Quantitative Stratification and the Regularity of Harmonic Map Flow

Jeff Cheeger, Robert Haslhofer and Aaron Naber[*]

Abstract

In this paper, we prove estimates and quantitative regularity results for the harmonic map flow. First, we consider H^1_{loc}-maps u defined on a parabolic ball $P \subset M^m \times \mathbb{R}$ and with target manifold N, that have bounded Dirichlet-energy and Struwe-energy. We define a *quantitative* stratification, which groups together points in the domain into quantitative weakly singular strata $\mathcal{S}^j_{\eta,r}(u)$ according to the number of approximate symmetries of u at certain scales, and prove that their tubular neighborhoods have small volume, namely $\text{Vol}\left(T_r(\mathcal{S}^j_{\eta,r}(u))\right) \leq Cr^{m+2-j-\varepsilon}$. In particular, this generalizes the known Hausdorff estimate $\dim \mathcal{S}^j(u) \leq j$ for the weakly singular strata of suitable weak solutions of the harmonic map flow. As an application, specializing to Chen-Struwe solutions with target manifolds that do not admit certain harmonic and quasi-harmonic spheres, we obtain refined Minkowski estimates for the singular set. This generalizes a result of Lin-Wang [LW99]. We also obtain L^p-estimates for the reciprocal of the regularity scale. The results are analogous to our results for mean curvature flow that we recently proved in [CHN13].

1   Introduction and main results

In this paper, we prove estimates and quantitative regularity results for the harmonic map flow. The results are analogous to our results for mean curvature flow that we recently proved in [CHN13].

Recalling the standard setting for the harmonic map flow [ES64], let M^m and N^n be closed Riemannian manifolds and $u_0 : M \to N$ be a smooth map; we can assume $N \subset \mathbb{R}^d$, and we write A for the second fundamental form. Chen-Struwe [CS89] proved the existence of weak solutions $u : M \times \mathbb{R}_+ \to N$ of the harmonic map flow starting at u_0,

$$\partial_t u = \Delta u + A(u)(\nabla u, \nabla u), \qquad u|_{t=0} = u_0. \qquad (1.1)$$

Having a general existence theory, a main question is then to study the regularity of these weak solutions.

Adapting Struwe's monotonicity formula [Str88] to the setting of their existence proof, Chen-Struwe proved that the parabolic Hausdorff dimension of the singular set $S \subset M \times \mathbb{R}_+$ is at most m, c.f. [CS89, Thm. 1.5]. We recall that the parabolic Hausdorff dimension is the Hausdorff dimension with respect to the space-time metric $d((x, t), (y, s)) = \max\{d_M(x, y), \sqrt{|t - s|}\}$; in particular, note that $\dim(M \times \mathbb{R}_+) = m + 2$. About ten years later, based on refined blowup-analysis, Lin-Wang proved that if the target manifold N doesn't admit certain harmonic and quasi-harmonic spheres, then the Hausdorff dimension of the singular set must be smaller [LW99]; this in turn was based on the sophisticated blow-up analysis of Lin in the elliptic setting [Lin99].

The goal of the present paper is to make these regularity results more quantitative. The main new ingredient that we develop and apply is the quantitative stratification technique [CN13a, CN13b, CHN13]; this allows us to turn infinitesimal statements from blowup analysis into more quantitative ones.

1.1 Quantitative stratification for general targets

Our first main result, Theorem 1.11, holds without any assumptions on the target manifold, and in fact can be formulated in a very general setting. It applies in particular to Chen-Struwe solutions, but it holds for any H^1_{loc}-map defined on a parabolic ball and with target N that satisfies the local energy bounds (1.2) and (1.3) below. In principle, we do not even need to assume that the equation (1.1) is satisfied weakly. However in practive the estimates (1.2) and (1.3) arise because u is a weak solution. This degree of generality is important for two reasons. To begin with there is more than one notion of a *weak* solution to (1.1). In reasonable situations, these notions should agree, but as of this point, this has not always been proved. Additionally, one may arrive at the estimates (1.2) and (1.3) in contexts where (1.1) only holds up to a bounded lower order term, allowing the estimates to apply to a much broader class of situations.

To describe this setting more precisely, let R be any sufficiently small radius, say less than one-quarter of the injectivity radius of M. We then consider H^1_{loc}-maps u defined on a parabolic ball $P_{4R} = B_{4R} \times (-(4R)^2, (4R)^2) \subset M \times \mathbb{R}$, and with target N. As usual when studying interior regularity, we will derive estimates on a somewhat smaller parabolic ball, say on P_R. Our estimates depend on a bound for the *scale invariant Dirichlet-energy*,

$$\sup_{X_0 \in P_{2R}} \sup_{r \leq 2R} \frac{1}{r^m} \int_{P_r(X_0)} |\nabla u|^2 \, dV dt \leq \Lambda_1, \tag{1.2}$$

and a bound for what we call – alluding to Struwe's monotonicity formula [Str88] – the *Struwe-energy*,

$$\sup_{X_0 \in P_{2R}} \int_{P_{2R}^-(X_0)} |(x - x_0) \cdot \nabla u + 2(t - t_0)\partial_t u|^2 \, e^{-\frac{|x - x_0|^2}{4|t - t_0|}} |t - t_0|^{-(m+2)/2} \, dV dt \leq \Lambda_2. \tag{1.3}$$

2

Here, we used the notation $X = (x, t)$ for points in space-time, and $P_r^-(X) = B_r(x) \times (t - r^2, t]$ for backwards parabolic balls. The integral in (1.3) is computed in geodesic coordinates. It follows from [CS89, Lemma 4.1, 4.2] and lower semicontinuity that Chen-Struwe solutions satisfy (1.2) and (1.3) for some constants $\Lambda_1, \Lambda_2 < \infty$ depending only on M, N and energy of the initial map u_0. To keep track of the setting we just described, we write $H^1_{\Lambda_1, \Lambda_2}(P_{4R}, N)$ for the space of all H^1_{loc}-maps with target N, defined on a sufficiently small parabolic ball P_{4R}, and with energy bounded as in (1.2) and (1.3).

