follows from the choice of τ that for $\varepsilon \le \varepsilon_0$

$$\partial B^e_{\frac{3}{4}\rho}(\xi_0) \cap \partial B^e_\tau(\pi_L(z)) \cap L = \{p_{1,k}, p_{2,k}\},$$

where $p_{1,k}, p_{2,k} \in (B^e_{\frac{\rho}{\sqrt{2}\cdot 32}}(\pi_L(z)) \cap L) \setminus \bigcup_m d_{k,m}$ are distinct points. Define the image points $z_{i,k} \in \operatorname{graph} u_k$ by

$$z_{i,k} = p_{i,k} + u_k(p_{i,k}).$$

Using the L^∞-estimates for u_k, we get for $\varepsilon \le \varepsilon_0$ that $\frac{5}{8}\rho < |z_{i,k} - \xi_0| < \frac{7}{8}\rho$, thus $\int_{B^e_{\frac{\rho}{8}}(z_{i,k})} |A^e_k|^2 \, d\mu^e_k \le \varepsilon^2$. Therefore we can again apply the Graphical Decomposition Lemma to the points $z_{i,k}$. Thus we get that

$$\mu^e_k \llcorner \overline{B^e_{\frac{\rho}{32}}(z_{i,k})} = \sum_{l=1}^{M_{i,k}(z_{i,k})} \mathcal{H}^2_e \llcorner \Big(\Big(\operatorname{graph} u^l_{i,k} \cup \bigcup_j P^{i,k,l}_j \Big) \cap \overline{B^e_{\frac{\rho}{32}}(z_{i,k})} \Big),$$

where the usual properties and estimates holds.

Now we have again that the points $z_{i,k}$ either lie in one of the graphs $u^l_{i,k}$ or can be connected to one of them. Without loss of generality let this be the graph corresponding to $u^1_{i,k}$. We will again drop the upper index. Since $z_{i,k} \in \operatorname{graph} u_k$ it follows that $\operatorname{dist}(z_{i,k}, L) \le c\varepsilon^{\frac{1}{6}}\rho$ and that $\operatorname{graph} u_{i,k}$ is connected to $\operatorname{graph} u_k$. Since the L^∞-norms of u_k and $u_{i,k}$ and their derivatives are small, we may assume (after translation and rotation as done before) that $L_{i,k} = L$.

By continuing with this procedure we get after a finite number of steps, depending not on ρ and k, an open cover of $\partial B^e_{\frac{3}{4}\rho}(\xi_0) \cap L$, which also covers the set

$$A(L) = \left\{ x + y \,\middle|\, x \in L, \operatorname{dist}\left(x, \partial B^e_{\frac{3}{4}\rho}(\xi_0) \cap L\right) < \frac{\rho}{\sqrt{2}\cdot 64}, y \in L^\perp, |y| < \frac{\rho}{\sqrt{2}\cdot 64} \right\}.$$

Now it can happen that after one "walk-around" we do not end up in the same disc of $\operatorname{spt}\mu_k \cap B^e_{\frac{\rho}{32}}(z)$ which contains the point z. But then we can proceed in a similar way and do another "walk-around". Now by construction, the "flatness" of the involved graph functions and the diameter bounds for the discs, every "walk-around" corresponds to a part of $\operatorname{spt}\mu_k$ with an area that is bounded from below by $c\rho^2$, where c is a universal constant independent of k and ρ. On the other hand we have that $\mu^e_k(B^e_\rho(\xi_0)) \le c\rho^2$. It follows that after a finite number of "walk-arounds" (which is bounded by a universal constant) we have to get back to the disc of $\operatorname{spt}\mu_k \cap B^e_{\frac{\rho}{32}}(z)$ which contains the point z.

We summarize the above procedure and the resulting properties in the following remark.

Remark 3.11. *If $\varepsilon \le \varepsilon_0$, then for each component Σ^p_k there exist a natural number k_p and a smooth function u^p_k defined on the rectangular set*

$$B^p_k = \left[\left(\left(\frac{3}{4} - \frac{1}{\sqrt{2}\cdot 64} \right)\rho, \left(\frac{3}{4} + \frac{1}{\sqrt{2}\cdot 64} \right)\rho \right) \times [0, 2\pi k_p) \right] \setminus \bigcup d^p_{k,m},$$

where the $d^p_{k,m}$ are closed discs in $\left(\left(\frac{3}{4} - \frac{1}{\sqrt{2}\cdot 64} \right)\rho, \left(\frac{3}{4} + \frac{1}{\sqrt{2}\cdot 64} \right)\rho \right) \times [0, 2\pi k_p)$, such that

$$\Sigma^p_k = \Big(R_p\left(\operatorname{graph} U^p_k\right) \cup \bigcup_j P^{k,p}_j \Big) \cap B^e_{\left(\frac{3}{4}+\frac{1}{128}\right)\rho}(\xi_0) \setminus B^e_{\left(\frac{3}{4}-\frac{1}{128}\right)\rho}(\xi_0),$$

where $\operatorname{graph} U_k^p = \{(re^{i\theta}, u_k^p(r,\theta)) \,|\, (r,\theta) \in B_k^p\}$ *and* R_p *denotes a rotation such that* $R_p(\mathbb{R}^2) = L_p$*, where* L_p *is the 2-dimensional plane with* $\xi_0 \in L_p$*. Moreover we have*

$$\sum_m \operatorname{diam} d_{k,m}^p + \sum_j \operatorname{diam} P_j^{k,p} \le c\varepsilon^{\frac{1}{2}}\rho, \quad \frac{1}{\rho}\|u_k^p\|_{L^\infty(B_k^p)} + \|\operatorname{D} u_k^p\|_{L^\infty(B_k^p)} \le c\varepsilon^{\frac{1}{6}}.$$

We may assume without loss of generality that the discs $d_{k,m}^p$ *are pairwise disjoint, since otherwise we can exchange two intersecting discs by one disc whose diameter is smaller than the sum of the diameters of the intersecting discs.*

Now let $\rho \le \rho_0$ and define the set

$$C_k(\xi_0) = \left\{ \sigma \in \left(\left(\frac{3}{4} - \frac{1}{256}\right)\rho, \left(\frac{3}{4} + \frac{1}{256}\right)\rho\right) \,\middle|\, \partial B_\sigma^e(\xi_0) \cap \bigcup_{p,j} P_j^{k,p} = \emptyset, \int_{\partial B_\sigma^e(\xi_0)} |A_k^e|^2 \, ds_k^e \le \frac{512}{\rho}\varepsilon^2 \right\}.$$

Again it follows from the diameter bounds, a simple Fubini argument and Lemma 5.2 that there exists a $\sigma \in \left(\left(\frac{3}{4} - \frac{1}{256}\right)\rho, \left(\frac{3}{4} + \frac{1}{256}\right)\rho\right)$ such that $\sigma \in C_k(\xi_0)$ for infinitely many $k \in \mathbb{N}$. For such a σ denote by

$$\left\{ \tilde{\Sigma}_k^q \,|\, 1 \le q \le Q_k \right\} \tag{54}$$

the images of the components of $f_k^{-1}(B_\sigma^e(\xi_0))$. By Remark 3.10, we get that Q_k is bounded by a universal constant which is independent of k and ρ.

Lemma 3.12. *Suppose that*

$$\frac{1}{2}\int |A_k^g|^2 \, d\mu_k^g \le 4\pi - \delta$$

for some $\delta > 0$ *(which holds in our case by Lemma 2.3). Then for* $\varepsilon \le \varepsilon_0$ *each* $\tilde{\Sigma}_k^q$ *is a topological disc, and moreover* $k_p = 1$ *for all* $1 \le p \le P_k$*.*

Proof. Fix $k \in \mathbb{N}$. First of all we construct a new immersed surface $\bar{\Sigma}_k$ such that ($\bar{\mu}_k$ denotes the associated Radon measure)

$$(i) \quad \bar{\mu}_k \llcorner B_\sigma^e(\xi_0) = \mu_k^e \llcorner B_\sigma^e(\xi_0),$$

$$(ii) \quad \left| \int_{B^e_{\left(\frac{3}{4}+\frac{1}{128}\right)\rho}(\xi_0) \setminus B_\sigma^e(\xi_0)} K_g \, d\bar{\mu}_k \right| \le c\varepsilon^{\frac{1}{3}}, \quad \text{where } K_g = \text{sectional curvature of } \bar{\Sigma}_k,$$

$$(iii) \quad \int_{M \setminus B^e_{\left(\frac{3}{4}+\frac{1}{128}\right)\rho}(\xi_0)} K_g \, d\bar{\mu}_k = 0.$$

To define $\bar{\Sigma}_k$ recall Remark 3.11 and notice that $\sum_{p,m} \operatorname{diam} d_{k,m}^p \le c\varepsilon^{\frac{1}{2}}\rho$. Now denote by M_k the number of all discs $d_{k,m}^p$. Because of the diameter estimate it follows for $\varepsilon \le \varepsilon_0$ that there exists an interval $I_k^p \subset \left(\left(\frac{3}{4} - \frac{1}{256}\right)\rho, \left(\frac{3}{4} + \frac{1}{128}\right)\rho\right)$ with $\mathcal{L}^1(I_k^p) \ge \frac{1}{512 M_k}\rho$, such that $(I_k^p \times [0, 2\pi k_p)) \cap \bigcup_m d_{k,m}^p = \emptyset$.

