Journal of Functional Analysis 289 (2025) 111006

Contents lists available at ScienceDirect

Journal of Functional Analysis

journal homepage: www.elsevier.com/locate/jfa

Regular Article

Low energy levels of harmonic spheres in analytic manifolds

Melanie Rupflin [1]

Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK

ARTICLE INFO

Article history:
Received 13 September 2023
Accepted 9 April 2025
Available online 14 April 2025
Communicated by Guido De Philippis

Keywords:
Harmonic maps
Lojasiewicz estimates
Quantitative stability of critical points

ABSTRACT

We consider the energy spectrum $\Xi_E(N)$ of harmonic maps from the sphere into a closed Riemannian manifold N. While a well known conjecture asserts that $\Xi_E(N)$ is discrete whenever N is analytic, for most analytic targets it is only known that any potential accumulation point of the energy spectrum must be given by the sum of the energies of at least two harmonic spheres. The lowest energy level that could hence potentially be an accumulation point of Ξ_E is thus $2E_{\min}$. In the present paper we exclude this possibility for generic 3 manifolds and prove additional results that establish obstructions to the gluing of harmonic spheres and provide Łojasiewicz-estimates for almost harmonic maps.

1. Introduction

Let (N, g_N) be a closed Riemannian manifold which we can assume without loss of generality to be embedded in a suitable Euclidean space $\mathbb{R}^N$ using Nash's embedding

E-mail address: melanie.rupflin@maths.ox.ac.uk.

[1] For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

https://doi.org/10.1016/j.jfa.2025.111006

theorem. We recall that a map $u : S^2 \to N$ is called harmonic if it is a critical point of the Dirichlet energy

$$E(u) = \tfrac{1}{2} \int_{S^2} |\nabla u|^2 dv_{g_{S^2}}.$$

Such harmonic maps are characterised by $\tau_{g_{S^2}}(u) = 0$ where the tension field is given by $\tau_{g_{S^2}}(u) = -\nabla^{L^2} E(u) = \Delta_{g_{S^2}} u + A(u)(\nabla u, \nabla u)$, A the second fundamental form of $N \hookrightarrow \mathbb{R}^N$.

We recall that any harmonic map from S^2 is (weakly) conformal and hence that the set of harmonic maps from the sphere coincides with the set of (weakly) conformal parametrisations of minimal spheres in N. In particular, the energy spectrum

$$\Xi_E(N, g_N) := \{E(u) : u : S^2 \to N \text{ harmonic, not constant } \}$$

of harmonic spheres agrees with the areas (counted with multiplicity) of (possibly branched) immersed minimal spheres.

While there are many known examples of smooth manifolds (N, g_N) for which $\Xi_E(N, g_N)$ has accumulation points, such as warped products with S^2 for warping functions with non-discrete critical values, a well known conjecture of Leon Simon and of Fang-Hua Lin [5] asserts that the energy spectrum of harmonic maps from S^2 into any closed analytic manifold is discrete.

In some very special cases this follows as a consequence of much stronger results such as an explicit characterisation of all minimal spheres in N and the resulting explicit knowledge of $\Xi_E(N)$. E.g. if the target is the round sphere then all harmonic spheres are described in stereographic coordinates $z \in \mathbb{C}$ by rational maps in either z or $\bar{z}$ and hence $\Xi(S^2, g_{S^2})$ is made up by the multiples of 4π.

Conversely, for more general analytic targets very little is known about Ξ_E with the few existing results following already from the compactness theory of (almost) harmonic maps developed in the 1990s in [14,1,10,4] and the seminal result of Simon [13] from 1983.

To describe the known properties of Ξ_E we first recall that sequences u_j of maps with bounded energy which are almost harmonic in the sense that

$$\|\tau_{g_{S^2}}(u_j)\|_{L^2(S^2, g_{S^2})} \to 0 \tag{1.1}$$

subconverge to a bubble tree without loss of energy or formation of necks: That is, there exists a harmonic base map $\omega_0 : S^2 \to N$ and a finite set $S \subset S^2$ so that $u_j \to \omega_0$ strongly in $H^2_{loc}(S^2 \setminus S)$ and weakly in $H^1(S^2)$ and so that near each point $p \in S$ the u_j are essentially described by a collection of highly concentrated harmonic spheres. Namely, for each such p there exists a finite collection of harmonic maps $\omega_k : \mathbb{R}^2 \to N$, points $p_j^k \to p$ and scales $\mu_j^k \to \infty$ so that, working in stereographic coordinates centred at p,

$$u_j(x) - \sum \left(\omega_k(\mu_j^k(x - p_j^k)) - \omega_k(\infty)\right) - \omega_0(x) \to 0 \text{ strongly in } L^\infty \cap H^1(B_r(p)). \quad (1.2)$$

We recall that the conformal invariance of the energy and the result [12] of Sacks-Uhlenbeck imply that a finite energy map $u : \mathbb{R}^2 \to N$ is harmonic if and only if $u \circ \pi^{-1} : S^2 \to N$ is a harmonic map from the sphere, $\pi(x) = (\frac{2x}{1+|x|^2}, \frac{1-|x|^2}{1+|x|^2}) : \mathbb{R}^2 \cup \{\infty\} \to S^2$ the inverse stereographic projection. In the following we can thus switch viewpoint and work on $\mathbb{R}^2$, or equivalently on $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, whenever this is more convenient and, except in the trivial case where $E(u_j) \to 0$, can pull-back a sequence of maps u_j satisfying (1.2) by suitable Möbius transforms to ensure that the base map ω_0 is non-constant.

If no bubbles form, i.e. if the maps converge strongly in $H^2(S^2)$, and if N is analytic then the seminal results [13] of Simon are applicable. These ensure that for every harmonic map $u^* : S^2 \to N$ there exists an $\varepsilon = \varepsilon(u^*) > 0$ so that a Łojasiewicz-estimate of the form

$$|E(u) - E(u^*)| \leq C\|\tau_{g_{S^2}}(u)\|^\alpha_{L^2(S^2, g_{S^2})} \text{ for some } \alpha = \alpha(u^*) \in (1, 2] \quad (1.3)$$

holds true for all maps in the neighbourhood

$$\mathcal{U}(u^*) = \{u : S^2 \to N : \|u - u^*\|_{L^\infty} < \varepsilon \text{ and } \|u - u^*\|_{H^1} < \varepsilon\}.$$

Conversely these results cannot be applied to study sequences of harmonic maps that undergo bubbling as such maps will not be H^1 close to a fixed harmonic map, or indeed to any set of harmonic maps for which the results of [13] yield Łojasiewicz-estimates on a *uniform* ε-neighbourhood, compare Remark 4.3. To the authors knowledge the only properties of the energy spectrum of harmonic maps into general analytic manifolds that are currently known are the ones that follow from the above mentioned compactness theory and from the work of Simon, i.e. that

- The lowest non-trivial energy level $E_{\min} := \inf \Xi_E(N)$ is achieved and isolated in the energy spectrum.
- Any potential accumulation point of $\Xi_E(N)$ must be of the form $\bar{E} = \sum_{i=0}^n E(\omega_i)$ for some non-trivial harmonic spheres ω_i, $i = 0, \dots, n \geq 1$, and such an accumulation point must be due to the existence of harmonic maps u_j with energies $\bar{E} \neq E(u_j) \to \bar{E}$ that converge to a non-trivial bubble tree.

The lowest level which might potentially be an accumulation point of the energy spectrum is hence $2E_{\min}$ and so a natural starting point to develop a better understanding of Ξ_E, and to gain new insight into the above mentioned conjecture, is to investigate whether energies of harmonic maps might accumulate at $2E_{\min}$. In the present paper we exclude this possibility for 3 manifolds under some natural assumptions on the minimal energy harmonic spheres

Theorem 1.1. *Let (N, g_N) be any analytic 3 manifold and let $E_{min} = \min \Xi_E(N, g_N)$ be the minimal energy of a non-trivial harmonic sphere in N. Suppose that harmonic spheres $\omega : S^2 \to N$ with this energy E_{min} are unbranched, non-degenerate critical points* *and that any intersection or self-intersection between such minimal spheres is transversal in the sense that the corresponding tangent spaces do not coincide.*

Then $2E_{min}$ cannot be an accumulation point of the energy spectrum $\Xi_E(N, g_N)$.

As the energy is invariant under conformal changes we know that any vector field of the form $Y = \frac{d}{d\varepsilon}|_{\varepsilon=0}\, \omega \circ M_\varepsilon$, M_ε a family of Möbius transforms with $M_{\varepsilon=0} = \mathrm{Id}$, is a Jacobi-field along ω, i.e. so that $d^2E(\omega)(Y, \cdot) = 0$. We hence say that ω is non-degenerate if all Jacobi-fields along ω are generated in this way.