Remark 1.4. In the literature, there is the notion of *suitable* weak solutions, i.e. weak solutions satisfying a parabolic stationarity condition, c.f. [Fel94, CLL95]. However, it is not clear whether or not the weak solutions constructed by Chen-Struwe actually satisfy this condition. This is one of the reasons, why we have chosen a more general setting merely requiring (1.2) and (1.3).

Let $u \in H^1_{\Lambda_1, \Lambda_2}(P_{4R}, N)$ be a H^1 function with bounded scale invariant Dirichlet energy and Struwe-energy. In the quantitative stratification we will group the points in P_R into different strata according to their number of approximate symmetries at bounded scales. To motivate the definition let us first discuss the nonquantitative notion of weak tangents: Given $X_0 \in P_R$ and $s < R$ we consider the rescaled map $u_{X_0, s} : P_1(0) \subset \mathbb{R}^m \times \mathbb{R} \to N$, $u_{X_0, s}(x, t) = u(\exp_{x_0}(sx), t_0 + s^2 t)$. Note first that by combining (1.2) and (1.3) we obtain scale invariant bounds

$$\sup_{X_0 \in P_R} \sup_{r \leq R} \frac{1}{r^{m-2}} \int_{P_r(X_0)} |\partial_t u|^2 \, dV dt \leq C(\Lambda_1, \Lambda_2). \tag{1.5}$$

By (1.2) and (1.5) for every *blowup sequence* u_{X_0, s_α}, $s_\alpha \to 0$ there exists a subsequence such that $u_{X_0, s_\alpha} \rightharpoonup \varphi$ weakly in H^1_{loc}. Such a limit φ is called a *weak tangent*, and by (1.3) every weak tangent is *backwardly selfsimilar*, i.e. $\varphi(x, t) = \varphi(\lambda x, \lambda^2 t)$ for all $x \in \mathbb{R}^m, t < 0$ and all $\lambda > 0$. It is important to keep track of the number of symmetries of φ. The number of spatial symmetries is the maximal d for which there exists a d-plane $V \subset \mathbb{R}^m$ such that $\varphi(x, t) = \varphi(x + v, t)$ for all $x \in \mathbb{R}^m, t < 0$ and all $v \in V$. With respect to the behavior in the time direction, we have to distinguish between the following three cases. *(i) static:* φ is independent of t for all $t \in \mathbb{R}$; *(ii) quasi-static:* φ is independent of t up to some time $T \in [0, \infty)$ but not for all t; *(iii) shrinking:* φ is not independent of time on $(-\infty, 0]$. Note that in the shrinking case we have $\varphi(x, t) = \psi(x/\sqrt{-t})$ for some function ψ that is not radially invariant. We then consider the number of space-time symmetries

$$D(\varphi) = \begin{cases} d(\varphi) & \text{if } \varphi \text{ is shrinking or quasi-static} \\ d(\varphi) + 2 & \text{if } \varphi \text{ is static} \end{cases} \tag{1.6}$$

We say that φ is *j-selfsimilar* if it is backwardly selfsimilar and $D(\varphi) \geq j$.

Remark 1.7. Consider a time-dependent map that is equal to some given stationary harmonic map up to some time T, and constant for later times. This map satisfies (1.2) and (1.3), and the blowups at the truncation time T are quasi-static. However, it seems to be unknown if quasi-static blowups

can actually occur in any "interesting" situation, e.g. for Chen-Struwe solutions with smooth initial data; it is known that they cannot occur if the target doesn't admit harmonic two-spheres.[1]

One could then proceed by grouping points $X \in P_R$ into weakly singular strata $\mathcal{S}^j(u)$ according to the number of symmetries of the weak tangents. Namely, $X \in \mathcal{S}^j(u)$ if no weak tangent φ at X has $D(\varphi) > j$. However, as mentioned above, we instead group points together into quantitative weakly singular strata $\mathcal{S}^j_{\eta,r}(u)$ according to the number of approximate symmetries at certain scales:

Definition 1.8. For each $\eta > 0$ and $0 < r < R$, we define the j-th quantitative weakly singular stratum

$$\mathcal{S}^j_{\eta,r}(u) := \left\{ X \in P_R : \int_{P_1(0)} \left| u_{X,s} - \varphi \right|^2 > \eta \text{ for all } s \in [r,R] \text{ and all } (j+1)\text{-selfsimilar } \varphi \right\}. \quad (1.9)$$

Remark 1.10. In particular, this recovers the standard stratification via $\mathcal{S}^j(u) = \bigcup_\eta \bigcap_r \mathcal{S}^j_{\eta,r}(u)$.

We write Vol for the $m + 2$ dimensional (parabolic) Hausdorff measure on space-time, and T_r for r-tubular neighborhoods with respect to the parabolic metric. Our main quantitative stratification theorem gives estimates for the volume of tubular neighborhoods of the quantitative weakly singular strata:

Theorem 1.11. *If $u : P_{4R} \subset M \times \mathbb{R} \to N$ is an H^1_{loc}-map with energy bounded as in (1.2) and (1.3), then the j-th quantitative weakly singular stratum $\mathcal{S}^j_{\eta,r}(u)$ satisfies the volume estimate*

$$\text{Vol}\left(T_r(\mathcal{S}^j_{\eta,r}(u)) \right) \leq C r^{m+2-j-\varepsilon} \qquad (0 < r < R), \qquad (1.12)$$

for some constant $C < \infty$ depending only on $\varepsilon, \eta, \Lambda_1, \Lambda_2$ and the geometry of $B_{4R} \subset M$ and N.