Let $I_k^p = (a_k^p, b_k^p)$ and $\varphi_p \in C^\infty((0,\infty) \times [0, 2\pi k_p))$ with $0 \le \varphi_p \le 1$ such that

$$\varphi_p = 1 \text{ on } (0, a_k^p) \times [0, 2\pi k_p), \quad \varphi_p = 0 \text{ on } (b_k^p, \infty) \times [0, 2\pi k_p), \quad |\operatorname{D}\varphi_p| \le \frac{c}{\rho} \text{ and } |\operatorname{D}^2 \varphi_p| \le \frac{c}{\rho^2}.$$

Now define new "components" $\bar{\Sigma}_k^p$ by

$$\bar{\Sigma}_k^p = \left(\left(R_p\left(\operatorname{graph} \bar{U}_k^p\right) \cup \bigcup_j P_j^{k,p}\right) \cap B^e_{\left(\frac{3}{4}+\frac{1}{128}\right)\rho}(\xi_0) \setminus B^e_{\left(\frac{3}{4}-\frac{1}{128}\right)\rho}(\xi_0)\right) \cup \left(L_p \setminus B^e_{\left(\frac{3}{4}+\frac{1}{128}\right)\rho}(\xi_0)\right),$$

where graph $\bar U_k^p$ is given by

$$\operatorname{graph} \bar U_k^p = \left\{ \left(re^{i\theta}, \varphi_p(r,\theta) u_k^p(r,\theta) \right) \,\middle|\, (r,\theta) \in B_k^p \right\},$$

and where again R_p denotes a rotation such that $R_p(\mathbb{R}^2) = L_p$. Namely we just "flattened out" the components Σ_k^p. Observe that by construction and Remark 3.11, outside of the ball $B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0)$, $\bar\Sigma_k^p$ is a k_p-fold covering of the plane L_p.

Now define the new surface $\bar\Sigma_k$ by

$$\bar\Sigma_k = \left(\left(f_k(\mathbb{S}^2) \cap B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0) \right) \setminus \bigcup_p \Sigma_k^p \right) \cup \bigcup_p \bar\Sigma_k^p.$$

Observe that by construction, $\bar\Sigma_k$ is an immersed surface given by an immersion $F_k : N_k \to M$, where N_k is obtained by gluing ends $E_k^p \cong \mathbb{R}^2 \setminus D$ to $f_k^{-1}(B^e_{(\frac{3}{4}-\frac{1}{128})\rho}(\xi_0))$ along $f_k^{-1}(\partial \Sigma_k^p \cap \partial B^e_{(\frac{3}{4}-\frac{1}{128})\rho}(\xi_0))$, such that outside the ball $B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0)$, $F_k|_{E_k^p}$ is a k_p-fold covering of the plane L_p.

By definition, (i) follows immediately. Since $\bar\Sigma_k \setminus B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0) = \bigcup_p L_p \setminus B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0)$, also (iii) follows directly. To prove property (ii) notice that

$$\int_{B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0)\setminus B^e_\sigma(\xi_0)} |K_g|\, d\bar\mu_k \;\le\; \int_{B^e_\rho(\xi_0)\setminus B^e_{\frac{\rho}{2}}(\xi_0)} |K_g|\, d\mu_k^e + \sum_p \int_{R_p\left(\operatorname{graph} \bar U_k^p\right)} |K_g|\, d\bar\mu_k.$$

Now the first integral on the right hand side can be estimated by

$$\int_{B^e_\rho(\xi_0)\setminus B^e_{\frac{\rho}{2}}(\xi_0)} |K_g|\, d\mu_k^e \le \frac{1}{2} \int_{B^e_\rho(\xi_0)\setminus B^e_{\frac{\rho}{2}}(\xi_0)} |A_k^e|^2\, d\mu_k^e \le \varepsilon^2.$$

The second integral can be estimated by

$$\int_{R_p\left(\operatorname{graph} \bar U_k^p\right)} |K_g|\, d\bar\mu_k \le \frac{1}{2} \int_{\operatorname{graph} \bar U_k^p} |A_e|^2\, d\bar\mu_k \le c \int_{B_k^p} |\,\mathrm{D}^2(\varphi_p u_k^p)|^2.$$

Because of the properties of the functions u_k^p and φ_p we have

$$|\,\mathrm{D}^2(\varphi_p u_k^p)|^2 \le c \left(|u_k^p|^2 |\,\mathrm{D}^2\varphi_p|^2 + |\,\mathrm{D}\, u_k^p|^2 |\,\mathrm{D}\,\varphi_p|^2 + |\varphi_p|^2 |\,\mathrm{D}^2 u_k^p|^2 \right) \le c\frac{\varepsilon^{\frac{1}{3}}}{\rho^2} + |\,\mathrm{D}^2 u_k^p|^2,$$

and therefore we get

$$\int_{B_k^p} |\,\mathrm{D}^2(\varphi_p u_k^p)|^2 \le c\varepsilon^{\frac{1}{3}} + c\int_{\operatorname{graph} U_k^p} |A_k^e|^2\, d\mu_k^e \le c\varepsilon^{\frac{1}{3}} + c\int_{B^e_\rho(\xi_0)\setminus B^e_{\frac{\rho}{2}}(\xi_0)} |A_k^e|^2\, d\mu_k^e \le c\varepsilon^{\frac{1}{3}}.$$

Thus property (ii) follows by summing over $1 \le p \le P_k \le c$.

Now denote by $N : \bar\Sigma_k \to \mathbb{S}^2$ the Gauß-map and notice that N is constant on each end. Therefore the degree of the Gauß-map $\deg(N)$ is half the Euler characteristic, and it follows from Gauß-Bonnet that

$$\deg(N) = \frac{1}{4\pi}\int K_g\, d\bar\mu_k = \frac{1}{4\pi}\int_{B^e_{(\frac{3}{4}+\frac{1}{128})\rho}(\xi_0)\setminus B^e_\sigma(\xi_0)} K_g\, d\bar\mu_k + \frac{1}{4\pi}\int_{B^e_\sigma(\xi_0)} K_g\, d\bar\mu_k.$$

Therefore we get that, using (ii) above,

$$\left| \int_{B^e_\sigma(\xi_0)} K_g\, d\bar\mu_k - 4\pi\deg(N) \right| \le c\varepsilon^{\frac{1}{3}}.$$

On the other hand it follows from the assumptions and Lemma 2.7 that

$$\left|\int_{B^e_\sigma(\xi_0)} K_g\, d\bar\mu_k\right| = \left|\int_{B^e_\sigma(\xi_0)} K_g\, d\mu^e_k\right| \le \frac{1}{2}\int_{B^e_\sigma(\xi_0)} |A^e_k|^2\, d\mu^e_k \le 4\pi - \frac{\delta}{2}$$

by choosing ρ_0 smaller if necessary. Since $\deg(N) \in \mathbb{Z}$, it follows for $\varepsilon \le \varepsilon_0$ that

$$\left|\int_{B^e_\sigma(\xi_0)} K_g\, d\mu^e_k\right| = \left|\int_{B^e_\sigma(\xi_0)} K_g\, d\bar\mu_k\right| \le c\varepsilon^{\frac{1}{3}}. \tag{55}$$

Now by the choice of σ we have for all $p = 1, \dots, P_k$ that

$$\Sigma^p_k \cap \partial B^e_\sigma(\xi_0) = \gamma_p,$$

where each γ_p is a closed, immersed smooth curve and where P_k is bounded by a universal constant. By construction and the choice of σ we have that $\gamma_p \cap \bigcup_j P^{k,p}_j = \emptyset$, therefore (see the almost graph representation of Σ^p_k above) γ_p is almost a flat circle of radius σ which can be parametrized on the interval $[0, 2\pi k_p)$. After some computations it follows from the choice of σ that (where κ denotes the geodesic curvature)

$$\left|\int_{\gamma_p} \kappa\, ds^e_k - 2\pi k_p\right| \le c\varepsilon^{\frac{1}{6}} + c\int_{\gamma_p} |A^e_k|\, ds^e_k \le c\varepsilon^{\frac{1}{6}} + c\sigma^{\frac{1}{2}}\left(\int_{\partial B^e_\sigma(\xi_0)} |A^e_k|^2\, ds^e_k\right)^{\frac{1}{2}} \le c\varepsilon^{\frac{1}{6}} + c\left(\frac{\sigma}{\rho}\right)^{\frac{1}{2}}\varepsilon \le c\varepsilon^{\frac{1}{6}},$$

and therefore it follows from the bound on P_k that

$$\left|\int_{\partial B^e_\sigma(\xi_0)} \kappa\, ds^e_k - 2\pi\sum_{p=1}^{P_k} k_p\right| \le c\varepsilon^{\frac{1}{6}}. \tag{56}$$