We recall that the results of Gulliver, Osserman and Royden [2] ensure that minimal surfaces in 3 manifolds that are locally area minimising cannot have any true branch points. Any harmonic sphere with energy $E_{\min}$ that is a local area minimiser will hence automatically be unbranched. Of course the example of an equatorial sphere in S^3 shows that harmonic spheres of minimal energy might not be local minimisers of the Area and to the author's knowledge it is not known whether there any minimal energy harmonic spheres in a 3 manifold that are branched.

The above theorem is valid also for manifolds which are not analytic but for which it is known that prime harmonic spheres of energy $2E_{\min}$ are non-degenerate critical points.

As the bumpy metric theorems obtained by White in [19–21] and by Moore in [8] and in Theorems 5.1.1 and 5.1.2 of [9], ensure that for generic metrics on manifolds of dimension at least 3 all prime harmonic maps are non-degenerate critical points, not branched and that all potential intersections and self-intersections are transverse we hence obtain

Corollary 1.2. *For generic* 3 *manifolds the possibility that $2E_{min}$ is an accumulation point of the energy spectrum $\Xi_E(N)$ is excluded and Łojasiewicz-estimates as stated in Theorems 1.3, 1.5 and 1.7 below hold true for any sequence of almost harmonic maps that converges to a bubble tree with a single bubble.*

We remark that while standard arguments show that the Łojasiewicz-estimate (1.3) holds true (with $\alpha = 2$) in an $\varepsilon = \varepsilon(u^*) > 0$ neighbourhood of any non-degenerate harmonic map u^* into a smooth manifold, these arguments cannot be used to obtain the above result as sequences of harmonic maps undergoing bubbling will not be in a neighbourhood of a fixed such map.

To prove Theorem 1.1 and the above corollary we need to exclude the possibility that there exists a sequence of harmonic maps $u_j : S^2 \to N$ with energy $2E_{\min} \neq E(u_j) \to 2E_{\min}$. We know that such a sequence cannot converge strongly as that would contradict [13], so (after pull-back by suitable Möbius transforms) must converge to a bubble tree with a base map ω_0 and a bubble ω_1 of energy $E(\omega_0) = E(\omega_1) = E_{\min}$.

We hence need to ask whether it is possible to glue increasingly concentrated harmonic spheres onto harmonic base maps in a way that results in a harmonic map whose energy is close, but not equal, to $E(\omega_0)+E(\omega_1)$. We will not only exclude this for harmonic spheres as considered in Theorem 1.1, but will establish obstructions to gluing harmonic spheres that apply in more general settings, including situations where the involved harmonic spheres are branched and have non-trivial, and even non-integrable, Jacobi-fields.

It is natural to distinguish between bubble trees for which the bubble ω_1 and base ω_0

(1) parametrise the same minimal surface with the same orientation
(2) parametrise the same minimal surface with the opposite orientation
(3) parametrise transversally intersecting minimal surfaces.

We note that the no-neck property of the convergence to a bubble tree ensures that we only have to consider maps ω_0 and ω_1 for which $\omega_0(p) = \omega_1(\infty)$, p the point at which the bubble forms. After rotating the domain, we can assume without loss of generality that the bubble is attached at the north pole, so in stereographic coordinates at $z = 0$. In the following we hence only have to consider maps $u_j : \hat{\mathbb{C}} \to N$ for which there exist $\mu_j \to \infty$ and $z_j \to 0$ for which

$$u_j(z) - [\omega_0(z) + \omega_1(\mu_j(z - z_j)) - \omega_1(\infty)] \to 0 \text{ in } \dot{H}^1 \cap L^\infty(\mathbb{C}). \tag{1.4}$$

Our analysis is not restricted to bubble trees for which ω_0 and ω_1 are unbranched but we can more generally consider bubble trees with base maps and bubbles which are obtained as a composition of a harmonic sphere $U^* : \hat{\mathbb{C}} \to N$ with $dU^*(q^*(0)) \neq 0$ and a rational map $q^* : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ of arbitrary degree, i.e. a non-constant map of the form $q^*(z) = \frac{p_1(z)}{p_2(z)}$, $p_{1,2}$ polynomials. As we can precompose U^* with a rotation we can restrict our attention to rational maps $q^*(z)$ which map 0 to 0 and so will always ask that $dU^*(0) \neq 0$.

We note that the conformal invariance of the energy ensures that any map of the form $U \circ q$, U harmonic and q rational, will be a harmonic map and that any vector field of the form $Y = \frac{d}{d\varepsilon}|_{\varepsilon=0}(U \circ q_\varepsilon)$, q_ε rational maps, is a Jacobi-field along $U \circ q_{\varepsilon=0}$. For higher order coverings of harmonic spheres it is hence natural to say that

Definition 1. The second variation of the energy is non-degenerate along a harmonic map ω of the form $\omega = U \circ q$ if every Jacobi-field Y along ω is generated by a variation of rational maps q_ε, i.e. given by $Y = \frac{d}{d\varepsilon}|_{\varepsilon=0}(U \circ q_\varepsilon)$.

We first consider the case where the base and the bubble parametrise the same minimal surface with the same orientation. Here we include settings in which $U^*(0)$ is a point of higher multiplicity of an immersed minimal sphere $U^*(\hat{\mathbb{C}})$ provided ω_0 and $\omega_1(\frac{1}{z})$ parametrise the same leaf of $U^*(\hat{\mathbb{C}})$ near $z = 0$, but will instead include settings in which they parametrise transversally intersecting leafs in the third case, compare Theorem 1.7 below.

So suppose that ω_0 and ω_1 are obtained by composing the same harmonic $U^* : \hat{\mathbb{C}} \to N$ with rational maps $q_0^*(z)$ and $q_1^*(\frac{1}{z})$ for which $q_0^*(0) = q_1^*(0) = 0$. We want to show that the only sequences of harmonic maps u_j that converge to such a bubble tree are higher degree coverings of $U^* : S^2 \to N$.

This does not follow from the existing theory since the size of the neighbourhoods on which Łojasiewicz-estimates are known to hold true shrink to zero as the maps undergo bubbling, compare Remark 4.3. If u_j is a sequence of harmonic maps that converges to such a bubble tree then the compactness theory from [14,1,10,4] ensures that $\|u_j - U^* \circ r_j\|_{L^\infty \cap \dot{H}^1} \to 0$ for some rational maps r_j. However the existing theory does not provide the quantitative estimates on the rate of this convergence which would allow one to know that the maps u_j are in the smaller and smaller neighbourhoods $\mathcal{U}(U^* \circ r_j)$ for which the results of Simon apply.

Conversely, our method allows us to prove that Łojasiewicz-estimates indeed hold true on a *uniform* ε neighbourhood around such a non-compact set of harmonic maps $\{U^* \circ r\}$ and hence that such sequences of harmonic maps cannot be responsible for an accumulation point of the energy spectrum. To be more precise, we show

Theorem 1.3. *Let (N, g_N) be a smooth Riemannian manifold of any dimension. Suppose that $U^* : \hat{\mathbb{C}} \to N$ is a harmonic map with $dU^*(0) \neq 0$ and $q_0^*, q_1^* : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ are rational maps with $q_1^*(0) = q_0^*(0) = 0$ so that d^2E is non-degenerate at $\omega_0(z) := U^* \circ {q_0}^*$ and $\omega_1(z) := U^* \circ {q_1}^*(\frac{1}{z})$ in the sense of Definition 1.*

Let (u_j) be any sequence of almost harmonic maps which converges to a bubble tree with base map ω_0 and bubble ω_1 as described in (1.1) and (1.4). Then there exists a constant C so that

$$|E(u_j) - [E(\omega_0) + E(\omega_1)]| \leq C \min(\|dE(u_j)\|_*, \log \mu_j \|dE(u_j)\|_*^2), \tag{1.5}$$

for all sufficiently large j and for $\mu_j \to \infty$ as in (1.4).

Furthermore, if the maps u_j are harmonic, then, after precomposing with a suitable rotation R_j of the domain, they can be written as

$$u_j(z) = U^*\big(q_0^j(z) + q_1^j(\tfrac{1}{\mu_j z})\big)$$

for rational maps q_i^j which converge smoothly to q_i^* when viewed as maps from S^2 to S^2.