Remark 1.13. In particular, this generalizes the known Hausdorff dimension estimate $\dim \mathcal{S}^j(u) \leq j$ for suitable weak solutions of the harmonic map flow in various ways. First, it replaces the weakly singular strata $\mathcal{S}^j(u)$ by the more effective quantitative weakly singular strata $\mathcal{S}^j_{\eta,r}(u)$. Second, it shows that tubular neighborhoods have small volume, i.e. it improves the Hausdorff estimate to a Minkowski estimate. Third, it applies to a much larger class of maps.

1.2 Higher regulartiy for certain targets

Following Lin [Lin99] and Lin-Wang [LW99], we consider target manifolds N that do not admit harmonic two-spheres, i.e. no nonconstant smooth harmonic maps $S^2 \to N$. By [LW99, Prop. 4.1] this will allow us to upgrade weak convergence in H^1_{loc} to strong convergence (see also the discussion on page 788 of [Lin99]).

---

[1]The situation is similar for the mean curvature flow: There, quasi-static planes occur when one truncates a Brakke flow at some time. Excluding this "trivial" example, it seems unknown if the quasi-static planes can actually occur as blowups of solutions with smooth embedded initial data; it is known that they cannot occur in the mean convex case.

Remark 1.14. By Sacks-Uhlenbeck [SU81, Thm. 5.7] the assumption that the target does not admit harmonic two-spheres implies that its universal cover must be contractible. See Section 4 for more on this.

To state our higher regularity results, we introduce the following definitions.

Definition 1.15 (regularity scale)**.** We define the regularity scale of a harmonic map flow u : $M \times I \to N$ at a point $X = (x, t)$ by

$$r_u(X) := \sup \left\{ r \geq 0 : \sup_{P_r(X)} r|\nabla u| + r^2|\nabla^2 u| \leq 1 \right\}. \tag{1.16}$$

Remark 1.17. The regularity scale controls all other geoemetric quantities. Indeed, standard interior estimates give scale invariant bounds for all (spatial and time) derivatives on a somewhat smaller parabolic ball, say on $P_{r_u(X)/2}(X)$.

Definition 1.18 (r-bad set)**.** Given $u : M \times I \to N$ and $r > 0$ we define the r-bad set

$$\mathcal{B}_r(u) := \{X \in M \times I : r_u(X) \leq r\}. \tag{1.19}$$

Our second main theorem shows that the bad set is quite small, provided the target doesn't admit harmonic two-spheres, and even smaller provided the target doesn't admit certain harmonic and quasi-harmonic spheres. We recall that a quasi-harmonic ℓ-sphere is a nonconstant smooth map $\psi : \mathbb{R}^\ell \to N$ that is a critical point of the functional $\int_{\mathbb{R}^\ell} |D\psi|^2 \, e^{-x^2/4}$. Such maps ψ correspond to selfsimilarly shrinking solutions $u(x, t) = \psi(x/\sqrt{-t})$ of the harmonic map flow.

Theorem 1.20. *Let M^m, N^n be closed manifolds and $u : M \times \mathbb{R}_+ \to N$ a Chen-Struwe solution of the harmonic map flow (1.1). If N does not admit harmonic two-spheres, then the bad set (c.f. Definition 1.18) satisfies the volume estimate*

$$\mathrm{Vol}\, (T_r(\mathcal{B}_r(u)) \cap \{t < T\}) \leq Cr^{4-\varepsilon} \qquad (0 < r < 1), \tag{1.21}$$

for some constant $C = C(M, N, u_0, T, \varepsilon) < \infty$. In particular, the parabolic Minkowski dimension of the singular set $\mathcal{S} \subset M \times \mathbb{R}_+$ is at most $m - 2$. More generally, if N doesn't admit harmonic ℓ-spheres for $\ell = 2, \ldots, k$ and quasi-harmonic ℓ-spheres for $\ell = 3, \ldots, k - 1$, then

$$\mathrm{Vol}\, (T_r(\mathcal{B}_r(u)) \cap \{t < T\}) \leq Cr^{k+2-\varepsilon} \qquad (0 < r < 1), \tag{1.22}$$

and the parabolic Minkowski dimension of the singular set is at most $m - k$.

Remark 1.23. Theorem 1.20 is a quantitative version of [LW99, Thm. 4.3]. In particular, it improves the Hausdorff estimate to a Minkowski estimate.

As an immediate consequence of Theorem 1.20, we obtain L^p-estimates for the reciprocal of the regularity scale. In particular, this gives integral estimates for the derivatives of u.

Corollary 1.24. *Let M^m, N^n be closed manifolds and $u : M \times \mathbb{R}_+ \to N$ a Chen-Struwe solution of the harmonic map flow (1.1). If N doesn't admit harmonic ℓ-spheres for $\ell = 2, \ldots, k$ and quasi-harmonic ℓ-spheres for $\ell = 3, \ldots, k-1$, then for every $p < k+2$ we have*

$$\int_0^T \int_M r_u^{-p} dV dt \le C,$$

for some $C = C(M, N, u_0, T, p) < \infty$. In particular, there are estimates for the derivatives of u,

$$\int_0^T \int_M |\nabla^j u|^{\frac{p}{j}} dV dt \le C_j, \tag{1.25}$$

for some constants $C_j = C_j(M, N, u_0, T, p) < \infty$ ($j = 1, 2 \ldots$).

We end this section with the following particularly important corollary. In [CN13b], the first general higher derivative estimates for *minimizing* harmonic maps were proved. By using Theorem 1.20 we can prove the same estimates for *stationary* harmonic maps when the target space does not admit any harmonic 2-spheres. Specifically:

Corollary 1.26. *Let M, N be closed manifolds and $u : M \to N$ a stationary harmonic map with $\int_M |\nabla u|^2 < \Lambda$. If N does not admit any harmonic ℓ-spheres for $\ell = 2, \ldots, k$, then for every $p < k+2$ we have*

$$\int_M reg_u^{-p} dV \le C,$$

for some $C = C(M, N, \Lambda, p) < \infty$, where $reg_u(x) \equiv \sup_{r \le 1}\{\sup_{B_r(x)} r|\nabla u| + r^2|\nabla^2 u| \le 1\}$. In particular, there are estimates for the higher derivatives of u,

$$\int_M |\nabla^j u|^{\frac{p}{j}} dV \le C_j, \tag{1.27}$$

for some constants $C_j = C_j(M, N, \Lambda, p) < \infty$ ($j = 1, 2 \ldots$).