Now the Euler characteristic of $\tilde\Sigma^q_k$ is given by

$$\chi(\tilde\Sigma^q_k) = 2(1 - g_q) - b_q,$$

where b_q is the number of boundary components of $\tilde\Sigma^q_k$ and g_q is the genus of the closed surface which arises by gluing b_q topological discs. Especially we have

$$b_q \ge 1 \quad \text{and} \quad \sum_{q=1}^{Q_k} b_q = P_k.$$

By summing over q we get that the Euler characteristic of $\bigcup_{q=1}^{Q_k} \tilde\Sigma^q_k$ is

$$\chi_E\left(\bigcup_{q=1}^{Q_k} \tilde\Sigma^q_k\right) = 2(Q_k - g) - P_k, \quad \text{where } g = \sum_{q=1}^{Q_k} g_q \ge 0.$$

Since $Q_k \le P_k$, we finally get that

$$P_k \ge 2(Q_k - g) - P_k = \frac{1}{2\pi}\int_{B^e_\sigma(\xi_0)} K_g\, d\mu^e_k + \frac{1}{2\pi}\int_{\partial B^e_\sigma(\xi_0)} \kappa\, ds^e_k \ge \sum_{p=1}^{P_k} k_p - c\varepsilon^{\frac{1}{6}} \ge P_k - c\varepsilon^{\frac{1}{6}}.$$

Since $2(Q_k - g) - P_k \in \mathbb{N}$, it follows for $\varepsilon \le \varepsilon_0$ that $P_k = 2(Q_k - g) - P_k$. Since $Q_k \le P_k$ we get that $Q_k = P_k$ and $g = 0$. Thus $g_q = 0$ and $b_q = 1$ for all q. This yields that the Euler characteristic of $\tilde\Sigma^q_k$ is 1 and therefore each $\tilde\Sigma^q_k$ is a topological disc. Moreover the estimate above yields $k_p = 1$. ■

Now define the sets

$$C_k^p = \left\{ s \in \left(0, \frac{\rho}{128}\right) \,\middle|\, \left(\left(\frac{3}{4}\rho + s\right) \times [0, 2\pi)\right) \cap \bigcup_m d_{k,m}^p = \emptyset \right\},$$

$$D_k^p = \left\{ s \in C_k^p \,\middle|\, \int_{R_p\left(\operatorname{graph} U_k^p|_{\left(\left(\frac{3}{4}\rho+s\right)\times[0,2\pi)\right)}\right)} |A_k^e|^2 \, ds_k^e \leq \frac{512}{\rho} \int_{\Sigma_k^p} |A_k^e|^2 \, d\mu_k^e \right\}.$$

By the diameter estimates for the discs $d_{k,m}^p$, again a simple Fubini-type argument and Lemma 5.2 there exists a $s \in \left(0, \frac{\rho}{128}\right)$ such that $s \in D_k^p$ for infinitely many k. It follows that u_k^p is defined on the line $\left(\frac{3}{4}\rho + s\right) \times [0, 2\pi)$. Now it follows from Lemma 3.12 that $R_p\left(\operatorname{graph} U_k^p|_{\left(\left(\frac{3}{4}\rho+s\right)\times[0,2\pi)\right)}\right)$ divides $f_k(\mathbb{S}^2)$ into two topological discs $\Sigma_1^{k,p}, \Sigma_2^{k,p}$, one of them, w.l.o.g. $\Sigma_1^{k,p}$, intersecting $B^e_{\frac{3}{4}\rho}(\xi_0)$. From the estimates for the function u_k^p and the choice of s we get that $\Sigma_1^{k,p} \subset B^e_{\left(\frac{3}{4}+\frac{1}{128}\right)\rho}(\xi_0)$, and Lemma 2.2 yields $\mu_k^e(\Sigma_1^{k,p}) \leq c\rho^2$.

According to the Lemma 5.1, let $w_k^p \in C^\infty\left(B^e_{\frac{3}{4}\rho+s}(\xi_0) \cap L_p, L_p^\perp\right)$ be an extension of $R_p(U_k^p)$ restricted to $\partial B^e_{\frac{3}{4}\rho+s}(\xi_0) \cap L_p$. In view of the estimates for u_k^p, and thus for w_k^p, we get that $\operatorname{graph} w_k^p \subset B^e_{\left(\frac{3}{4}+\frac{1}{128}\right)\rho}(\xi_0)$.

Now we can define the surface $\tilde{\Sigma}_k$ by

$$\tilde{\Sigma}_k = \left(f_k(\mathbb{S}^2) \setminus \bigcup_p \Sigma_1^{k,p} \right) \cup \bigcup_p \operatorname{graph} w_k^p,$$

and we can do exactly the same as in the proof of Lemma 3.6 to get the same power decay as for the good points, but now for balls around the bad points. But by definition the bad points do not allow a decay like this, and therefore we have proved that there are no bad points.

Up to now we have shown that the limit measure μ is locally given by $C^{1,\alpha} \cap W^{2,2}$-graphs. In the next step we show that there exists a $C^{1,\alpha} \cap W^{2,2}$-immersion $f : \mathbb{S}^2 \hookrightarrow M$ such that μ is the Radon measure associated to this immersion f. To prove this we will apply a result of Breuning [Breu], which involves so called generalized (r, λ)-immersions (for the Definition see Definition 5.5).

For that recall Lemma 3.4 and Lemma 3.7, namely for every $\xi \in \operatorname{spt} \mu$ there exist a radius $r_\xi > 0$ and a natural number $K_\xi \in \mathbb{N}$ such that

(i) $\mu_k \llcorner B^e_{r_\xi}(\xi) = \sum_{l=1}^{M_\xi} \mathcal{H}^2 \llcorner \left(\left(\operatorname{graph} u_l^k \cup \bigcup_j P_j^{k,l} \right) \cap B^e_{r_\xi}(\xi) \right)$ for $k \geq K_\xi$, where u_l^k are smooth functions defined on appropriate planes L_l with the usual properties and estimates,

(ii) $\mu \llcorner B^e_{r_\xi}(\xi) = \sum_{l=1}^{M_\xi} \mathcal{H}^2 \llcorner \left(\operatorname{graph} u_l \cap B^e_{r_\xi}(\xi) \right)$, where u_l are $C^{1,\alpha} \cap W^{2,2}$-functions defined on L_l.

For $\xi \in \operatorname{spt} \mu$ let $\rho_\xi := \sup\{r_\xi > 0 \text{ such that (i) and (ii) holds}\}$. Since $\operatorname{spt} \mu$ is compact, it follows that $\rho := \inf\{\rho_\xi : \xi \in \operatorname{spt} \mu\} > 0$. Notice that (i) and (ii) holds for ρ instead of r_ξ.

By compactness of $\operatorname{spt} \mu$ there exist $\{\xi_1, \ldots, \xi_I\} \subset \operatorname{spt} \mu$ such that $\operatorname{spt} \mu \subset \bigcup_{i=1}^I B^e_{\frac{\rho}{4}}(\xi_i)$. From the Hausdorff distance sense convergence it also follows that $f_k(\mathbb{S}^2) \subset \bigcup_{i=1}^I B^e_{\frac{\rho}{4}}(\xi_i)$ for k sufficiently large.

Now (i) yields $\mu_k \llcorner B^e_\rho(\xi_i) = \sum_{l=1}^{M_{\xi_i}} \mathcal{H}^2 \llcorner \left(\left(\operatorname{graph} u_l^{k,i} \cup \bigcup_j P_j^{k,l,i} \right) \cap B^e_\rho(\xi_i) \right)$. By the diameter estimates for the $P_j^{k,l,i}$ and the selection principle 5.2 there exists a $\frac{\bar{\rho}}{2} \in (\frac{\rho}{4}, \frac{\rho}{2})$ such that $\partial B_{\bar{\rho}}(\xi_i) \cap \bigcup_{l,j} P_j^{k,l,i} = \emptyset$

for all $i \in \{1, \dots, I\}$ and infinitely many k.

Of course we still have that $f_k(\mathbb{S}^2) \cup \operatorname{spt}\mu \subset \bigcup_{i=1}^I B^e_{\frac{\bar\rho}{2}}(\xi_i)$, and also the graphical decomposition as in (i) and (ii) still holds in $B^e_{\bar\rho}(\xi_i)$.