Here and in the following $\|\cdot\|_*$ denotes a weighted H^{-1} norm which is defined in Remark 4.1 and which is chosen so that the above estimate (1.5) is invariant under pull-back by Möbius transforms. We note that $\|\tau_{g_{S^2}}(u)\|_{L^2(S^2, g_{S^2})}$, which dominates $\|dE(u)\|_*$, does not have this property. However, while it is natural work with $\|dE(u)\|_*$ to analyse the energy spectrum, it is also of interest to obtain L^2-Łojasiewicz-estimates e.g. to analyse the asymptotic behaviour of the harmonic map heat flow

$$\partial_t u = \tau_{g_{S^2}}(u). \tag{1.6}$$

We shall hence also prove

Corollary 1.4. *For sequences of maps u_j as in Theorem 1.3 we can bound*

$$|E(u_j) - E^*| \leq C\|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2,g_{S^2})}. \tag{1.7}$$

A consequence of this corollary is that solutions $u(t)$ of the flow (1.6) for which sequences $u(t_j)$, $t_j \to \infty$, are known to subconverge to a bubble tree as considered in Theorem 1.3 must indeed converge exponentially fast to a unique base map $u_\infty : S^2 \to N$ as $t \to \infty$ away from a unique point $p_\infty \in S^2$ in the sense of both $L^2(S^2, g_{S^2})$ and $C^k_{loc}(S^2 \setminus \{p_\infty\})$.

In the one bubble case this extends results that were previously obtained by Topping [15] in the case where N is the round 2-sphere and that were later extended by Liu and Yang in [6] to compact Kähler manifolds with nonnegative holomorphic bisectional curvature. We note that the results from [15,6] are not restricted to the one bubble case but apply to sequences of maps and solutions of the flow which converge to any bubble tree for which the base map and the bubbles are all parametrisations of the same minimal spheres with the same orientation.

The above corollary suggests that while the special geometric structure of the target is crucially used in the proofs of [15,6], analogue results might be valid for more general targets and one would expect that the main property needed is the non-degeneracy of the second variation of the energy at the underlying harmonic spheres $U^* \circ q_i^*$, which for $N = S^2$ follows from [3].

Next we want to show that it is impossible to glue a highly concentrated bubble onto a base map which parametrises the same minimal surface, but with opposite orientation. Indeed, we will prove more generally that for almost harmonic maps that converge to such a bubble tree the energy defect and the bubble scale are controlled in terms of $\|dE(u)\|_*$.

Theorem 1.5. *Let (N, g_N) be a smooth Riemannian manifold of any dimension, let $U^* : \hat{\mathbb{C}} \to N$ be any harmonic map with $dU^*(0) \neq 0$ and let $q_0^*, q_1^* : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be any rational maps with $q_1^*(0) = q_0^*(0) = 0$ so that d^2E is non-degenerate at $\omega_0(z) := U^* \circ q_0{}^*$ and $\omega_1(z) := U^* \circ q_1{}^*(\frac{1}{\bar{z}})$ in the sense of Definition 1.*

Let (u_j) be any sequence of almost harmonic maps which converges to a bubble tree with base map ω_0 and bubble ω_1 as described in (1.1) and (1.4). Then the bubble scale must satisfy

$$\mu_j^{-\min(n_0^*, n_1^*)} \leq C\|dE(u_j)\|_*,$$

$n_i^* \in \mathbb{N}$ the order of the zero of q_i^* at $z = 0$, while the energy defect is controlled by

$$|E(u_j) - [E(\omega_0) + E(\omega_1)]| \leq C\|dE(u_j)\|_*.$$

In particular, no sequence of harmonic maps can converge to such a bubble tree.

For maps to the round sphere, L^2-Łojasiewicz-estimates with exponent $\alpha = 2$ and exponential bounds on the bubble scale were proven by Topping in [16]. In this setting there are only two types of harmonic spheres, namely holomorphic or antiholomorphic maps. The results of [16] apply for maps that converge to any bubble tree which is so that for each singular point p the bubbles forming at p are all of the same type, where for points p at which bubbles form that have a different type than the base map one needs to impose additionally that the base map is not branched.

Topping already observed in [17] that even for maps into (S^2, g_{S^2}) one cannot expect to prove L^2-Łojasiewicz-estimates with exponent $\alpha = 2$ if the base map is branched at a point where a bubble of a different type forms and we remark that such settings have been addressed very recently by Waldron in [18].

We also recall that any L^2-Łojasiewicz-estimate with an exponent $\alpha > 1$ is sufficient to prove convergence results for solutions of harmonic map flow. As this is our main motivation to also prove Łojasiewicz-estimates that involve $\|\tau_{g_{S^2}}(u)\|_{L^2(S^2,g_{S^2})}$ rather than the weaker, and scaling invariant $\|dE(u)\|_*$, we shall hence simply show

Corollary 1.6. *For any sequence (u_j) of almost harmonic maps as in Theorem 1.5 we can bound*

$$|E(u_j) - [E(\omega_0) + E(\omega_1)]| \leq C\|\tau_{g_{S^2}}(u_j)\|^{\alpha}_{L^2(S^2,g_{S^2})} \tag{1.8}$$

for an exponent $\alpha > 1$ that only depends on U^* and $q^*_{0,1}$.

We finally want to show that for maps into 3 manifolds it is impossible to glue two harmonic spheres ω_0 and ω_1 which are so that ω_0 and $\omega_1(\frac{1}{z})$ parametrise transversally intersecting minimal surfaces, respectively transversally intersecting leafs of the same minimal surface, near $z = 0$.

Our result in this setting also applies for harmonic spheres which are degenerate critical points, i.e. which have Jacobi-fields that are not generated by variations of rational maps. Indeed, we can even deal with harmonic spheres ω_i which have non-integrable Jacobi-fields, i.e. for which $d^2E(\omega_i)$ has null directions that are not tangential to the set of harmonic maps near ω_i.

The one thing we want to ask is that for maps ω_i that are obtained as higher order coverings of a map U_i^* this structure is reflected also at the level of Jacobi-fields. Namely, if q_i^* has degree strictly greater than 1 then we ask that all Jacobi-fields along $\omega_i = U_i^* \circ q_i^*$ are generated by a combination of Jacobi-fields along U_i^* and Jacobi-fields that are generated by variations of the rational maps: that is, in this case we will assume that

$$\ker(L_{\omega_i}) = \{W \circ q_i^*, W \in \ker(L_{U_i^*})\} + T_{\omega_i}\{U_i^* \circ q : q : S^2 \to S^2 \text{ rational }\}, \tag{1.9}$$

L_u the Jacobi-operator along u. Here we note that the fact that $W \circ q^*$ is a Jacobi-field along $U^* \circ q^*$ whenever W is a Jacobi-field along U^* follows from the conformal invariance of the energy and the resulting formula

$$\tau(U \circ q) = \tfrac{1}{2}|\nabla q|^2 \tau(U \circ q) \tag{1.10}$$

for the transformation of the tension under conformal changes.

As our final main result we prove

Theorem 1.7. *Let (N, g_N) be any analytic 3 manifold and let $U^*_{0,1} : \hat{\mathbb{C}} \to N$ be any harmonic spheres with $U^*_1(0) = U^*_0(0)$ and $dU^*_{0,1}(0) \neq 0$ which are so that the tangent spaces $T_{U^*_i(0)} U^*_i(\hat{\mathbb{C}})$ to the corresponding minimal surfaces $U_i(\hat{\mathbb{C}})$ do not coincide. Let q^*_0, q^*_1 be any rational maps with $q^*_0(0) = q^*_1(0) = 0$, assumed to satisfy (1.9) if their degree is at least 2.*

Then for any sequence (u_j) of almost harmonic maps that converges to a bubble tree with base map $\omega_0 = U^*_0 \circ q^*_0$ and bubble $\omega_1(z) = U^*_1(q^*_1(\frac{1}{z}))$ as described in (1.1) and (1.4) we can bound the bubble scale by

$$\mu_j^{-\min(n^*_0, n^*_1)} \leq C \|dE(u_j)\|_*$$

and we have a Łojasiewicz-estimate of the form

$$|E(u) - [E(\omega_1) + E(\omega_0)]| \leq C \|dE(u_j)\|_*.$$

In particular, no sequence of harmonic maps can converge to such a bubble tree.

Remark 1.8. If $U^*_{0,1}$ are non-degenerate critical points, then the above theorem remains valid also without the assumption that (N, g_N) is analytic.

We note that in contrast to Theorems 1.3 and 1.5 which are valid for target manifolds of arbitrary dimension, here we have to impose that the target is 3 dimensional. We cannot expect to obtain the same repulsion effect if we were to consider maps into higher dimensional manifolds such as $N = S^2 \times S^2 \hookrightarrow \mathbb{R}^6$ where e.g. $u_j(\cdot) := (\pi(\cdot), \pi(\frac{1}{\mu_j \cdot})) : \mathbb{R}^2 \to N$, $\mu_j \to \infty$, are harmonic maps which converge to a bubble tree for which the tangent spaces of the base $\omega_0 = (\pi, \pi(\infty))$ and the bubble $\omega_1 = (\pi(\infty), \pi(\frac{1}{\cdot}))$ intersect transversally in a point.

Strategy of our proofs:
To prove our main results we use a general method that was first developed in the joint work [7] of Malchiodi, Sharp and the author on H-surfaces. A crucial aspect of this method is that it allows us to prove Łojasiewicz estimates for sequences of (almost) critical points of an energy E that converge to a singular limit, here a bubble tree, without having to analyse the properties of such general (almost) critical points.