Organization of the paper. In Section 2, we prove Theorem 1.11. In Section 3, we prove an ε-regularity lemma, and use this in combination with Theorem 1.11 to prove Theorem 1.20. In Section 4, we provide some examples of compact manifolds which do not admit harmonic 2-spheres. In particular, we provide examples which do not have nonpositive sectional curvature.

2 Volume estimates for the quantitative strata

In this section, we prove Theorem 1.11 following closely our argument for Brakke flows in [CHN13]. We first prove a quantitative rigidity lemma and decompose P_R into a union of sets, according to

the behavior of points at different scales. By virtue of a quantitative differentiation argument, we show that the number of sets in this decomposition grows at most polynomially. We then establish a cone-splitting lemma for weak tangents and prove, roughly speaking, that at their good scales points in $\mathcal{S}^j_{\eta,r}(u)$ line up along at most j-dimensional subspaces. Finally, we conclude the argument by constructing a suitable covering of $\mathcal{S}^j_{\eta,r}(u)$ and computing its volume. For ease of notation, in the following we will pretend $R = 1$ and $P_4 = P_4(0) \subset \mathbb{R}^m \times \mathbb{R}$; the general case works similarly.

2.1 Energy decomposition

The goal of this subsection is to decompose P_1 into a union of sets E_{T^β}, according to the behavior of points at different scales. As in [CN13a, CN13b, CHN13] it will be of crucial importance that we can deal separately with each individual set E_{T^β}, all of whose points have the same $\{0, 1\}$-valued β-tuple T^β of good and bad scales.

Definition 2.1. A map $u \in H^1_{\Lambda_1,\Lambda_2}(P_{4r}(X), N)$ is (ε, r, j)-*selfsimilar* at $X = (x, t)$ if there exists a j-selfsimilar map φ such that
$$\int_{P_1(0)} |u_{X,r} - \varphi|^2 < \varepsilon$$
If φ is shrinking with respect to a plane V^j, we put $W_X = (x + V) \times \{t\}$. If φ is quasi-static with respect to V^j up to time T, we put $W_X = (x + V) \times (-\infty, t + r^2 T]$. If φ is static with respect to V^{j-2}, we put $W_X = (x + V) \times \mathbb{R}$. We say that u is (ε, r, j)-*selfsimilar at X with respect to W_X*.

Definition 2.2. Given $u \in H^1_{\Lambda_1,\Lambda_2}(P_4, N)$, for $X_0 \in P_1$ and $1/2 > r_1 > r_2$, we define the (r_1, r_2)-*Struwe energy* by

$$\mathcal{W}_{r_1,r_2}(u, X_0) := \int_{P^-_{r_1}(X_0) \backslash P^-_{r_2}(X_0)} |(x - x_0) \cdot \nabla u + 2(t - t_0)\partial_t u|^2 \, e^{-\frac{|x - x_0|^2}{4|t - t_0|}} \, |t - t_0|^{-(m+2)/2} \, dVdt. \qquad (2.3)$$

Lemma 2.4 (Quantitative Rigidity)**.** *For all $\varepsilon > 0$, $\Lambda_1, \Lambda_2 < \infty$, m and N there exists $\delta = \delta(\varepsilon, \Lambda_1, \Lambda_2, m, N) > 0$, such that if $u \in H^1_{\Lambda_1,\Lambda_2}(P_4, N)$ satisfies*

$$\mathcal{W}_{r,\delta r}(u, X) \leq \delta, \qquad (2.5)$$

for some $X \in P_1$ and some $r \in (0, 1/2)$, then u is $(\varepsilon, r, 0)$-selfsimilar at X.

Proof. If not, then there exist a sequence of maps $u_\alpha \in H^1(P_4, N)$ with

$$\int_{P^-_1 \backslash P^-_{1/\alpha}} |x \cdot \nabla u_\alpha + 2t\partial_t u_\alpha|^2 \, e^{-\frac{|x|^2}{4|t|}} \, |t|^{-(m+2)/2} \, dVdt \leq 1/\alpha, \qquad (2.6)$$

and scale invariant bounds as in (1.2) and (1.5), but such that

$$\int_{P_1} |u_\alpha - \varphi|^2 \geq \varepsilon \qquad (2.7)$$

7

for all 0-selfsimilar φ. However, it follows from (1.2), (1.5) and (2.6) that, after passing to a subsequence, $u_\alpha \rightharpoonup \varphi$ weakly in H^1 for some 0-selfsimilar φ. Since the convergence is strong in L^2, for α large enough we obtain a contradiction with (2.7). □

Let $u \in H^1_{\Lambda_1,\Lambda_2}(P_4, N)$ and $X \in P_1$. Given constants $0 < \gamma < 1/2$ and $\delta > 0$ and an integer $q < \infty$ (these parameters will be fixed suitably in Section 2.3), let K be the number of $\alpha > q$ such that

$$\mathcal{W}_{\gamma^{\alpha-q},\gamma^{\alpha+q}}(u, X) > \delta.$$

Since the Struwe-energy is bounded by Λ_2 it follows that

$$K \leq (2q + 1)\delta^{-1}\Lambda_2. \tag{2.8}$$

Otherwise, there would be at least $\delta^{-1}\Lambda_2$ disjoint intervalls of the form $(\gamma^{\alpha-q}, \gamma^{\alpha+q})$ with $\mathcal{W}_{\gamma^{\alpha-q},\gamma^{\alpha+q}}(u, X) > \delta$. This is an instance of *quantitative differentiation* (see [Che12] for a general perspective).