First consider $f_k(\mathbb{S}^2) \cap B^e_{\bar\rho}(\xi_1)$. We replace the pimples $\{P_j^{k,l,1}\}_{l,j}$ of $f_k(\mathbb{S}^2) \cap B^e_{\bar\rho}(\xi_1)$ with the extension Lemma 5.1 as done in the proof of Lemma 3.6 by graphs of functions with small C^1-norms defined on the discs $d^{l,1}_{k,m}$. It follows that the sum of the areas of all these graphs is bounded by $c\sum_m (\operatorname{diam} d^{l,1}_{k,m})^2 \le c\varepsilon\bar\rho$, which follows from the diameter estimates for the discs. Notice that by the choice of $\bar\rho$, no pimple intersects $\partial B^e_{\bar\rho}(\xi_1)$, and we obtain a new $C^{1,1}$-immersion $f^1_k : \mathbb{S}^2 \hookrightarrow M$ such that

$$f_k(\mathbb{S}^2) \setminus B^e_{\bar\rho}(\xi_1) = f^1_k(\mathbb{S}^2) \setminus B^e_{\bar\rho}(\xi_1), \quad f^1_k(\mathbb{S}^2) \cap B^e_{\bar\rho}(\xi_1) = \bigcup_{l=1}^{M_1} \operatorname{graph} w_l^{k,1} \cap B^e_{\bar\rho}(\xi_1). \tag{57}$$

Moreover the above area estimate yields $\mu^1_k(M) \le \mu_k(M) + c\varepsilon\bar\rho$, where μ^1_k denotes the Radon measure associated to f^1_k. Observe that by construction $w_l^{k,1} : L^1_l \cap B^e_{\bar\rho}(\xi_1) \to (L^1_l)^\perp$ are $C^{1,1}$-functions satisfying $\frac{1}{\bar\rho}\|w_l^{k,1}\|_{L^\infty} + \|\operatorname{D} w_l^{k,1}\|_{L^\infty} \le c\varepsilon^{\frac16} + \delta_k$, where $\delta_k \to 0$. By construction of the limit graphs representing μ (see the part after Lemma 3.6 and Lemma 3.7) we have that $w_l^{k,1} \to u_{l,1}$ uniformly, where $u_{l,1}$ are the graph functions representing μ, namely $\mu \llcorner B^e_{\bar\rho}(\xi_1) = \sum_{l=1}^{M_1} \mathcal{H}^2 \llcorner \big(\operatorname{graph} u_{l,1} \cap B^e_{\bar\rho}(\xi_1)\big)$.

Now consider a point $\xi_j \in \{\xi_1, \dots, \xi_I\}$ such that $B^e_{\frac{\bar\rho}{2}}(\xi_1) \cap B^e_{\frac{\bar\rho}{2}}(\xi_j) \ne \emptyset$, without loss of generality $j = 2$. Recall that $\mu_k \llcorner B^e_{\bar\rho}(\xi_2) = \sum_{l=1}^{M_2} \mathcal{H}^2 \llcorner \Big(\Big(\operatorname{graph} u_l^{k,2} \cup \bigcup_j P_j^{k,l,2}\Big) \cap B^e_{\bar\rho}(\xi_2)\Big)$, where $u_l^{k,2}$ are smooth functions defined on appropriate planes L^2_l.

Observe that $f^1_k(\mathbb{S}^2) \cap B^e_{\bar\rho}(\xi_1) \cap B^e_{\bar\rho}(\xi_2) = \bigcup_{l=1}^{M_1} \operatorname{graph} w_l^{k,1} \cap B^e_{\bar\rho}(\xi_1) \cap B^e_{\bar\rho}(\xi_2)$, and because of the C^1-estimates for $w_l^{k,1}$ and $u_l^{k,2}$ and the diameter estimate for the pimples, these functions can be written as graphs over the planes L^2_l satisfying analogous estimates. We conclude that $f^1_k(\mathbb{S}^2) \cap B^e_{\bar\rho}(\xi_1) \cap B^e_{\bar\rho}(\xi_2) = \bigcup_{l=1}^{M_1} \operatorname{graph} w_l^{k,2} \cap B^e_{\bar\rho}(\xi_1) \cap B^e_{\bar\rho}(\xi_2)$, where now the functions $w_l^{k,2}$ are defined on the planes L^2_l. From (57), the graphical representation of $f_k(\mathbb{S}^2) \cap B^e_{\bar\rho}(\xi_2) \setminus B^e_{\bar\rho}(\xi_1)$ and the choice of $\bar\rho$, we can replace the pimples inside $B^e_{\bar\rho}(\xi_2) \setminus B^e_{\bar\rho}(\xi_1)$ with new graphs as done before obtaining a new $C^{1,1}$-immersion $f^2_k : \mathbb{S}^2 \hookrightarrow M$ which is the union of graphs (without pimples) in both balls such that the corresponding graph functions converge uniformly to the graph functions representing μ, and such that $\mu^2_k(M) \le \mu_k(M) + 2c\varepsilon\bar\rho$, where μ^2_k denotes the Radon measure associated to f^1_k.

Repeating the above procedure I times we obtain a $C^{1,1}$-immersion $\tilde f_k := f^I_k : \mathbb{S}^2 \hookrightarrow M$ such that $\mu^I_k(M) \le \mu_k(M) + Ic\varepsilon\bar\rho$, where μ^I_k denotes the Radon measure associated to f^I_k. Because of the uniform area estimate given in Proposition 2.1 we have on particular $\mu^I_k(M) \le C$.

Now we show that $\tilde f_k$ is actually a generalized (r, λ)-immersion. Recall that $\operatorname{spt}\mu \subset \bigcup_{i=1}^I B^e_{\bar\rho}(\xi_i)$ is an open cover of $\operatorname{spt}\mu$. By Lebesgue's Lemma there exists the Lebesgue number $\tilde\rho > 0$ such that for every $\xi \in \operatorname{spt}\mu$ we have $B^e_{\tilde\rho}(\xi) \subset B^e_{\bar\rho}(\xi_i)$ for some $i \in \{1, \dots, I\}$. Now observe that also $\tilde f_k(\mathbb{S}^2)$ converges to $\operatorname{spt}\mu$ in the Hausdorff distance sense (which follows from the uniform convergence of the corresponding graphs), thus $B_{\frac{\tilde\rho}{2}}(\tilde f_k(\mathbb{S}^2)) \subset \bigcup_{i=1}^I B^e_{\bar\rho}(\xi_i)$ for k sufficiently large. Let $p \in \mathbb{S}^2$ and observe that $B^e_{\frac{\tilde\rho}{2}}(\tilde f_k(p)) \subset B^e_{\bar\rho}(\xi_i)$ for some i. Therefore by construction of $\tilde f_k$ we have

$$\tilde f_k(\mathbb{S}^2) \cap B^e_{\frac{\tilde\rho}{2}}(\tilde f_k(p)) = \bigcup_{l=1}^{M_i} \operatorname{graph} w_l^{k,i} \cap B^e_{\frac{\tilde\rho}{2}}(\tilde f_k(p)),$$

where $w_l^{k,i} : L^i_l \cap B^e_{\bar\rho}(\xi_i) \to (L^l_j)^\perp$ are $C^{1,1}$-functions satisfying $\|\operatorname{D} w_l^{k,i}\|_{L^\infty} \le c\varepsilon^{\frac16} + \delta_k$, where $\delta_k \to 0$. Now recall that by Nash's embedding theorem we can assume that our ambient manifold M is isometrically embedded in some $\mathbb{R}^p$. Denote by $A^k_{p,L^i_l}:\mathbb{R}^p\to\mathbb{R}^p$ an Euclidean isometry which maps the origin to $\tilde f_k(p)$ and the subspace $\mathbb{R}^2\times\{0\}$ onto $\tilde f_k(p)+(L^i_l-\pi^i_l(\tilde f_k(p)))$, where π^i_l denotes the orthogonal projection onto L^i_l. We get that

$$\tilde f_k(\mathbb{S}^2)\cap B^e_{\frac{\tilde\rho}{2}}(\tilde f_k(p))=\bigcup_{l=1}^{M_i}A^k_{p,L^i_l}(\operatorname{graph}\tilde w^{k,i}_l\cap B^e_{\frac{\tilde\rho}{2}}(0)),$$

where $\tilde w^{k,i}_l:\mathbb{R}^2\cap B^e_{\tilde\rho}(0)\to(\mathbb{R}^2)^\perp$ are $C^{1,1}$-functions given by

$$\tilde w^{k,i}_l(x)=\left(A^k_{p,L^i_l}\right)^{-1}\left(w^{k,i}_l\left(A^k_{p,L^i_l}(x)-\left(\tilde f_k(p)-\pi^i_l(\tilde f_k(p))\right)\right)-\left(\tilde f_k(p)-\pi^i_l(\tilde f_k(p))\right)\right),$$

and which satisfy $\|\operatorname{D}\tilde w^{k,i}_l\|_{L^\infty}\le c\varepsilon^{\frac16}+\delta_k$, where $\delta_k\to0$. Now denote by $U^k_{\frac{\tilde\rho}{2},p}\subset\mathbb{S}^2$ the component of $(\pi\circ(A^k_{p,L^i_l})^{-1}\circ\tilde f_k)^{-1}(\mathbb{R}^2\cap B^e_{\frac{\tilde\rho}{2}}(0))$ containing p, where $\pi:\mathbb{R}^p\to\mathbb{R}^2$ is the projection on the first two coordinates. By construction we have $(A^k_{p,L^i_l})^{-1}\circ\tilde f_k(U^k_{\frac{\tilde\rho}{2},p})=\operatorname{graph}\tilde w^{k,i}_l\cap B^e_{\frac{\tilde\rho}{2}}(0)$ for some $l\in\{1,\dots,M_i\}$.

Finally, for given $\lambda<\frac14$ we get that for $\varepsilon\le\varepsilon_0$ and k sufficiently large that $\tilde f_k:\mathbb{S}^2\hookrightarrow M$ is a generalized $(\frac{\tilde\rho}{2},\lambda)$-immersion, namely $\{\tilde f_k\}_{k\in\mathbb{N}}\subset\mathcal{F}^1_C(r,\lambda)$ for $r=\frac{\tilde\rho}{2}$ and $\lambda<\frac14$.