Instead, if we can construct a suitable finite dimensional (non-compact) manifold $\mathcal{Z}$ of *singularity models* so that the restriction of the energy to $\mathcal{Z}$ has the right properties, then this method allows us to obtain Łojasiewicz-estimates that are valid on a *uniform* ε neighbourhood of $\mathcal{Z}$ based on the properties of the energy and its variation on $\mathcal{Z}$.

These singularity models will in general not be critical points of the energy, but will always be so that $\|dE(\mathfrak{z})\|_*$ is small as they serve as models for the behaviour of almost critical points. The key properties $\mathcal{Z}$ needs to satisfy are that

(i) The second variation of the energy is uniformly definite orthogonal to $\mathcal{Z}$
(ii) For each $\mathfrak{z} \in \mathcal{Z}$ which is not a critical point of E we can identify a direction on $\mathcal{Z}$ in which the variation of E has a sign and suitable scaling. To be more precise, we need that for each such $\mathfrak{z} \in \mathcal{Z}$ there exists a unit element $y_\mathfrak{z} \in T_\mathfrak{z}\mathcal{Z}$ so that

$$Q(\mathfrak{z}, y_\mathfrak{z}) := \frac{\|dE(u)\|_*\big(\|dE(u)\|_* + \|d^2E(u)(y_\mathfrak{z}, \cdot)\|_*\big)}{dE(u)(y_\mathfrak{z})} \text{ is small.} \tag{1.11}$$

Our proofs hence consist of three main steps: the construction of the manifold $\mathcal{Z}$ of singularity models, the analysis of the energy and its variations on $\mathcal{Z}$, including the identification of the dominating direction $y_\mathfrak{z}$ for which (1.11) holds, and finally the arguments of how these properties of E on $\mathcal{Z}$ yield our main results.

After rotating the domain, we will only need to consider maps for which the bubble is attached at $z = 0$ and in settings where the $U_{0,1}^*$ are non-degenerate, the singularity models will be obtained by carefully gluing maps of the form $U_i^*(q_0(z) + q_1(\frac{1}{\mu z}) + c_i)$, $i = 0, 1$, for $c_i \in \mathbb{C}$ small and q_i rational maps with $q_i(0) = 0$ which are close to q_i^*.

In case (1) where $U_0^* = U_1^*$, the set $\mathcal{Z}$ contains all relevant critical points (corresponding to the case where $c_0 = c_1$), but also non-harmonic maps $\mathfrak{z}$ which are obtained from gluing maps for which $c_0 \neq c_1$ and hence $U_0^*(c_0) \neq U_1^*(c_1)$. These need to be included as they correspond to directions in which the second variation degenerates as $\mu \to \infty$, compare Remark 4.3. For these non-harmonic elements of $\mathcal{Z}$ the dominating term in the energy expansion will always come from the zeroth order deviation $\delta := \mathrm{dist}(U_0^*(c_0), U_1^*(c_1))$ between the maps we glue. In case (1) we will hence always choose $y_\mathfrak{z} \in T_\mathfrak{z}\mathcal{Z}$ so that it corresponds to a variation that increases this number δ, compare Lemma 3.7.

For the other cases (2) and (3) our set of singularity models will *not* contain any critical points, and indeed a key part of our claims is the *non-existence* of any critical points which look like bubble trees of types (2) or (3). For these cases, the choice of $y_\mathfrak{z}$ will also be more involved. In case (2) both zeroth and first order deviations in the Taylor expansions of the maps that we glue can dominate the behaviour of the energy and its variation. As a result, the choice of direction will depend on the relative size of δ and (a suitable polynomial of) μ^{-1}, and we will consider variations that increase δ as described above only if δ dominates, and otherwise use variations that increase the energy by varying one of the rational maps in a suitable way, compare Lemma 3.5.

For settings as considered in Theorem 1.7 where the U_i^* are allowed to have non-integrable Jacobi-fields, we will need to expand the class of maps that we glue to obtain our singularity maps to include maps $U_i(q_0(z) + q_1(\frac{1}{\mu z}) + c_i)$ for certain U_i which are themselves not harmonic. As a result, in this case (3), the choice of $y_\mathfrak{z}$ will depend on the relation between δ, μ^{-1} and the tension of the maps U_i, and in the case where the later dominates we will consider a third type of variation which will increase $E(\mathfrak{z})$ by increasing $E(U_i)$, compare Lemma 3.8.

Outline of the paper:
The paper is organised as follows: We begin by explaining the precise construction of the singularity models in Section 2. We then collect all key properties of E and its variations on this set $\mathcal{Z}$ in a series of lemmas in the subsequent Section 3. We however postpone their rather technical proofs until Sections 5 and 6, as an understanding of these proofs is not required to complete the proofs of the main results which are carried out in Section 4.

2. Construction of the singularity models

In this section we construct the singularity models which are maps $\mathfrak{z}: S^2 \to N$ that model the behaviour of sequences of almost harmonic maps that converge to a bubble tree with base $\omega_0 = U_0^* \circ q_0^*$ and bubble $\omega_1(z) = U_1^* \circ q_1^*(\frac{1}{z})$ where here and in the following $U_i^* : \hat{\mathbb{C}} \to N$ always denote fixed harmonic maps with $dU_i^*(0) \neq 0$ while q_i^* denote fixed elements of

$$\mathcal{R} := \{q : \hat{\mathbb{C}} \to \hat{\mathbb{C}} \text{ rational with } q(0) = 0\}. \tag{2.1}$$

It suffices to construct singularity models $\mathfrak{z}: S^2 \to N$ for which the bubble is attached at the north pole and we denote the corresponding set of maps by $\mathcal{Z}_0$. Once we have constructed $\mathcal{Z}_0$ we can then obtain the full manifold of singularity models as

$$\mathcal{Z} := \{\mathfrak{z}_0 \circ R_p : p \in S^2,\, \mathfrak{z}_0 \in \mathcal{Z}_0\}, \tag{2.2}$$

where $R_p \in SO(3)$ denotes the rotation which maps the plane containing 0, p and the north pole $\pi(0)$ to itself and which is so that $R_p(p) = \pi(0)$ (with the convention that $R_p = \mathrm{Id}$ if $p = \pi(0)$).

To construct these maps $\mathfrak{z} \in \mathcal{Z}_0$ we work in fixed stereographic coordinates where the base and bubble will be represented by harmonic maps $U_i^* : \hat{\mathbb{C}} \to N$ and rational maps $q_i^* : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$. As discussed in the introduction we need to analyse the three cases where either $U_0^* = U_1^*$ or $U_0^*(\bar{z}) = U_1^*(z)$ or where $U_{0,1}^*$ parametrise minimal spheres in a 3 manifold that intersect transversally in $U_1^*(0) = U_0^*(0)$.

We can assume without loss of generality that the rational maps $q_{0,1}^*$ are so that the leading order coefficients in $q_i^*(z) = a_{n_i^*}(q_i^*)z^{n_i^*} + \sum_{j>n_i^*} a_j(q_i^*)z^j$ are so that $|a_{n_i^*}(q_i^*)| = 1$. Indeed, to ensure this for q_0^* we can simply replace U_0^* by $U_0^*(c\cdot)$ for a suitable $c > 0$, while replacing the chosen scales μ_i in the bubble tree convergence (1.4) by $c\mu_i$ allows us to assume that this holds true for q_1^* while still preserving the relations $U_0^* = U_1^*$ respectively $U_0^*(z) = U_1^*(\bar{z})$.

To obtain a manifold $\mathcal{Z}$ which is so that d^2E is definite orthogonal to $\mathcal{Z}$ we first need to determine suitable sets $\mathcal{H}_1^\sigma(U_i^*)$ of maps which are so that the tangent space to

$$\mathcal{H}^\sigma(\omega_i) := \{U \circ q : U \in \mathcal{H}_1^\sigma(U_i^*), \quad q \in \mathcal{R}^\sigma(q_i^*)\} \tag{2.3}$$

at $\omega_i = U_i^* \circ q_i^*$ coincides with the space $\ker(L_{\omega_i})$ of Jacobi-fields along ω_i. Here and in the following $\sigma > 0$ denotes a small constant that is chosen later and that is in particular small enough so that all maps in

$$\mathcal{R}^\sigma(q_i^*) := \{q \in \mathcal{R} : \|\pi \circ q \circ \pi^{-1} - \pi \circ q_i^* \circ \pi^{-1}\|_{C^k(S^2)} < \sigma\} \tag{2.4}$$

have the same degree as q_i^* and are so that $|a_{n_i^*}(q_i)| \in (\frac{1}{2}, 2)$. Here and in the following we can choose $k \geq 2$ to be any fixed number.