For each point $X \in P_1$, to keep track of its behavior at different scales, we define a $\{0, 1\}$-valued sequence $(T_\alpha(X))_{\alpha \geq 1}$ as follows. By definition, $T_\alpha(X) = 1$ if $\alpha \leq q$ or $\mathcal{W}_{\gamma^{\alpha-q},\gamma^{\alpha+q}}(u, X) > \delta$, and $T_\alpha(X) = 0$ if $\alpha > q$ and $\mathcal{W}_{\gamma^{\alpha-q},\gamma^{\alpha+q}}(u, X) \leq \delta$. Then, for each β-tuple $(T^\beta_\alpha)_{1 \leq \alpha \leq \beta}$, we put

$$E_{T^\beta}(u) = \{X \in P_1 \mid T_\alpha(X) = T^\beta_\alpha \text{ for } 1 \leq \alpha \leq \beta\}. \tag{2.9}$$

A priori there are 2^β possible sets $E_{T^\beta}(u)$. However, by the above, $E_{T^\beta}(u)$ is empty whenever T^β has more than

$$Q := (2q + 1)\delta^{-1}\Lambda_2 + q$$

nonzero entries. Thus, we have constructed a decomposition of P_1 into at most β^Q nonempty sets $E_{T^\beta}(u)$.

2.2 Cone-splitting

The goal of this subsection is to prove Corollary 2.12 which says, roughly speaking, that at their good scales points line up in a tubular neighborhood of a well defined *almost planar* set. Here, the set of points that we call δ-*good at scales between Ar and r/A $(A > 1)$* is defined as

$$L_{Ar,r/A,\delta}(u) = \{X \in P_1 : \mathcal{W}_{Ar,A^{-1}r}(u, X) \leq \delta\}. \tag{2.10}$$

A key role is played by the following cone-splitting principle and its quantitative version (Lemma 2.11). Similar ideas played a key role in [CN13a, CN13b, CHN13].

Cone-splitting principle. Assume that φ is j-selfsimilar at 0 with respect to W and 0-selfsimilar at $Y = (y, s) \notin W$. Then we have the following implications:

- If $W = V^j \times \{0\}$ and

- if $s = 0$, then $y \notin V$ and φ is $(j+1)$-selfsimilar at 0 with respect to $\mathrm{span}\{y, V^j\} \times \{0\}$.

- if $s \neq 0$ and $y \in V$, then φ is j-selfsimilar at 0 and quasistatic with respect to $V^j \times (-\infty, \max\{s, 0\}]$.

- if $s \neq 0$ and $y \notin V$, then φ is $(j+1)$-selfsimilar and quasistatic with respect to $\mathrm{span}\{y, V^j\} \times (-\infty, \max\{s, 0\}]$.

- If $W = V^j \times (-\infty, T]$ and

  - if $y \in V$, then $s > T$ and φ is j-selfsimilar at 0 and quasistatic with respect to $V^j \times (-\infty, s]$.

  - $y \notin V$, then φ is $(j+1)$-selfsimilar and quasistatic with respect to $\mathrm{span}\{y, V^j\} \times (-\infty, \max\{s, T\}]$.

- If $W = V^{j-2} \times \mathbb{R}$, then $y \notin V$ and φ is $(j+1)$-selfsimilar and static with respect to $\mathrm{span}\{y, V^{j-2}\} \times \mathbb{R}$.

From an argument by contradiction, we immediately obtain the following quantitative refinement.

Lemma 2.11 (Cone-splitting lemma). *For all $\varepsilon, \rho > 0$, $R, \Lambda_1, \Lambda_2 < \infty, m$ and N there exists a constant $\delta = \delta(\varepsilon, \rho, R, \Lambda_1, \Lambda_2, m, N) > 0$ with the following property. If $u \in H^1_{\Lambda_1, \Lambda_2}(P_{4R}, N)$ satisfies*

1. *u is (δ, R, j)-selfsimilar at 0 with respect to W.*

2. *There exists $Y = (y, s) \in P_1 \setminus T_\rho(W)$ such that u is $(\delta, 2, 0)$-selfsimilar at Y,*

then we have the following implications:

- *If $W = V^j \times \{0\}$ and*

  - *if $|s| < \rho^2$, then $d(y, V) \geq \rho$ and u is $(\varepsilon, 1, j+1)$-selfsimilar at 0 with respect to $\mathrm{span}\{y, V^j\} \times \{0\}$.*

  - *if $|s| \geq \rho^2$ and $d(y, V) < \rho$, then u is $(\varepsilon, 1, j)$-selfsimilar at 0 with respect to $V^j \times (-\infty, \max\{s, 0\}]$.*

  - *if $|s| \geq \rho^2$ and $d(y, V) \geq \rho$, then u is $(\varepsilon, 1, j+1)$-selfsimilar at 0 with respect to $\mathrm{span}\{y, V^j\} \times (-\infty, \max\{s, 0\}]$.*

- *If $W = V^j \times (-\infty, T]$ and*

  - *if $d(y, V) < \rho$, then $s \geq T + \rho^2$ and u is $(\varepsilon, 1, j)$-selfsimilar at 0 with respect to $V^j \times (-\infty, s]$.*

– *if $d(y, V) \geq \rho$, then u is $(\varepsilon, 1, j + 1)$-selfsimilar at 0 with respect to* $\mathrm{span}\{y, V^j\} \times (-\infty, \max\{s, T\}]$.

• *If $W = V^{j-2} \times \mathbb{R}$, then $d(y, V) \geq \rho$ and u is $(\varepsilon, 1, j+1)$-selfsimilar with respect to* $\mathrm{span}\{y, V^{j-2}\} \times \mathbb{R}$.