By the compactness Theorem 5.6 for generalized (r,λ)-immersions of Breuning [Breu], there exist a generalized $(\frac{\tilde\rho}{2},\lambda)$-function $f:\mathbb{S}^2\hookrightarrow M$ (see Definition 5.5) and diffeomorphisms $\phi_k:\mathbb{S}^2\to\mathbb{S}^2$ such that $\tilde f_k\circ\phi_k\to f$ uniformly. Let us briefly recall Breuning's construction of the limit f: Let $q\in\mathbb{S}^2$ and $q_k=\phi_k(q)$. By the uniform convergence of $\tilde f_k\circ\phi_k$ we have that for k sufficiently large $B^e_{\frac{\tilde\rho}{2}}(\tilde f_k(q_k))\subset B^e_{\tilde\rho}(\xi_i)$ for some i. By the construction above we know that for each large k

$$(A^k_{q_k,L^i_l})^{-1}\circ\tilde f_k(U^k_{\frac{\tilde\rho}{2},q_k})=\operatorname{graph}\tilde w^{k,i}_l\cap B^e_{\frac{\tilde\rho}{2}}(0)\quad\text{for some } l\in\{1,\dots,M_i\}.$$

Now Breuning proves that there exist λ-Lipschitz functions $\tilde u^i_l$ such that

$$\tilde w^{k,i}_l\to\tilde u^i_l\quad,\quad A^k_{q_k,L^i_l}\to A_{q,L^i_l}\quad\text{and}\quad(A_{q,L^i_l})^{-1}\circ f(U_{\frac{\tilde\rho}{2},q})=\operatorname{graph}\tilde u^i_l\cap B^e_{\frac{\tilde\rho}{2}}(0).\tag{58}$$

On the other hand we know from the representation of the limit measure μ that

$$\mu\llcorner B^e_{\tilde\rho}(\xi_i)=\sum_{l=1}^{M_i}\mathcal{H}^2\llcorner\left(\operatorname{graph}u^i_l\cap B^e_{\tilde\rho}(\xi_i)\right),$$

where u^i_l are $C^{1,\alpha}\cap W^{2,2}$-functions defined on the planes L^i_l. By construction of these limit graphs as carried out before we get that the function given by

$$A^k_{q_k,L^i_l}\left(\tilde w^{k,i}_l\left(\left(A^k_{q_k,L^i_l}\right)^{-1}\left(x+\tilde f_k(q_k)-\pi^i_l(\tilde f_k(q_k))\right)\right)\right)+\left(\tilde f_k(q_k)-\pi^i_l(\tilde f_k(q_k))\right)$$

converges uniformly to u^i_l. Since $\tilde f_k(q_k)=\tilde f_k(\phi_k(q))\to f(q)$, it follows from (58) that

$$\tilde u^i_l(x)=\left(A_{q,L^i_l}\right)^{-1}\left(u^i_l\left(A_{q,L^i_l}(x)-\left(f(q)-\pi^i_l(f(q))\right)\right)-\left(f(q)-\pi^i_l(f(q))\right)\right).$$

Therefore the function $\tilde u^i_l$ is actually $C^{1,\alpha}\cap W^{2,2}$ and $A_{q,L^i_l}(\operatorname{graph}\tilde u^i_l)=\operatorname{graph}u^i_l$. Thus

$$f(U_{\frac{\tilde\rho}{2},q})=A_{q,L^i_l}(\operatorname{graph}\tilde u^i_l\cap B^e_{\frac{\tilde\rho}{2}}(0))=\operatorname{graph}u^i_l\cap B^e_{\frac{\tilde\rho}{2}}(f(q)).$$

We have therefore shown that the generalized $(\frac{\tilde\rho}{2},\lambda)$-function $f:\mathbb{S}^2\hookrightarrow M$ is actually a $C^{1,\alpha}\cap W^{2,2}$-immersion and that μ is the Radon measure associated to the immersion f.

Finally we show that f satisfies the Euler-Lagrange equation and is smooth. First we prove

$$E(f) = \inf\{E(F) | F \in C^1 \cap W^{2,2}(\mathbb{S}^2, M) \text{ immersed}\}. \tag{59}$$

A standard approximation argument implies that the right hand side equals the infimum $\inf_{[\mathbb{S}^2,M]} E(f)$ among smooth immersions. Therefore, (59) follows if we prove the lower semicontinuity of the functional, i.e.

$$E(f) \leq \liminf_{k\to\infty} E(f_k). \tag{60}$$

For this we employ results about curvature varifolds due to Hutchinson [Hu1]. For convenience of the reader, we recall the main points. For an open set $U \subset \mathbb{R}^n$, let $f \in C^1 \cap W^{2,2}_{loc}(\Sigma, U)$ be a properly immersed surface with induced metric g. For any vector field $Y \in C^1_c(\Sigma, \mathbb{R}^n)$ we have the first variation formula

$$\int_\Sigma \operatorname{div}_g Y \, d\mu_g = -\int_\Sigma \langle H, Y\rangle \, d\mu_g. \tag{61}$$

The projection $P(p) : \mathbb{R}^n \to \mathbb{R}^n$ onto the tangent space $T_p f = Df(p) T_p\Sigma$ is given by

$$P = g^{\alpha\beta}\langle \partial_\alpha f, \cdot\,\rangle \partial_\beta f \in C^0 \cap W^{1,2}_{loc}(\Sigma, \mathbb{R}^{n\times n}).$$

We define a vector-valued bilinear form $\overline{B}(p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ of class L^2_{loc} by the formula

$$\overline{B}(e_i, e_j) = g^{\alpha\beta}\langle \partial_\alpha f, e_i\rangle (\partial_\beta P)\cdot e_j.$$

Note that $\overline{B}(\partial_\alpha f, \cdot\,) = \partial_\alpha P$. Now we take $Y = X \circ G_f$ in (61), where $X \in C^1(U \times \mathbb{R}^{n\times n}, \mathbb{R}^n)$ has compact support in the first variable and $G_f(p) = (f(p), P(p))$ is the Gauß map. Compute

$$\begin{aligned}
\operatorname{div}_g Y &= g^{\alpha\beta}\langle \partial_\alpha Y, \partial_\beta f\rangle \\
&= g^{\alpha\beta}\langle (D_x X)\circ G_f\, \partial_\alpha f, \partial_\beta f\rangle + g^{\alpha\beta}\langle (D_P X)\circ G_f\, \partial_\alpha P, \partial_\beta f\rangle \\
&= \operatorname{tr}^{Tf}(D_x X)\circ G_f + (\partial_{P^k_j} X^i)\circ G_f\, \overline{B}^k_{ij}.
\end{aligned}$$

The integral 2-varifold V_f induced by f on $G_2(U) = U \times G(2,p)$ has weight measure $\mu_f = \mathcal{H}^2 \llcorner \theta_f$, where $\theta_f(x) = \# f^{-1}\{x\}$ is the multiplicity function, and we have

$$\int_{G_2(U)} \phi(x,P)\, dV_f(x,P) = \int_U \phi(x, T_x\mu_f)\, d\mu_f(x) \quad \text{for all } \phi \in C^0_c(G_2(U)).$$

Following Section 5.2 in [Hu1], we show that V_f has generalized curvature given by

$$B(x) = \frac{1}{\theta_f(x)} \sum_{p\in f^{-1}\{x\}} \overline{B}(p) \quad \text{for } x \in f(\Sigma).$$

We put $B = 0$ outside $f(\Sigma)$. To prove the claim we must verify that

$$\int_{G_2(U)} \operatorname{tr}^P(D_x X)\, dV_f + \int_{G_2(U)} \partial_{P^k_j} X^i\, B^k_{ij}\, dV_f = -\int_{G_2(U)} \langle B_{ii}, X\rangle\, dV_f. \tag{62}$$

This will follow from the first variation identity above, recalling that $T_x\mu_f$ exists for μ_f-almost every $x \in U$, and hence $T_x\mu_f = T_p f$ for $p \in f^{-1}\{x\}$. We compute

$$\int_{G_2(U)} \operatorname{tr}^P(D_x X)(x,P)\, dV_f(x,P) = \int_\Sigma \operatorname{tr}^{T_p f}(D_x X)(G_f(p))\, d\mu_g(p).$$

Secondly,

$$\int_{G_2(U)} (\partial_{P^k_j} X^i)(x,P) B^k_{ij}(x)\, dV_f(x,P) \quad = \quad \int_\Sigma (\partial_{P^k_j} X^i)(G_f(p)) \overline{B}^k_{ij}(p)\, d\mu_g(p).$$

Similarly,

$$\int_{G_2(U)} \langle B_{ii}(x), X(x,P)\rangle\, dV_f(x,P) \quad = \quad \int_\Sigma \big\langle \overline{B}_{ii}(p), X(G_f(p))\big\rangle\, d\mu_g(p).$$

To calculate $\overline{B}$, we first observe that $\overline{B}(N,\cdot\,) = 0$ if N is normal along f. We further calculate

$$\begin{aligned}
\overline{B}(\partial_\alpha f, \partial_\beta f) &= \partial_\alpha P \cdot \partial_\beta f = \partial^2_{\alpha\beta} f - P \partial^2_{\alpha\beta} f = A_{\alpha\beta} \\
\overline{B}(\partial_\alpha f, N) &= \partial_\alpha P \cdot N = -P \partial_\alpha N = -g^{\beta\gamma} \langle \partial_\alpha N, \partial_\beta f\rangle \partial_\gamma f = g^{\beta\gamma} \langle N, A_{\alpha\beta}\rangle \partial_\gamma f.
\end{aligned}$$

In particular $\overline{B}_{ii} = H$ which completes the proof of (62). We will also need that $B \in L^1_{loc}(\mu_f)$ is uniquely determined by (62), see [Hu1], Proposition 5.2.2.