If ω_i is non-degenerate, i.e. if all Jacobi-fields are generated by variations of rational maps, then we can simply choose

$$\mathcal{H}_1^\sigma(U_i^*) := \{U_i^*(\cdot + c) : c \in \mathbb{C},\, |c| < \sigma\}.$$

More generally if all Jacobifields along U_i^* are integrable we first choose a manifold $\mathcal{H}_0(U_i^*)$ of harmonic maps of dimension $\dim(\ker L_{U_i^*}) - 6$ which is so that each harmonic map close to U_i^* can be written uniquely in the form $U \circ M$ for a Möbius transform M and an element U of $\mathcal{H}_0(U_i^*)$. We then set

$$\mathcal{H}_1^\sigma(U_i^*) := \{U(\cdot + c) : c \in \mathbb{C},\, |c| < \sigma,\, U \in \mathcal{H}_0(U_i^*) \text{ with } \|(U - U_i^*) \circ \pi^{-1}\|_{C^k(S^2)} < \sigma\} \tag{2.5}$$

and note that (1.9) ensures that $\mathcal{H}^\sigma(\omega_i)$ defined by (2.3) is indeed so that $T_{\omega_i}\mathcal{H}^\sigma(\omega_i) = \ker(L_{\omega_i})$.

In the more involved case where U_i^* has non-integrable Jacobi-fields we need to additionally include certain non-harmonic maps in $\mathcal{H}^\sigma(\omega_i)$, and hence in $\mathcal{H}_1^\sigma(U_i^*)$ and $\mathcal{H}_0(U_i^*)$, to ensure that also these non-integrable Jacobi-fields correspond to directions that are tangent to $\mathcal{H}^\sigma(\omega_i)$. In this case, we first choose $\mathcal{H}_0(U_i^*)$ as described in detail in Appendix A of [11] as a manifold of smooth maps with

$$T_{U_i^*}\{U \circ M : U \in \mathcal{H}_0(U_i^*), M \text{ Möbius transform}\} = \ker(L_{U_i^*}) \tag{2.6}$$

which has the key properties that for every $U \in \mathcal{H}_0(U_i^*)$ we can bound

$$\|\tau_{g_{S^2}}(U \circ \pi^{-1})\|_{C^k(S^2)} \leq C\|\tau_{g_{S^2}}(U \circ \pi^{-1})\|_{L^2(S^2, g_{S^2})} \tag{2.7}$$

and find a variation U_ε in $\mathcal{H}_0(U_i^*)$ of U with $\|\partial_\varepsilon(U_\varepsilon \circ \pi^{-1})\|_{C^k} \leq C$ for which

$$\tfrac{d}{d\varepsilon}|_{\varepsilon=0} E(U_\varepsilon) \geq \|\tau_{g_{S^2}}(U \circ \pi^{-1})\|_{L^2(S^2, g_{S^2})}, \tag{2.8}$$

see Lemma 2.1 of [11].

We then again define $\mathcal{H}_1^\sigma(U_i^*)$ by (2.5) and note that this manifold inherits the properties (2.7) and (2.8), while (1.9) ensures that $T_{\omega_i}\mathcal{H}^\sigma(\omega_i) = \ker(L_{\omega_i})$ for $\mathcal{H}^\sigma(\omega_i)$ defined by (2.3).

Having chosen $\mathcal{H}_1^\sigma(U_i^*)$ in this way we now obtain our singularity models by carefully gluing a highly concentrated copy of an element of $\mathcal{H}_1^\sigma(U_1^*)$ onto an element of $\mathcal{H}_1^\sigma(U_0^*)$. That is we want to construct the elements $\mathfrak{z} = \mathfrak{z}(U_0 \circ q_0, U_1 \circ q_1(\frac{1}{\mu \cdot}))$ of $\mathcal{Z}_0$ in a way that $\mathfrak{z} \approx U_0(q_0(z))$ away from $z = 0$ while $\mathfrak{z}(z) \approx U_1 \circ q_1(\frac{1}{\mu z})$ near $z = 0$, for $U_i \in \mathcal{H}_1^\sigma(U_i^*)$, $q_i \in \mathcal{R}^\sigma(q_i^*)$, $i = 0, 1$ and a large scaling factor μ.

Remark 2.1. We will always consider elements that are obtained from maps in $\mathcal{H}_1^\sigma(\omega_i)$ and scaling factors $\mu > \bar{\mu}$ for a sufficiently large number $\bar{\mu}$ and a sufficiently small number $\sigma > 0$ (depending only on U_i^*, q_i^*) and in the following all claims and estimates are to be understood to hold true after increasing $\bar{\mu}$ and decreasing σ if necessary.

To describe this gluing construction we first assign to each $q_0 \in \mathcal{R}^\sigma(q_0^*)$ and each $q_1(\frac{1}{\mu z})$ that we obtain from an element $q_1 \in \mathcal{R}^\sigma(q_1^*)$ and a scaling factor $\mu > \bar{\mu}$ the numbers

$$\mu_0 := |a_{n_0^*}(q_0)|^{-1/n_0^*} \text{ and } \mu_1 := \mu |a_{n_1^*}(q_1)|^{-1/n_1^*}. \tag{2.9}$$

We note that while there are multiple ways of representing the same rational map $q_1(\frac{1}{\mu z})$ using an element of $\mathcal{R}^\sigma(q_1^*)$ and a factor $\mu > \bar{\mu}$, the number μ_1 is uniquely determined by $q_1(\frac{1}{\mu z})$ and while μ_0 is of order 1, μ_1 will be of order μ.

Some of the gluing construction below will be carried out on annuli

$$A = \mathbb{D}_{r_0} \setminus \mathbb{D}_{r_1}, \quad A^* := (\mathbb{D}_{2r_1} \setminus \mathbb{D}_{r_1}) \cup (\mathbb{D}_{r_0} \setminus \mathbb{D}_{\frac{1}{2}r_0}) \text{ and } \hat{A} = \mathbb{D}_{2\hat{r}} \setminus \mathbb{D}_{\frac{1}{2}\hat{r}} \tag{2.10}$$

whose radii $\mu^{-1} \ll r_1 \ll \hat{r} \ll r_0 \ll 1$ are determined by $\mu_{0,1}$ defined in (2.9) via

$$r_0 := \mu_0 f_{\mu_1/\mu_0}^{-1}, \quad r_1 := \mu_1^{-1} f_{\mu_1/\mu_0} \text{ and } \hat{r} := \mu_0^{\frac{1}{2}} \mu_1^{-\frac{1}{2}} \quad . \tag{2.11}$$

Here $\mu \mapsto f_\mu$ is a fixed non-decreasing function which is so that

$$\mu |\partial_\mu f_\mu| \leq f_\mu \text{ while } \tfrac{\log f_\mu}{\log \mu} \leq \sigma_1 \text{ and } f_\mu^{-1} \leq \sigma_1 \tag{2.12}$$

for $\mu > \bar{\mu}$ and a small constant $\sigma_1 \in (0, \frac{1}{4}]$. We use the same convention for σ_1 as for σ, i.e. ask in the following that all claims hold provided f_μ satisfies this assumption for a sufficiently small number $\sigma_1 > 0$ that only depends on $U_{0,1}^*$ and $q_{0,1}^*$.

The precise choice of f_μ is not important in the construction and it will be useful later on that we can fix f_μ according to our needs as for the proof of Theorems 1.7 it will be convenient to work with $f_\mu = \log \mu$ while in the proof of Corollary 1.6 we will want to choose $f_\mu = \mu^{\sigma_1}$.

Remark 2.2. We note that the choice of the above radii and annuli is so that the construction is invariant under a rescaling $z \mapsto \lambda z$ of the stereographic coordinates and symmetric with respect to interchanging the roles of the base and the bubble as we change the coordinates according to $z \mapsto \frac{\mu_0}{\mu_1 z}$.

To construct our singularity models we will first modify the maps

$$u_1 := U_1 \circ q_\mu \text{ and } u_0 := U_0 \circ q_\mu \text{ for } q_\mu(z) := q_0(z) + q_1(\tfrac{1}{\mu z}) \tag{2.13}$$

on A so that they agree upto first order errors in $|z| + |\mu z|^{-1}$, then further modify the resulting maps on all of $\mathbb{R}^2$ so that they agree upto second order, before gluing these maps together on the central annulus $\hat{A}$ and projecting onto N.

To this end we let $\hat{\gamma} = \hat{\gamma}_{U_1(0),U_0(0)} : [0,1] \to N$ be the shortest geodesic from $\hat{\gamma}(0) = U_1(0)$ to $\hat{\gamma}(1) = U_0(0)$. We note that $\hat{\gamma}$ is well defined as $\delta(U_0, U_1) := \mathrm{dist}_N(U_0(0), U_1(0))$ is small for all elements of $\mathcal{H}^\sigma(U_i^*)$ as such maps are C^k close to U_i^* and as $U_0^*(0) = U_1^*(0)$.