Using also Lemma 2.4, by induction/contradiction we now obtain:

Corollary 2.12 (Line-up in tubular neighborhoods)**.** *For all $\mu, \nu > 0$, $\Lambda_1, \Lambda_2 < \infty$, m and N there exist $\delta = \delta(\mu, \nu, \Lambda_1, \Lambda_2, m, N) > 0$ and $A = A(\mu, \nu, \Lambda_1, \Lambda_2, m, N) < \infty$ such that the following holds: If $u \in H^1_{\Lambda_1, \Lambda_2}(P_4, N)$ and $X \in L_{Ar, r/A, \delta}(u)$ for some $r \leq 1/A$, then there exists $0 \leq \ell \leq m + 2$ and W^ℓ_X such that*

1. *u is (μ, r, ℓ)-selfsimilar at X with respect to W^ℓ_X ,*

2. *$L_{Ar, r/A, \delta}(u) \cap P_r(X) \subseteq T_{\nu r}(W^\ell_X)$.*

To deal with the quasistatic case we also need the following lemma.

Lemma 2.13 (Quantitative behavior in the quasistatic case)**.** *For all $\varepsilon, \gamma > 0$, $\Lambda_1, \Lambda_2 < \infty$, m and N there exists $\delta = \delta(\varepsilon, \gamma, \Lambda_1, \Lambda_2, m, N) > 0$, such that the following holds: If $u \in H^1_{\Lambda_1, \Lambda_2}(P_4, N)$ is $(\delta, 1, \ell)$-selfsimilar at 0 with respect to $W = V^\ell \times (-\infty, T]$ and if $Y = (y, s) \in P_{1-2\gamma}$ with $s \leq T - (2\gamma)^2$ then u is $(\varepsilon, \gamma, \ell + 2)$-selfsimilar at Y with respect to $W = (y + V^\ell) \times \mathbb{R}$.*

Proof. If not, passing to limits we obtain φ that is ℓ-selfsimilar on P_1 with respect to $W = V^\ell \times (-\infty, T]$ and a point $Y = (y, s) \in \overline{P}_{1-2\gamma}$ with $s \leq T - (2\gamma)^2$ such that φ is not $(\varepsilon, \gamma, \ell + 2)$-selfsimilar at Y with respect to $W = (y + V^\ell) \times \mathbb{R}$, a contradiction. $\square$

2.3 Conclusion of the argument

Proof of Theorem 1.11. Let $\varepsilon, \eta, \Lambda_1, \Lambda_2, m, N$ be as in the statement of the theorem. It is convenient to choose $\gamma := c_0(m)^{-\frac{2}{\varepsilon}}$, where $c_0(m)$ is a geometric constant that only depends on the dimension and will appear below (roughly a doubling constant). Now we apply Corollary 2.12 with $\nu = \gamma/2$ and $\mu \leq \eta$ small enough such that also the below application of Lemma 2.13 is justified, and get constants δ and A. Choose an integer q, such that $\gamma^q \leq 1/(2A)$. Setting $Q := \lfloor (2q + 1)\delta^{-1}\Lambda_2 \rfloor + q$, from the argument in Section 2.1, for any u satisfying the assumptions of the theorem we get a decomposition of P_1 into at most β^Q nonempty sets $E_{T^\beta}(u)$.

Lemma 2.14 (Covering lemma)**.** *There exists $c_0(m) < \infty$ such that each set $\mathcal{S}^j_{\eta, \gamma^\beta}(u) \cap E_{T^\beta}(u)$ can be covered by at most $c_0(c_0\gamma^{-(m+2)})^Q(c_0\gamma^{-j})^{\beta - Q}$ balls of radius γ^β.*

Proof. We will recursively define a covering. For $\beta = 0$ pick some minimal covering of $\mathcal{S}^j_{\eta,\gamma^0}(u)$ by balls of radius 1 with centers in $\mathcal{S}^j_{\eta,\gamma^0}(u) \cap P_1$. Note that $\mathcal{S}^j_{\eta,\gamma^{\beta+1}}(u) \subset \mathcal{S}^j_{\eta,\gamma^\beta}(u)$. Let T^β be the β-tupel obtained from dropping the last entry from $T^{\beta+1}$. Then we also have $E_{T^{\beta+1}}(u) \subset E_{T^\beta}(u)$.

Recursion step. For each ball $P_{\gamma^\beta}(X)$ in the covering of $\mathcal{S}^j_{\eta,\gamma^\beta}(u) \cap E_{T^\beta}(u)$, take a minimal covering of $P_{\gamma^\beta}(X) \cap \mathcal{S}^j_{\eta,\gamma^{\beta+1}}(u) \cap E_{T^{\beta+1}}(u)$ by balls of radius $\gamma^{\beta+1}$ with centers in $P_{\gamma^\beta}(X) \cap \mathcal{S}^j_{\eta,\gamma^{\beta+1}}(u) \cap E_{T^{\beta+1}}(u)$.

Let us now explain that this covering has indeed the desired properties. First observe that, for all β, the number of balls in a minimal covering from the recursion step is at most

$$c(m)\gamma^{-(m+2)}. \tag{2.15}$$

However, if $T_\beta(X) = 0$, then $X \in L_{2A\gamma^\beta, 2\gamma^\beta/A, \delta}(u)$ and Corollary 2.12 gives us $0 \leq \ell \leq n+2$ and W^ℓ_X such that

1. u is $(\mu, 2\gamma^\beta, \ell)$-selfsimilar at X with respect to W^ℓ_X ,

2. $L_{2A\gamma^\beta, 2\gamma^\beta/A, \delta}(u) \cap P_{2\gamma^\beta}(X) \subseteq T_{\gamma^{\beta+1}}(W^\ell_X)$.