Next consider a sequence of varifolds $V_k \to V$ weakly in $G_2(U)$, and functions $\psi_k \in L^2(V_k, \mathbb{R}^m)$ with

$$C_0 := \lim_{k\to\infty} \|\psi_k\|_{L^2(V_k)} < \infty.$$

Define the linear functionals $\Lambda_k : C^0_c(G_2(U), \mathbb{R}^m) \to \mathbb{R}$, $\Lambda_k(\phi) = \int_{G_2(U)} \langle \phi, \psi_k\rangle\, dV_k$. Clearly

$$|\Lambda_k(\phi)| \leq \|\psi_k\|_{L^2(V_k)} V_k(\operatorname{spt}\phi)^{\frac12} \|\phi\|_{C^0(U)}.$$

By the Banach-Alaoglu theorem, we have $\Lambda_k \to \Lambda$ in $C^0_c(G_2(U))'$ for a subsequence, and we get

$$|\Lambda(\phi)| \leq C_0\, V(\operatorname{spt}\phi)^{\frac12} \|\phi\|_{C^0(U)} \quad \text{ for } \phi \in C^0_c(G_2(U), \mathbb{R}^m).$$

By the theorem of Riesz, the functional Λ has a representation

$$\Lambda(\phi) = \int_{G_2(U)} \langle \phi, \nu\rangle\, d|\Lambda|,$$

where $|\Lambda|$ is the variation measure and $\nu : G_2(U) \to \mathbb{R}^m$ is Borel measurable with $|\nu| = 1$ almost everywhere with respect to $|\Lambda|$. But $|\Lambda|$ is absolutely continuous with respect to V, hence we have $|\Lambda| = V \llcorner \theta$ for some function $\theta \in L^1_{loc}(V, \mathbb{R}^+_0)$. Put $\psi = \theta\nu \in L^1_{loc}(V, \mathbb{R}^m)$ to obtain

$$\int_{G_2(U)} \langle \phi, \psi\rangle\, dV = \lim_{k\to\infty} \int_{G_2(U)} \langle \phi, \psi_k\rangle\, dV_k \quad \text{ for all } \phi \in C^0_c(G_2(U), \mathbb{R}^m).$$

Now for any $\phi \in C^0_c(G_2(U), \mathbb{R}^m)$ we can estimate

$$\Lambda(\phi) = \int_{G_2(U)} \langle \phi, \psi\rangle\, dV \leq C_0 \lim_{k\to\infty} \|\phi\|_{L^2(V_k)} = C_0 \|\phi\|_{L^2(V)}.$$

Thus Λ extends continuously to $L^2(V, \mathbb{R}^m)$ and hence $\psi \in L^2(V, \mathbb{R}^m)$. Moreover, for any $\eta \in C^0_c(U, \mathbb{R}^+_0)$ we get by Cauchy-Schwarz

$$\int_U \eta |\psi|^2\, dV \leq \Big(\int_U \eta |\psi|^2\, dV\Big)^{\frac12} \liminf_{k\to\infty} \Big(\int_U \eta |\psi_k|^2\, dV_k\Big)^{\frac12}.$$

Canceling we obtain

$$\int_U \eta |\psi|^2\, dV \leq \liminf_{k\to\infty} \int_U \eta |\psi_k|^2\, dV_k.$$

We now return to the setting of immersed surfaces. Let $f_k \in C^1 \cap W^{2,2}_{loc}(\Sigma_k, U)$, $f \in C^1 \cap W^{2,2}_{loc}(\Sigma, U)$ be properly immersed. Assume that

$$\|A_{f_k}\|_{L^2(\Sigma_k)} \leq C_0, \quad \text{and} \quad V_{f_k} \to V_f \quad \text{as varifolds in } U.$$

Let us fix a cutoff function $\eta \in C^0_c(U, \mathbb{R}^+_0)$. From the above we see $|\overline{B}|^2 = 2|A|^2$ and

$$\begin{aligned}
\int_U \eta(x)|B(x)|^2 \, d\mu_f &= \int_U \eta(x)\theta_f(x)^{-2}\Big|\sum_{p\in f^{-1}\{x\}} \overline{B}(p)\Big|^2 \, d\mu_f(x) \\
&\leq \int_U \eta(x) \sum_{p\in f^{-1}\{x\}} |\overline{B}(p)|^2 \, d\mathcal{H}^2(x) \\
&= 2\int_\Sigma \eta \circ f |A|^2 \, d\mu_g.
\end{aligned}$$

In order to have equality for f in this argument, we make the technical assumption that f is injective. It now follows that B_{f_k} is bounded in $L^2(V_k)$, and $V_k \llcorner B_{f_k}$ converges to $V \llcorner B$ as varifolds, for some $B \in L^2(V)$. Taking limits in (62) shows that V has generalized second fundamental form equal to B, hence we have $B = B_f$ by uniqueness. We conclude

$$\begin{aligned}
\int_\Sigma \eta \circ f |A_f|^2 \, d\mu_g &= \frac{1}{2}\int_{G_2(U)} \eta \, |B|^2 \, dV_f \\
&\leq \frac{1}{2}\liminf_{k\to\infty} \int_{G_2(U)} \eta \, |B_{f_k}|^2 \, dV_{f_k} \\
&\leq \liminf_{k\to\infty} \int_{G_2(U)} \eta \circ f_k |A_{f_k}|^2 \, d\mu_{g_k}.
\end{aligned}$$

This proves the local lower semicontinuity of the functional $E(f)$. Finally, assume that $f : \Sigma \to M \cap U$ where $M \subset \mathbb{R}^n$ is a C^2 submanifold. The second fundamental forms in M and in $\mathbb{R}^n$ differ only by a first order term, more precisely

$$\int_\Sigma |A|^2 \, d\mu_g = \int_\Sigma |A^{\mathbb{R}^n}|^2 \, d\mu_g - \int_\Sigma |A^M_{Tf\times Tf}|^2 \, d\mu_g.$$

Here by $A^{\mathbb{R}^n}$ we mean the second fundamental form in $\mathbb{R}^n$, while A now refers to the second fundamental form in M. Extending the second fundamental form A^M of $M \subset \mathbb{R}^n$ to $TM^\perp$ by zero, we may write

$$\begin{aligned}
\int_\Sigma \eta \circ f |A^M_{Tf\times Tf}|^2 \, d\mu_g &= \int_U \eta(x) \sum_{p\in f^{-1}\{x\}} |A^M(x)(P(p)e_i, P(p)e_j|^2 \, d\mathcal{H}^2(x) \\
&= \int_{G_2(U)} \eta(x)|A^M(x)(Pe_i, Pe_j)|^2 \, dV_f(x, P).
\end{aligned}$$

The last expression is continuous under the convergence $V_{f_k} \to V_f$. Therefore the L^2 integral of the second fundamental form in M is also lower semicontinuous.

To prove the lower semicontinuity of the full functional $E(f)$, we cover the image of the limit surface by neighborhoods on which we have a local graph description (with pimples for the f_k). Then we choose a subordinate partition of unity and apply the above lower semicontinuity statement to each of the graphs. Summing up yields the desired result.

Now by construction and lower semicontinuity it follows that the limit immersion f minimizes E among $C^1 \cap W^{2,2}$-immersions, in particular it satisfies the Euler-Lagrange equation.