To transition between $U_1(0)$ and $U_0(0)$ on $A = \mathbb{D}_{r_0} \setminus \mathbb{D}_{r_1}$ we use

$$\gamma := \hat{\gamma} \circ \phi_{r_0,r_1} \tag{2.14}$$

where $\phi_{r_0,r_1} = \phi_{r_0,r_1}(|x|) : \mathbb{R}^2 \to [0,1]$ is defined by

$$\phi_{r_0,r_1}(r) := \begin{cases} (1 - \phi(\frac{r}{r_1}))c_{r_0,r_1} \log(\frac{r}{r_1}) & \text{for } r \leq 2r_1 \\ c_{r_0,r_1} \log(\frac{r}{r_1}) = 1 - c_{r_0,r_1} \log(\frac{r_0}{r}) & \text{for } r \in (2r_1, \frac{1}{2}r_0) \\ 1 - \phi(\frac{2r}{r_0})c_{r_0,r_1} \log(\frac{r_0}{r}) & \text{for } r \geq \frac{1}{2}r_0 \end{cases} \tag{2.15}$$

for $c_{r_0,r_1} := (\log(\frac{r_0}{r_1}))^{-1}$ and a fixed cut-off function $\phi \in C_c^\infty([0,2),[0,1])$ with $\phi \equiv 1$ on $[0,1]$.

We note that ϕ_{r_0,r_1} is chosen so that $\Delta\phi_{r_0,r_1} \equiv 0$ outside of $A^* = (\mathbb{D}_{2r_1} \setminus \mathbb{D}_{r_1}) \cup (\mathbb{D}_{r_0} \setminus \mathbb{D}_{\frac{1}{2}r_0})$ and so that $\phi_{r_0,r_1}|_{\mathbb{D}_{r_1}} \equiv 0$ while $\phi_{r_0,r_1}|_{(\mathbb{D}_{r_0})^c} \equiv 1$. Hence $\gamma : \mathbb{R}^2 \to N$ is given by a harmonic map outside of A^* and transitions between $\gamma|_{\mathbb{D}_{r_1}} \equiv U_1(0)$ and $\gamma|_{\mathbb{D}^c_{r_0}} \equiv U_0(0)$. Setting

$$\tilde{\gamma}_{U_1} := \gamma - U_1(0) \text{ and } \tilde{\gamma}_{U_0} := \gamma - U_0(0) \tag{2.16}$$

we can hence change our original maps $u_{0,1}$ into maps $u_1(z) + \tilde{\gamma}_{U_1}(z)$ and $u_0(z) + \tilde{\gamma}_{U_0}(z)$ that agree upto error terms of order $O(|z| + \frac{1}{\mu|z|})$ on A. Next we set

$$\beta := U_0 - U_1, \quad j_0 := -d\beta(0)(q_{0^*}(\tfrac{1}{\mu z})) \text{ and } j_1 := d\beta(0)(q_{1^*}(z)) \tag{2.17}$$

where here and in the following we use the notation $0^* = 1$ and $1^* = 0$ to stress that to obtain the (first order) modification j_0 of the base map u_0 we use the rational map q_1 that corresponds to the bubble and vice versa. This ensures that the resulting maps

$$v_1(z) := u_1(z) + \tilde{\gamma}_{U_1}(z) + d\beta(0)(q_0(z)) = u_1(z) + \tilde{\gamma}_{U_1}(z) + j_1(z) \tag{2.18}$$

and

$$v_0(z) := u_0(z) + \tilde{\gamma}_{U_0}(z) - d\beta(0)(q_{1,\mu}(z)) = u_0(z) + \tilde{\gamma}_{U_0}(z) + j_0(z), \tag{2.19}$$

agree upto a second order error term $v_0 - v_1 = \beta(q_\mu) - \beta(0) - d\beta(0)(q_\mu)$ in $q_\mu = q_0 + q_{1,\mu}$.

We finally interpolate between these maps on the central annulus $\hat{A} = \mathbb{D}_{2\hat{r}} \setminus \mathbb{D}_{\frac{1}{2}\hat{r}}$ and project onto N. That is we define our singularity model as

$$\mathfrak{z} := \pi_N((1 - \varphi_{\hat{r}})v_1 + \varphi_{\hat{r}} v_0), \tag{2.20}$$

where we set $\varphi_{\hat{r}}(z) := \bar{\varphi}(\hat{r}^{-1}|z|)$ for a fixed smooth function $\bar{\varphi} : [0,\infty) \to [0,1]$ which is so that $\bar{\varphi}(t) = 0$ for $t \leq \frac{1}{2}$ while $\bar{\varphi}(t) = 1$ for $t \geq 2$.

We note that these maps are well defined for $U_{0,1} \in \mathcal{H}^\sigma(U^*_{0,1})$ and $q_{0,1} \in \mathcal{R}^\sigma(q^*_{0,1})$ as the image of $(1 - \varphi_{\hat{r}})v_1 + \varphi_{\hat{r}} v_0$ will be contained in a small tubular neighbourhood of N where the nearest point projection is well defined.

Remark 2.3. An important feature of our construction is that the modifications to the maps u_i are not just supported on an annulus, but rather on all of $\mathbb{R}^2$. In contrast to maps we would get from a more ad hoc gluing construction this means that $dE(\mathfrak{z})$ has a significant part that is supported on $\mathbb{D}_{r_1}$ and $\mathbb{D}^c_{r_0}$. This will be crucial in the proof of our main results later on as variations of elements of $\mathcal{Z}_0$ induced by variations of one of the rational maps are mainly concentrated on $\mathbb{D}_{r_1}$ respectively $\mathbb{D}^c_{r_0}$ and hence can only be used to obtain directions $y_\mathfrak{z} \in T_\mathfrak{z}\mathcal{Z}$ satisfying (1.11) if $dE(\mathfrak{z})$ is so that it interacts suitably with such directions.

3. Key lemmas

Here we collect the key properties on the behaviour of the energy and its variations on $\mathcal{Z}$ that we need to prove our main results. We postpone the proofs of these lemmas to Sections 5 and 6, as an understanding of these technical proofs is not required for the proofs of our main results.

As these estimates will all be invariant under precomposition with a rotation, it suffices to state and prove them for elements of the set $\mathcal{Z}_0$ of singularity models for which the bubble is attached at the north pole. We can furthermore assume that $q_1(\frac{1}{\mu z})$ is represented by a map q_1 for which $|a_{n_1^*}(q_1)| = 1$ and hence use that μ_1 and μ_0 defined in (2.9) are so that $\mu_1 = \mu$ while $\mu_0 \sim 1$.

Given $\mathfrak{z} = \mathfrak{z}(U_0 \circ q_0, U_1 \circ q_1(\frac{1}{\mu \cdot})) \in \mathcal{Z}_0$ we will work with respect to the weighted H^1-norm $\|\cdot\|_\mathfrak{z}$ and the corresponding inner product $\langle \cdot, \cdot \rangle_\mathfrak{z}$ which are given in stereographic coordinates by

$$\|w\|^2_\mathfrak{z} := \int_{\mathbb{R}^2} |\nabla w|^2 + \rho^2_\mathfrak{z} |w|^2 dx \text{ for } \rho^2_\mathfrak{z} := |\nabla \pi_{\mu_0(\mathfrak{z})^{-1}}|^2 + |\nabla \pi_{\mu_1(\mathfrak{z})}|^2, \tag{3.1}$$

$\pi_\lambda(z) := \pi(\lambda z)$. We note that this definition respects the symmetry of the construction, compare Remark 2.2, and that we define $\|\cdot\|_\mathfrak{z}$ for general elements $\mathfrak{z} = \mathfrak{z}_0 \circ R_p$ of $\mathcal{Z}$ by the analogue formula in stereographic coordinates centred at p.

We recall that the second variation of the energy is given by

$$d^2E(\mathfrak{z})(v,w) = \int \nabla v \nabla w - A(u)(\nabla u, \nabla u)A(u)(v,w) \text{ for } v,w \in \Gamma(\mathfrak{z}^*TN), \tag{3.2}$$

compare e.g. [11, Lemma 3.1], and note that d^2E is uniformly bounded with respect to this norm, i.e. so that

$$|d^2E(\mathfrak{z})(w,v)| \lesssim \|w\|_{\mathfrak{z}} \|v\|_{\mathfrak{z}} \text{ for all } v,w \in \Gamma(\mathfrak{z}^*TN), \tag{3.3}$$

compare Section 6. Here and in the following we write $T_1 \lesssim T_2$ if we can bound $T_1 \leq CT_2$ for a constant C that only depends on the limiting configuration, i.e. on $U_{0,1}^*$ and $q_{0,1}^*$ while we write $T_1 \sim T_2$ if both $T_1 \lesssim T_2$ and $T_2 \lesssim T_1$.