Since $X \in \mathcal{S}^j_{\eta,\gamma^\beta}(u)$ we must have $\ell \leq j$. Since $E_{T^\beta}(u) \subset L_{2A\gamma^\beta, 2\gamma^\beta/A, \delta}(u)$ this implies the following better estimate for the number of balls in a minimal covering:

$$c(m)\gamma^{-j}. \tag{2.16}$$

Indeed, the estimate is clear in the cases $W^\ell_X = (x + V^\ell) \times \{t\}$ and $W^\ell_X = (x + V^{\ell-2}) \times \mathbb{R}$, the case $W^\ell_X = (x + V^\ell) \times (-\infty, T]$ requires some extra thought, but in fact only if $T \leq t + (2\gamma^\beta)^2$ and $\ell \geq j-1$ which we will assume now. So, if $Y = (y, s) \in P_{\gamma^\beta}(X) \cap \mathcal{S}^j_{\eta,\gamma^{\beta+1}}(u) \cap E_{T^{\beta+1}}(u)$, then by Lemma 2.13 we conclude $s \geq T - (2\gamma^{\beta+1})^2$ and thus (2.16) holds also in the quasistatic case. By the quantitative differentiation argument, the better estimate (2.16) applies with at most Q exceptions. This proves the lemma. $\qquad\square$

We will now conclude the proof of Theorem 1.11 by estimating the volume of the covering. The volume of balls in $\mathbb{R}^{m,1}$ satisfies

$$\mathrm{Vol}(P_{\gamma^\beta}(X)) = w_m(\gamma^\beta)^{m+2}, \tag{2.17}$$

which together with the choice of γ and the fact that polynomials grow slower than exponentials, i.e. with

$$c_0^\beta = (\gamma^\beta)^{-\frac{\varepsilon}{2}},$$

$$\beta^Q \leq c(m, Q)(\gamma^\beta)^{-\frac{\varepsilon}{2}},$$

11

gives (recalling again the decomposition of P_1 into at most β^Q nonempty sets $S^j_{\eta,\gamma^\beta}(u) \cap E_{T^\beta}(u)$ and the Covering Lemma 2.14)

$$
\begin{aligned}
\mathrm{Vol}(S^j_{\eta,\gamma^\beta}(u)) &\le \beta^Q \left[c_0 (c_0 \gamma^{-(m+2)})^Q (c_0 \gamma^{-j})^{\beta-Q} \right] w_m (\gamma^\beta)^{m+2} \\
&\le c(m,Q,\varepsilon) \beta^Q c_0^\beta (\gamma^\beta)^{m+2-j} \\
&\le c(m,Q,\varepsilon) (\gamma^\beta)^{m+2-j-\varepsilon} .
\end{aligned}
\tag{2.18}
$$

From the above, for all $0 < r < 1$, we get

$$
\begin{aligned}
\mathrm{Vol}(S^j_{\eta,r}(u)) &\le c(m,Q,\varepsilon)(\gamma^{-1} r)^{m+2-j-\varepsilon} \\
&\le c(\varepsilon,\eta,\Lambda_1,\Lambda_2,m,N) r^{m+2-j-\varepsilon} .
\end{aligned}
$$

It follows that

$$
\mathrm{Vol}\left(T_r(S^j_{\eta,r}(u)) \right) \le C r^{m+2-j-\varepsilon} \qquad (0 < r < 1) .
$$

for another constant $C = C(\varepsilon,\eta,\Lambda_1,\Lambda_2,m,N)$, and this finishes the proof of Theorem 1.11. $\qquad\square$

3 Higher regularity for certain targets

To finish the proofs of Theorem 1.20 and Corollary 1.24, we prove the following ε-regulartiy lemma. Roughly speaking, it says that enough approximate degrees of symmetry imply regularity.

Lemma 3.1 (ε-regularity). *If N does not admit harmonic two-spheres, then there exists an $\varepsilon = \varepsilon(m,N,\Lambda_1,\Lambda_2,R) > 0$ such that the following holds. Every Chen-Struwe solution $u \in H^1_{\Lambda_1,\Lambda_2}(P_{4R}, N)$ that is $(\varepsilon, 2r, m-1)$-selfsimilar at 0 for some $r \le R$, satisfies $r_u(0) \ge r$.*

More generally, if N does not admit harmonic ℓ-spheres for $\ell = 2,\ldots,k$ and quasi-harmonic ℓ-spheres for $\ell = 3,\ldots,k-1$, then it suffices to assume that u is $(\varepsilon, r, m-k+1)$-selfsimilar.

Proof. If not, there are Chen-Struwe solutions $u_\alpha : P_2 \to N$ with bounded scale invariant energy that are $(\alpha^{-1}, 2, m-1)$-selfsimilar at 0, but such that $r_{u_\alpha}(0) < 1$. By [LW99, Prop. 4.1] after passing to a subsequence we can find a strong limit $u_\alpha \to u$ in $H^1_{\mathrm{loc}}(P_2, N)$. This limit is a $(m-1)$-selfsimilar weak solution of the harmonic map flow that satisfies the monotonicity formula and the standard ε-regularity lemma, and whose singular set has vanishing m-dimensional parabolic Hausdorff measure, see again [LW99, Prop. 4.1]. Thus u must be of the form $u(x,t) = \psi(\frac{y}{|y|})$, where we have written $x = (y,z)$ with $y \in \mathbb{R}^3$. Since by assumption the target does not admit harmonic two-spheres, u must be constant. Then the standard ε-regularity lemma [Str88, Thm. 5.3] together with the strong convergence in H^1_{loc} implies that $r_{u_\alpha}(0) > 1$ for α large enough; a contradiction. Arguing similarly, the more general statement follows by induction. $\qquad\square$

Proof of Theorem 1.20. Since the initial map is smooth, the regularity scale is bounded below for small times. Thus, by covering $M \times \mathbb{R}_+$ with parabolic balls, we can reduce the problem to the local

setting of Theorem 1.11. After this reduction, Lemma 3.1 implies that $\mathcal{B}_r(u) \subseteq \mathcal{S}_{\eta,r}^{m-k}(u)$ for η small enough, and the claim follows from the volume estimate of Theorem 1.11. $\qquad \square$