To compute the Euler-Lagrange equation, let $f \in W^{2,2} \cap C^{1,\alpha}(U, \mathbb{R}^3)$, $f(x) = (x, u(x))$, be a graph given in local coordinates in M. The functional $E(f)$ is then given by

$$E(f) = \frac{1}{2} \int_U \sqrt{\det g}\, g^{\alpha\gamma} g^{\beta\lambda} h(P^\perp D_\alpha \partial_\beta f, D_\gamma \partial_\lambda f),$$

where $h = h_{ij}$ is the Riemannian metric on M, and

$$\begin{aligned} g_{\alpha\beta} &= (h \circ f)(\partial_\alpha f, \partial_\beta f), \\ P^\perp &= \mathrm{Id} - g^{\alpha\beta}(h \circ f)(\partial_\alpha f, \cdot)\partial_\beta f, \\ D_\alpha \partial_\beta f &= (0, \partial^2_{\alpha\beta} u) + (\Gamma \circ f)(\partial_\alpha f, \partial_\beta f). \end{aligned}$$

Here $\Gamma = \Gamma^k_{ij}$ are the Christoffel symbols of M. The functional thus has the general form

$$E(f) = \int_U \Big(A^{\alpha\beta\gamma\lambda}(x, u, Du) \partial^2_{\alpha\beta} u\, \partial^2_{\gamma\lambda} u + B^{\alpha\beta}(x, u, Du) \partial^2_{\alpha\beta} u + C(x, u, Du) \Big),$$

where A, B, C are smooth functions, and specifically for $e_3 = (0, 1) \in \mathbb{R}^3$

$$A^{\alpha\beta\gamma\lambda}(x, u, Du) = \frac{1}{2} \sqrt{\det g}\, g^{\alpha\gamma} g^{\beta\lambda}\, h\big(P^\perp e_3, e_3\big)\, \partial^2_{\alpha\beta} u\, \partial^2_{\gamma\lambda} u.$$

We see that a bound for Du implies an ellipticity condition

$$A^{\alpha\beta\gamma\lambda} \xi_{\alpha\beta} \xi_{\gamma\lambda} = \frac{1}{2} \sqrt{\det g}\, \|P^\perp e_3\|^2_h\, \|\xi\|^2_g \geq \lambda |\xi|^2 > 0.$$

It is now straightforward to check that the Euler-Lagrange equation satisfies all the conditions of Lemma 3.2 in [SiL], provided that Du is bounded. Hence we get that u belongs locally to $W^{3,2} \cap C^{2,\alpha}$ for some $\alpha > 0$, and that the L^2 integral of $D^3 u$ satisfies a power decay. As in [SiL] we can refer to [MCB] to conclude that u is in fact smooth. Therefore Theorem 1.1 is proved.

4 Proof of Theorem 1.2

The proof of Theorem 1.2, namely the problem of minimizing the functional

$$W_1(f) = \int_{\mathbb{S}^2} \Big(\frac{1}{4} |H|^2 + 1 \Big)\, d\mu_g$$

in the class of immersions $f : \mathbb{S}^2 \hookrightarrow M$, where M is a closed, three-dimensional Riemannian manifold with sectional curvature $K^M \leq 2$ and moreover $R^M(\overline{x}) > 6$ for some point $\overline{x} \in M$, is very similar to the proof of Theorem 1.1. Here we summarize the different steps of the proof and point out the differences to the proof of Theorem 1.1.

Again we use the concept of minimizing sequences. Therefore let $f_k : \mathbb{S}^2 \hookrightarrow M$ be a minimizing sequence of immersed closed surfaces for the functional W_1 and denote by μ_k the Radon measure on M associated to f_k. Obviously we have that $\mu_k(M) \leq W_1(f_k) \leq C$ uniformly in k. Therefore there exists a Radon measure μ on M such that, up to subsequences,

$$\mu_k \to \mu \quad \text{weakly as Radon measures,} \tag{63}$$

and as before the monotonicity formula Lemma 2.6 yields

$$\operatorname{spt} \mu_k \to \operatorname{spt} \mu \quad \text{in the Hausdorff distance sense.} \tag{64}$$

Observe that, since $R^M(\overline{x}) > 6$ for some point $\overline{x} \in M$, it follows similar to Lemma 2.3 that

$$\inf_{f \in [\mathbb{S}^2, M]} W_1(f) < 4\pi. \tag{65}$$

Using Lemma 2.7, it follows that Proposition 2.5 also holds for E replaced by W_1, which yields that we again have a lower diameter bound, namely

$$\operatorname{diam}_h(\operatorname{spt}\mu) \geq \liminf_k(\operatorname{diam}_h \operatorname{spt}\mu_k) > 0. \tag{66}$$

The next Lemma states an important upper bound for the functional E in terms of the functional W_1 and is a direct consequence of equation (3).

Lemma 4.1. *Let M be a compact Riemannian manifold with sectional curvature $K^M \leq 2$, and let $f : \mathbb{S}^2 \hookrightarrow M$ be a smooth immersion. It follows that*

$$E(f) \leq 2W_1(f) - 4\pi$$

It follows that $\limsup_{k\to\infty} E(f_k) < 4\pi$. Moreover it follows from this uniform upper bound that we can define the bad points with respect to $\varepsilon > 0$ as in Definition 3.1, and that also Remark 3.2, Lemma 3.3, the Graphical Decomposition Lemma 3.4 and the lower density bound in Proposition 3.5 hold in exactly the same way.

Now observe that the proof of the power decay of the L^2-norm of the second fundamental form in Lemma 3.6 carries over analogously up to equation (20) (for the following notation see the proof of Lemma 3.4). Now, since f_k is a minimizing sequence for the functional W_1, we have that

$$W_1(\tilde{f}_k) \geq W_1(f_k) - \varepsilon_k, \quad \text{where } \varepsilon_k \to 0. \tag{67}$$

Equation (3) yields (using that the sectional curvature is bounded by compactness of the manifold M)

$$\sum_{l=1}^{M_k} \int_{\operatorname{graph} w_k^l} |A|^2 \, d\mathcal{H}^2 + c\sum_{l=1}^{M_k} \mathcal{H}^2(\operatorname{graph} w_k^l) \geq \int_{B^e_{\frac{\rho}{16}}(\xi)} |A_k|^2 \, d\mu_k - c\mu_k(B^e_{\frac{\rho}{16}}(\xi)) - \varepsilon_k. \tag{68}$$

Using that $\mathcal{H}^2(\operatorname{graph} w_k^l) \leq c\rho^2$ by the estimates for w_k^l, it follows from Lemma 2.2 and Lemma 2.7 that (22) holds in this setting. The rest of the proof is again the same as before. This shows that also Lemma 3.4 holds.

Now we can construct the limit graph functions as done before after the proof of Lemma 3.6, and show in the same way as before that the limit measure μ is locally (around the good points) given by the sum of the 2-dimensional Hausdorff measure restricted to these limit graphs, namely that Lemma 3.7 holds. Observe that also Proposition 3.8 holds, thus the limit measure is given by $C^{1,\alpha} \cap W^{2,2}$-graphs away from the bad points.

To exclude the bad points, we can do the same as before. Observe that the crucial Lemma 3.12 also holds, because by Lemma 4.1 and (65) the assumption

$$\frac{1}{2}\int |A_k^g|^2 \, d\mu_k^g \leq 4\pi - \delta$$

for some $\delta > 0$ are satisfied. Thus μ is locally given by $C^{1,\alpha} \cap W^{2,2}$-graphs.

As before, using [Breu], it follows that there exists a $C^{1,\alpha} \cap W^{2,2}$-immersion $f : \mathbb{S}^2 \hookrightarrow M$ such that μ is the Radon measure associated to this immersion f. To conclude that f is actually smooth, observe that by construction and lower semicontinuity f satisfies the Euler-Lagrange equation for the functional W_1. By equation (3) the functionals E and W_1 differ only by a topological constant and a multiple of K_f^M, which is a smooth function of u, $\mathrm{D}\,u$ in graph coordinates. Therefore the conditions of Lemma 3.2 in [SiL] are again satisfied. Hence we get that u belongs locally to $W^{3,2} \cap C^{2,\alpha}$ for some $\alpha > 0$, and that the L^2 integral of $D^3 u$ satisfies a power decay. As in [SiL] we can refer to [MCB] to conclude that u is in fact smooth. Therefore also Theorem 1.2 is proved.

5 Appendix

5.1 Some useful Lemmas

In this subsection we state some useful results we need for proving regularity. Lemma 5.1 is an extension result adapted to the cut-and-paste procedure we use and is proved in [Schy].

Lemma 5.1. *Let L be a 2-dimensional plane in $\mathbb{R}^n$, $x_0 \in L$ and $u \in \mathrm{C}^\infty\left(U, L^\perp\right)$, where $U \subset L$ is an open neighborhood of $L \cap \partial B_\rho(x_0)$. Moreover let $|\mathrm{D}\, u| \leq c$ in U. Then there exists a function $w \in \mathrm{C}^\infty(\overline{B_\rho(x_0)}, L^\perp)$ with the following properties:*

$$1.)\quad w = u \quad \text{and} \quad \frac{\partial w}{\partial \nu} = \frac{\partial u}{\partial \nu} \quad \text{on } \partial B_\rho(x_0),$$

$$2.)\quad \frac{1}{\rho}||w||_{L^\infty(B_\rho(x_0))} \leq c(n)\left(\frac{1}{\rho}||u||_{L^\infty(\partial B_\rho(x_0))} + ||\mathrm{D}\, u||_{L^\infty(\partial B_\rho(x_0))}\right),$$

$$3.)\quad ||\mathrm{D}\, w||_{L^\infty(B_\rho(x_0))} \leq c(n)||\mathrm{D}\, u||_{L^\infty(\partial B_\rho(x_0))},$$

$$4.)\quad \int_{B_\rho(x_0)} |\mathrm{D}^2\, w(x)|^2 \,\mathrm{d}\, x \leq c(n)\rho \int_{\mathrm{graph}\, u_{|\partial B_\rho(x_0)}} |A(x)|^2\, d\mathcal{H}^1,$$

where $d\mathcal{H}^1$ is the 1-dimensional Euclidean Hausdorff measure.

The second lemma is a useful selection principle proved in [SiL].