The first key property we need is that the choice of $\mathcal{H}_1^\sigma(U_{0,1}^*)$ in the construction of $\mathcal{Z}_0$ ensures that the second variation is uniformly definite orthogonal to $\mathcal{Z}$, i.e. that

Lemma 3.1. *There exist $c_0 > 0$ depending only on $U_{0,1}^*$ and $q_{0,1}^*$ so that for every $\mathfrak{z} \in \mathcal{Z}_0$ we can split $T_{\mathfrak{z}}^{\perp}\mathcal{Z}$ orthogonally into subspaces $\mathcal{V}_{\mathfrak{z}}^{\pm}$ which are so that*

$$\pm d^2E(\mathfrak{z})(w^\pm, w^\pm) \geq c_0 \|w^\pm\|_{\mathfrak{z}}^2 \text{ and } d^2E(\mathfrak{z})(w^+, w^-) = 0 \text{ for all } w^\pm \in \mathcal{V}_{\mathfrak{z}}^{\pm}. \tag{3.4}$$

To state all other properties of the energy E on $\mathcal{Z}$ we now associate to each $\mathfrak{z} = \mathfrak{z}(U_i, q_i, \mu)$ the quantities

$$\delta := \text{dist}_N(U_1(0), U_0(0)), \quad \mathcal{T} := \max(\mathcal{T}_0, \mathcal{T}_1) \text{ and } \bar{\nu} := \max(\nu_1, \nu_0) \tag{3.5}$$

where we set $\mathcal{T}_i := \|\tau_{g_{S^2}}(U_i \circ \pi^{-1})\|_{L^2(S^2, g_{S^2})}$ and write $q_i(z) = \sum_j a_j(q_i)z^j$ near $z = 0$ to define

$$\nu_1 := \max_{j=1,\dots,n_1^*} |a_j(q_1)|\mu^{-j}, \quad \nu_0 := \max_{j=1,\dots,n_0^*} |a_j(q_0)|\mu^{-j}, \tag{3.6}$$

$n_{0,1}^* \geq 1$ the order of the zero of the fixed rational maps $q_{0,1}^*$ at $z = 0$. We remark that these quantities respect the symmetry discussed in Remark 2.2 and recall furthermore that $\mathcal{T}_i$ also controls the C^k-norm of the tension of $U_i \circ \pi^{-1}$, compare (2.7).

The first quantitative estimate we need is a bound on the dual norm $\|dE(\mathfrak{z})\|_* := \sup\{dE(\mathfrak{z})(w) : \|w\|_{\mathfrak{z}} = 1\}$ of the first variation and in Section 6 we shall prove

Lemma 3.2. *For any $\mathfrak{z} \in \mathcal{Z}_0$ we have*

$$\|dE(\mathfrak{z})\|_* \lesssim \|\beta\|\bar{\nu}(\log \mu)^{\frac{1}{2}} + \tfrac{\delta(\log f_\mu)^{\frac{1}{2}}}{\log \mu} + \mathcal{T}.$$

for δ, $\mathcal{T}$ and $\bar{\nu}$ as in (3.5) and $\beta = U_0 - U_1$.

Here and in the following we compute the norms of $\beta = U_0 - U_1$ and $\partial_\varepsilon \beta$ as norms of the corresponding maps from S^2 i.e. set

$$\|\beta\| := \|\beta \circ \pi^{-1}\|_{C^2(S^2)} \text{ and } \|\partial_\varepsilon \beta\| := \|\partial_\varepsilon \beta \circ \pi^{-1}\|_{C^2(S^2)}.$$

In the following we consider variations $\mathfrak{z}_\varepsilon = \mathfrak{z}(U_{0,1}^{(\varepsilon)}, q_{0,1}^{(\varepsilon)}, \mu)$ in $\mathcal{Z}$ which are induced by variations of $U_i^{(\varepsilon)}$ in $\mathcal{H}_1^\sigma(U_i^*)$ and of $q_i^{(\varepsilon)}$ in $\mathcal{R}^\sigma(q_i^*)$. We do not impose the assumption that $|a_{n_i}^*(q_i^{(\varepsilon)})| = 1$ for $\varepsilon \neq 0$ but allow for general variations of q_i in $\mathcal{R}^\sigma(q_i)$ since this allows us to represent a variation of $\mu_\varepsilon = \mu(1+\varepsilon)$ through a variation of

$$q_1^{(\varepsilon)}(z) = q_1(\tfrac{1}{1+\varepsilon} z). \tag{3.7}$$

To simplify the notation we will drop the index ε whenever there is no room for confusion and use the convention that ∂_ε is evaluated at $\varepsilon = 0$ unless specified otherwise. We will only ever consider variations for which

$$\|\partial_\varepsilon(\pi \circ q_i \circ \pi^{-1})\|_{C^k(S^2)} \lesssim 1 \text{ and } \|\partial_\varepsilon U_i\|_{C^k(S^2)} \lesssim 1, \quad i = 0, 1 \tag{3.8}$$

as these correspond to variations in $\mathcal{Z}_0$ for which $\|\partial_\varepsilon \mathfrak{z}\|_{\mathfrak{z}} \lesssim 1$. We note that for such variations also $\|\partial_\varepsilon \mathfrak{z}\|_{L^\infty} \lesssim 1$ while $|\partial_\varepsilon \mu_i| \lesssim \mu_i$.

We associate to each such variation the quantities

$$\begin{aligned} &\eta_0 := |\partial_\varepsilon U_1(0)| + |\partial_\varepsilon U_0(0)| \text{ and} \\ &\eta_\tau := \|\partial_\varepsilon \tau_{g_{S^2}}(U_1 \circ \pi^{-1})\|_{C^0(S^2)} + \|\partial_\varepsilon \tau_{g_{S^2}}(U_0 \circ \pi^{-1})\|_{C^0(S^2)} \end{aligned} \tag{3.9}$$

and furthermore set $\eta_{rat} = 0$ whenever $\partial_\varepsilon q_0 = \partial_\varepsilon q_1 = 0$ while otherwise we let $\eta_{rat} := 1$ and choose $j^* \geq 1$ so that both $\partial_\varepsilon q_0$ and $\partial_\varepsilon q_1$ are of order $O(z^{j^*})$ near $z = 0$.

The dual norm $\|d^2E(\mathfrak{z})(\partial_\varepsilon \mathfrak{z}, \cdot)\|_* := \sup\{\partial_\varepsilon dE(\mathfrak{z}_\varepsilon)(w) : w \in \Gamma(\mathfrak{z}^*TN), \|w\|_{\mathfrak{z}} = 1\}$ of the second variation of E along such variations is controlled by the following lemma which we will prove in Section 6.

Lemma 3.3. *For any variation $\mathfrak{z}_\varepsilon$ for which (3.8) holds we can bound*

$$\|d^2E(\mathfrak{z})(\partial_\varepsilon \mathfrak{z}, \cdot)\|_* \lesssim \left[(\|\beta\| + \|\partial_\varepsilon \beta\|)\bar{\nu} + \eta_{rat}\mu^{-j^*}\right] \log(\mu)^{\frac{1}{2}} + \tfrac{(\delta+\eta_0)(\log f_\mu)^{\frac{1}{2}}}{\log \mu} + \mathcal{T} + \eta_{\mathcal{T}}, \tag{3.10}$$

for δ, $\mathcal{T}$ and $\bar{\nu}$ as in (3.5), $\beta = U_0 - U_1$ and $\eta_\cdot$ as in (3.9).

Remark. For changes of U_i in $\mathcal{H}_1^\sigma(U_i^*) = \{U(\cdot + c), U \in \mathcal{H}_0(U_i^*), |c| \leq \sigma\}$ that are induced only by translations on the domain, rather than also a change of the map in $\mathcal{H}_0(U_i^*)$, we can bound $\eta_\tau \lesssim \mathcal{T}$ since $\mathcal{T}_i$ controls the C^1 norm of the tension of U_i, compare (2.7).

For more general variations in $\mathcal{H}_1^\sigma(U_i^*)$ we can use that $\mathcal{H}_0(U_i^*)$ is chosen in a way that all elements of $T_{U_i^*}\mathcal{H}_0(U_i^*)$ are Jacobi-fields along U_i^*, and hence so that

$P_{U_i^*\circ\pi^{-1}}(\partial_\varepsilon \tau_{g_{S^2}}(U_\varepsilon \circ \pi^{-1})) = 0$ for all variation U_ε of U_i^* in $\mathcal{H}_0(U_i^*)$. As $\mathcal{H}_1^\sigma(U_i^*)$ is contained in a small C^k neighbourhood of U_i^* and as the tension is tangential, and hence

$$\begin{aligned}\eta_{\mathcal{T}} &= \sum_i \|\partial_\varepsilon\big(P_{U_{i,\varepsilon}\circ\pi^{-1}}(\tau_{g_{S^2}}(U_{i,\varepsilon}\circ\pi^{-1})\big)\|_{C^0(S^2)} \\ &\lesssim \mathcal{T} + \sum_i \|P_{U_i\circ\pi^{-1}}(\partial_\varepsilon \tau_{g_{S^2}}(U_{i,\varepsilon}\circ\pi^{-1}))\|_{C^0(S^2)}\end{aligned}$$

this ensures that $\eta_{\mathcal{T}}$ will be smaller than any given positive constant if $\sigma > 0$ is chosen sufficiently small.