Proof of Corollary 1.24. Using the layer-cake formula, this follows immediately from Theorem 1.20, and standard interior estimates (Remark 1.17). $\qquad \square$

4 Examples without harmonic 2-spheres

We end by considering compact Riemannian manifolds N which admit no harmonic 2-spheres. It has become a common phenomena in the theory of harmonic maps that when the target space N does not admit any harmonic 2-spheres then the regularity of a harmonic mapping improves. In particular, a *stationary* harmonic map has the same regularity properties as a minimizing harmonic map, see [Lin99] and Theorem 1.20 and Corollary 1.26. It is known from the work of Sacks-Uhlenbeck that in this case the universal cover of N must be contractible. However in the literature at this point the only examples of such spaces N have nonpositive sectional curvature. In particular, for these examples other methods can be used to arrive at much stronger regularity results. The goal of this Section is simply to record a few examples of compact manifolds without harmonic 2-spheres which do not have nonpositive sectional curvature.

Theorem 4.1. *Let N be one of the following:*

1. *A 3-dimensional infranil manifold equipped with a left invariant metric.*

2. *A 3-dimensional solv-manifold with a left invariant metric.*

Then N admits no harmonic 2-spheres. In particular, products of such spaces with each other or with spaces of nonpositive sectional curvature also admit no harmonic 2-spheres.

Question 4.2. Do there exist any nilmanifolds N with left invariant metrics that admit harmonic 2-spheres? What can be said about the existence of harmonic k-spheres?

We are indebted to Harold Rosenberg for the following simple argument.

Proof. View N is the quotient of its universal cover $\tilde{N}$ by a discrete cocompact group Γ acting on the left by isometries. Since S^2 is simply connected we can lift any mapping into N to $\tilde{N}$, Since the image of a harmonic 2-sphere is a branched minimal surface, it suffices to show that $\tilde{N}$ contains no branched minimal 2-spheres. Let us first deal with the case when N is nilpotent. In particular, unless $\tilde{N}$ is euclidean space, it is the Heisenberg group. If η is the lie algebra of $\tilde{N}$ equipped with an inner product, let e_1, e_2, e_3 be an orthonormal basis with $[\eta, \eta] = \text{span}\{e_3\}$. A standard computation shows that the closed normal subgroup $\tilde{N}' \subseteq \tilde{N}$ associated to the subalgebra $\eta' \equiv \text{span}\{e_2, e_3\}$ is minimal. The family of hypersurfaces $n \cdot \tilde{N}'$ with $n \in N$ gives a foliation of N

by minimal hypersurfaces which is preserved by the left action of Γ, with orbit space $\tilde{N}/\tilde{N}' = \mathbf{R}$. If $S^2 \to \tilde{N}$ is a branched minimal sphere, then there exists $n \in N$ such that $n \cdot N'$ touches the S^2 from one side. By a standard maximum principle argument, this is a contradiction unless the image of S^2 is a point.

If N is a compact solv-manifold, then we can pick an orthonormal basis of the lie algebra e_1, e_2, e_3 such that $[e_1, e_2] = 0$, $[e_1, e_3] = e_1$, $[e_2, e_3] = -e_2$. If we let $\eta' = \mathrm{span}\{e_1, e_2\}$ be a subalgebra then the brackets allow us to compute the mean curvature of the corresponding closed normal subgroup as $\langle \nabla_{e_1} e_1, e_3 \rangle + \nabla_{e_2} e_2, e_3 \rangle = 1 - 1 = 0$. Now one can argue as in the nilpotent case to finish the proof. $\qquad\square$

References

[Che12] J. Cheeger. Quantitative differentiation: a general formulation. *Comm. Pure Appl. Math.*, 65(12):1641–1670, 2012.

[CHN13] J. Cheeger, R. Haslhofer, and A. Naber. Quantitative stratification and the regularity of mean curvature flow. *Geom. Funct. Anal.*, 23(3):828–847, 2013.

[CLL95] Y. M. Chen, J. Li, and F.-H. Lin. Partial regularity for weak heat flows into spheres. *Comm. Pure Appl. Math.*, 48(4):429–448, 1995.

[CN13a] J. Cheeger and A. Naber. Lower bounds on Ricci curvature and quantitative behavior of singular sets. *Invent. Math.*, 191(2):321–339, 2013.

[CN13b] J. Cheeger and A. Naber. Quantitative stratification and the regularity of harmonic maps and minimal currents. *Comm. Pure and Appl. Math.*, 66(6):965–990, 2013.

[CS89] Y. M. Chen and M. Struwe. Existence and partial regularity results for the heat flow for harmonic maps. *Math. Z.*, 201(1):83–103, 1989.

[ES64] J. Eells, Jr. and J. H. Sampson. Harmonic mappings of Riemannian manifolds. *Amer. J. Math.*, 86:109–160, 1964.

[Fel94] M. Feldman. Partial regularity for harmonic maps of evolution into spheres. *Comm. Partial Differential Equations*, 19(5-6):761–790, 1994.

[Lin99] F.-H. Lin. Gradient estimates and blow-up analysis for stationary harmonic maps. *Ann. of Math. (2)*, 149(3):785–829, 1999.

[LW99] F.-H. Lin and C. Y. Wang. Harmonic and quasi-harmonic spheres. *Comm. Anal. Geom.*, 7(2):397–429, 1999.

[Str88]   M. Struwe. On the evolution of harmonic maps in higher dimensions. *J. Differential Geom.*, 28(3):485–502, 1988.

[SU81]   J. Sacks and K. Uhlenbeck. The existence of minimal immersions of 2-spheres. *Ann. of Math. (2)*, 113(1):1–24, 1981.

JEFF CHEEGER AND ROBERT HASLHOFER, COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, 251 MERCER STREET, NEW YORK, NY 10012, USA

AARON NABER, DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MA 02139, USA

E-mail: robert.haslhofer@cims.nyu.edu, cheeger@cims.nyu.edu, anaber@math.mit.edu