Lemma 5.2. *Let $\delta > 0, I \subset \mathbb{R}$ a bounded interval and $A_k \subset I, k \in \mathbb{N}$, measurable sets with $\mathcal{L}^1(A_k) \geq \delta$ for all k. Then there exists a set $A \subset I$ with $\mathcal{L}^1(A) \geq \delta$, such that each point $x \in A$ lies in A_k for infinitely many k.*

The third lemma is a decay result we need to get the power decay for the L^2-norm of the second fundamental form in Lemma 3.6.

Lemma 5.3. *Let $g : (0, b) \to [0, +\infty)$ be a bounded function such that*

$$g\,(x) \leq \gamma g(2x) + Cx^\alpha \quad \text{for all } x \in \left(0, \frac{b}{2}\right),$$

where $\alpha > 0$, $\gamma \in (0, 1)$, and $C \geq 0$ is a constant. Then there exists a $\beta \in (0, 1)$ and a constant $C = C\left(\gamma, \alpha, b, ||g||_{L^\infty(0,b)}\right)$ such that

$$g(x) \leq Cx^\beta \quad \text{for all } x \in (0, b)\,.$$

The last statement is a generalized Poincaré inequality proved in [SiL].

Lemma 5.4. *Let $\mu > 0$, $\delta \in \left(0, \frac{\mu}{2}\right)$ and $\Omega = B_\mu^{\mathbb{R}^2}(0) \backslash E$, where $E \subset \mathbb{R}^2$ is measurable with $\mathcal{L}^1(p_1(E)) \leq \frac{\mu}{2}$ and $\mathcal{L}^1(p_2(E)) \leq \delta$, where p_1 is the projection onto the x-axis and p_2 is the projection onto the y-axis. Then for any $f \in C^1(\Omega)$ there exists a point $(x_0, y_0) \in \Omega$ such that*

$$\int_\Omega |f - f(x_0, y_0)|^2 \leq C\mu^2 \int_\Omega |\mathrm{D}\, f|^2 + C\delta\mu \sup_\Omega |f|^2$$

where C is an absolute constant.

5.2 Definitions and properties of generalized (r,λ)-immersions

Here we recall the definitions and properties of generalized (r,λ)-immersions $f:\mathbb{S}^2 \hookrightarrow M \subset \mathbb{R}^p$ appearing in [Breu].
We call a mapping $A:\mathbb{R}^p \to \mathbb{R}^p$ an Euclidean isometry, if there is a rotation $R \in SO(p)$ and a translation $T \in \mathbb{R}^p$, such that $A(x) = Rx + T$ for all $x \in \mathbb{R}^p$.
For a given point $q \in \mathbb{S}^2$ and a given 2-plane $E \in G(p,2)$ let $A_{q,E}:\mathbb{R}^p \to \mathbb{R}^p$ be an Euclidean isometry which maps the origin to $f(q)$ and the subspace $\mathbb{R}^2 \times \{0\} \subset \mathbb{R}^p$ onto $f(q) + E$.
Let $U^E_{r,q} \subset \mathbb{S}^2$ be the q-component of the set $(\pi \circ A^{-1}_{q,E} \circ f)^{-1}(B_r)$, where $\pi:\mathbb{R}^p \to \mathbb{R}^2$ is the projection on the first two coordinates.

Definition 5.5. *An immersion $f:\mathbb{S}^2 \hookrightarrow M \subset \mathbb{R}^p$ is called a generalized (r,λ)-immersion, if for each point $q \in \mathbb{S}^2$ there is an $E = E(q) \in G(p,2)$, such that $A^{-1}_{q,E} \circ f(U^E_{r,q})$ is the graph of a differentiable function $u: B_r \to (\mathbb{R}^2)^\perp$ with $\|\mathrm{D}\,u\|_{C^0(B_r)} \le \lambda$.*
The set of generalized (r,λ)-immersions is denoted by $\mathcal{F}^1(r,\lambda)$. Moreover let $\mathcal{F}^1_V(r,\lambda)$ be the set of all immersions $f \in \mathcal{F}^1(r,\lambda)$ such that $\mu_g(\mathbb{S}^2) \le V$, where μ_g is the induced area measure.
A continuous function $f:\mathbb{S}^2 \hookrightarrow M \subset \mathbb{R}^p$ is called a (r,λ)-function, if for each point $q \in \mathbb{S}^2$ there is an $E = E(q) \in G(p,2)$, such that $A^{-1}_{q,E} \circ f(U^E_{r,q})$ is the graph of a Lipschitz function $u: B_r \to (\mathbb{R}^2)^\perp$ with with Lipschitz constant λ. The set of (r,λ)-functions is denoted by $\mathcal{F}^0(r,\lambda)$.

Now we recall the Compactness Theorem in [Breu], Theorem 0.5.

Theorem 5.6. *Let $\lambda \le \frac{1}{4}$. Then $\mathcal{F}^1_V(r,\lambda)$ is relatively compact in $\mathcal{F}^0(r,\lambda)$ in the following sense: Let $f_k:\mathbb{S}^2 \hookrightarrow M \subset \mathbb{R}^p$ be a sequence in $\mathcal{F}^1_V(r,\lambda)$. Then, after passing to a subsequence, there exists a function $f \in \mathcal{F}^0(r,\lambda)$ and a sequence of diffeomorphisms $\phi_k:\mathbb{S}^2 \to \mathbb{S}^2$, such that $f_k \circ \phi_k$ is uniformly Lipschitz bounded and converges uniformly to f.*

References

[Al] W. K. Allard, *On the first variation of a varifold,* Annals of Math. Vol. 95, (1972), 417–491.

[Breu] P. Breuning, *Immersions with local Lipschitz representation,* Ph. D. Thesis, Freiburg, (2011).

[Dan] B. Daniel, *Isometric immersions into 3-dimensional homogeneous manifolds,* Comment. Math. Helv., Vol. 82 (2007), 87–131

[GT] Gilbarg, D. and Trudinger, N.S., *Elliptic partial differential equations of second order,* Springer (2001)

[Hu1] J. E. Hutchinson, *Second fundamental form for varifolds and the existence of surfaces minimizing curvature,* Indiana Math. Journ., Vol. 35, Num. 1, (1986), 45–71.

[KS] E. Kuwert, R. Schätzle, *Removability of isolated singularities of Willmore surfaces,* Annals of Math., Vol. 160, Num. 1, (2004), 315–357.

[LM] T. Lamm, J. Metzger *Small surfaces of Willmore type in Riemannian manifolds*, Int. Math. Res. Not. IMRN. 19 (2010), 3786–3813.

[LMS] T. Lamm, J. Metzger, F. Schulze *Foliations of asymptotically flat manifolds by surfaces of Willmore type*, Math. Ann. 350 (2011), 1–78.

[Lan] J. Langer, *A compactness theorem for surfaces with L^p-bounded second fundamental form,* Math. Ann., Vol. 270, (1985), 223–234.

[LY] P. Li, S. T. Yau, *A new conformal invariant and its applications to the Willmore conjecture and the first eigenvalue on compact surfaces,* Invent. Math., Vol. 69, (1982), 269–291 .

[Mon1] A. Mondino, *Some results about the existence of critical points for the Willmore functional,* Math. Zeit., Vol. 266, Num. 3, (2010), 583–622.

[Mon2] A. Mondino, *The conformal Willmore Functional: a perturbative approach,* Journal of Geometric Analysis (Online First), (2011), 1–48.

[MonVar] A. Mondino, *Existence of integral m-varifolds minimizing $\int |A|^p$ and $\int |H|^p$, $p > m$, in Riemannian manifolds,* arXiv:1010.4514, submitted, (2010).

[MCB] C. B. Morrey, *Multiple integrals in the calculus of variations*, Springer Verlag (1966).

[Riv] T. Rivière, *Variational principles for immersed surfaces with L^2-bounded second fundamental form,* arXiv:1007.2997 (2010)

[Schy] J. Schygulla, *Willmore minimizers with prescribed isoperimetric ratio,* to appear in Archiv. Rational Mech. Anal. (2011).

[SiProc] L. Simon, *Existence of Willmore surfaces,* Miniconf. on Geom. and P.D.E. (Canberra, 1985), Proc. Centre Math. Anal., Vol. 10, Australian Nat. Univ., Canberra, (1986), 187–216.

[SiL] L. Simon, *Existence of surfaces minimizing the Willmore functional,* Comm. Anal. Geom., Vol. 1, Num. 2, (1993), 281–325.

[SiGMT] L. Simon, *Lectures on geometric measure theory,* Proc. Centre for Math. Analysis Australian National University, Vol.3, Canberra, Australia (1983).

[ST] R. Souam, E. Toubiana *Totally umbilic surfaces in homogeneous 3-manifolds,* Comment. Math. Helv., Vol 84, (2009), 673–704.

[Will] T.J. Willmore, *Riemannian geometry,* Oxford Science Publications, Oxford University Press (1993).

Ernst Kuwert, Mathematisches Institut, Universität Freiburg, Eckerstraße 1, 79104 Freiburg, Germany
EMAIL: ernst.kuwert@math.uni-freiburg.de

Andrea Mondino, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
EMAIL: andrea.mondino@sns.it

Johannes Schygulla, Mathematisches Institut, Universität Freiburg, Eckerstraße 1, 79104 Freiburg, Germany
EMAIL: johannes.schygulla@math.uni-freiburg.de