To analyse the first variation of E along $\mathcal{Z}$ we will in particular use

Lemma 3.4. *Let $U_{0,1}^*$ be any harmonic maps from S^2 into a smooth Riemannian manifold N of any dimension, let $q_{0,1}^* : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be any rational maps with $q_{0,1}^*(0) = 0$ and let $\mathcal{Z}_0$ be the corresponding set of singularity models defined in Section 2. Then we have*

$$dE(\mathfrak{z})(\partial_\varepsilon \mathfrak{z}) = \sum_{i=0,1} \int_{\Omega_i} j_i \partial_\varepsilon \Delta u_i + \pi c_{r_0,r_1} \tfrac{d}{d\varepsilon}\delta(U_0, U_1)^2 + \sum_{i=0,1} \deg(q_i) \tfrac{d}{d\varepsilon} E(U_i) + err \tag{3.11}$$

for any variation $\mathfrak{z}_\varepsilon$ in $\mathcal{Z}_0$ which satisfies (3.8). Here u_i, j_i and γ are as defined in (2.13), (2.17) and (2.14), $\Omega_1 := \mathbb{D}_{\hat{r}} = \Omega_0^c$, $c_{r_0,r_1} = (\log(\frac{r_0}{r_1}))^{-1}$ and err denotes an error term that is bounded by $|err| \lesssim R_0$ for

$$\begin{aligned}R_0 :=&\bar{\nu}\|\beta\|\big[\mu^{-\frac{1}{2}} + \eta_0 + \eta_{\mathcal{T}}\big] + \tfrac{\delta}{\log\mu}\big[\delta + f_\mu{}^{-2} + \eta_0[(\log\mu)^{-1} + f_\mu{}^{-1}]\big] \\ &+ \mathcal{T}\big[\mu^{-1}\log\mu + \eta_0(\log\mu)^{-1} f_\mu{}^{-1}\big],\end{aligned} \tag{3.12}$$

δ, $\mathcal{T}$ and $\bar{\nu}$ as in (3.5), $\beta = U_0 - U_1$ and $\eta_\cdot$ as in (3.9).

As $c_{r_0,r_1} \sim (\log\mu)^{-1}$ and as we shall see that $|\int_{\Omega_i} j_i \partial_\varepsilon \Delta u_i| \lesssim \|\beta\|(\nu_{i^*} + \mu^{-\frac{1}{2}}\bar{\nu})$ for $\nu_{1^*} = \nu_0$ and $\nu_{0^*} = \nu_1$ this ensures in particular that

$$|\partial_\varepsilon E(\mathfrak{z}_\varepsilon)| \lesssim \|\beta\|\bar{\nu} + \tfrac{\delta}{\log\mu} + \mathcal{T} \tag{3.13}$$

for any variation $\mathfrak{z}_\varepsilon$ in $\mathcal{Z}$ for which $\|\partial_\varepsilon \mathfrak{z}\| \lesssim 1$.

In the following we will consider three types of variations, which geometrically correspond to

(V1) Variations which do not change the surfaces $U_i(\hat{\mathbb{C}})$, but move one of the points $U_i(0)$ along $U_i(\hat{\mathbb{C}})$. Such variations will be obtained by considering translations on the domain.

(V2) Variations which do not change the surfaces $U_i(\hat{\mathbb{C}}) \subset N$, nor the points $U_i(0)$ of these surfaces where we glue. Such variations will be obtained from varying one of the rational maps q_i in a way that preserves the condition $q_i(0) = 0$.

(V3) Variations which change one of the surface $U_i(\hat{\mathbb{C}})$.

Roughly speaking we will want to consider variations of type (V1) if the influence of δ dominates, of type (V2) if the influence of $\|\beta\|\bar{\nu}$ dominates and of type (V3) if tension terms $\mathcal{T}$ dominate.

For variations of type (V1) the third term in (3.11) vanishes and $\eta_{rat} = 0$ while $\eta_\tau \lesssim \mathcal{T}$ so we will have

$$|\partial_\varepsilon E(\mathfrak{z}_\varepsilon)| \lesssim \bar{\nu} + \tfrac{\delta}{\log \mu} + \tfrac{\mathcal{T}}{\log \mu}. \tag{3.14}$$

For variations of type (V2) both the second and the third term in (3.11) vanish and $\eta_0 = \eta_\tau = 0$, so we get a further improvement of this bound. In particular, for variations induced by only a change of q_1 we will see that

$$|\partial_\varepsilon E(\mathfrak{z}_\varepsilon)| \lesssim \|\beta\|\nu_0 + \|\beta\|\nu_1\mu^{-\frac{1}{2}} + \tfrac{\delta^2}{\log \mu} + \tfrac{\delta}{f_\mu^2 \log \mu} + \mathcal{T}\mu^{-1}\log \mu \tag{3.15}$$

and we note that the same bound also applies to $\mu|\partial_\mu E(\mathfrak{z}(U_{0,1}, q_{0,1}, \mu)|$ since variations of μ can be represented by variations of q_1 as described in (3.7).

To prove our main results we now want to identify variations for which $dE(\mathfrak{z})(\partial_\varepsilon \mathfrak{z})$ scales like the dominating term in the above estimates.

To state these lemmas in a way that makes them applicable in the proofs of all of our main results we recall that we always deal with settings for which the following assumption is satisfied.

Assumption 1. We ask that $U^*_{0,1} : \hat{\mathbb{C}} \to N$ are harmonic maps with $U^*_0(0) = U^*_1(0)$ and $dU^*_{0,1}(0) \neq 0$ for which one of the following holds

(A1) We have $U^*_0 = U^*_1$ and this map is a non-degenerate critical point;

(A2) We have $U^*_0(z) = U^*_1(\bar{z})$ and these maps are non-degenerate critical points;

(A3) The target manifold N is 3 dimensional and the tangent spaces $T_{U^*_i(0)}U^*_i(\hat{\mathbb{C}})$ do not coincide.

We recall that the assumption (A1) corresponds to settings as considered in Theorem 1.3 and Corollary 1.4, (A2) to Theorem 1.5 and Corollary 1.6 and (A3) to Theorem 1.7 and also note that $\mathcal{T} = 0$ for both (A1) and (A2) and that will see that $\|\beta\| \lesssim \delta$ for settings satisfying (A1).

In practice this will mean that for settings of type (A1) we will only ever consider variations of type (V1), that for settings satisfying (A2) we will consider variations of either type (V1) or (V2), while for settings as in (A3) all three types of variations will be relevant.

In Section 5.3 we shall prove that we can choose variations of type (V2) for which the following holds.

Lemma 3.5. *Let $U_{0,1}^*$ be so that either (A2) or (A3) of Assumption 1 holds, let $q_{0,1}^* \in \mathcal{R}$ and let $\mathcal{Z}_0$ be the corresponding set of singularity models defined in Section 2.*

Then there exists a constant $c_1 > 0$ so that for every $\mathfrak{z} \in \mathcal{Z}_0$ and each $i \in \{0,1\}$ there is a variation $\mathfrak{z}_\varepsilon$ that is induced by a variation of q_i which satisfies (3.8) for which

$$dE(\mathfrak{z})(\partial_\varepsilon \mathfrak{z}_\varepsilon) \geq c_1 |d\beta(0)|^2 \nu_{i^*} - C\big[\|\beta\|\bar{\nu}\mu^{-\frac{1}{2}} + \tfrac{\delta^2}{\log\mu} + \tfrac{\delta}{f_\mu^2 \log\mu} + \mathcal{T}\mu^{-1}\log\mu\big], \tag{3.16}$$

where we write for short $\nu_{0^*} = \nu_1$ and $\nu_{1^*} = \nu_0$, recall that f_μ satisfies (2.12) and that δ, $\mathcal{T}$, ν. and $\bar{\nu}$ are as in (3.5) and (3.6) while $\beta = U_0 - U_1$.

To obtain variations of type (V1) which change the distance between the points $U_i(0)$ that we glue, we will be able to use

Lemma 3.6. *Let U_i^* be as in Assumption 1 and let $U_{0,1} \in \mathcal{H}_1^\sigma(U_{0,1}^*)$. Then there exists a variation of the form $U_{i,\varepsilon} = U_i(\cdot + \varepsilon a)$, $a \in \mathbb{R}^2$ with $|a| \lesssim 1$, so that*

$$\tfrac{d}{d\varepsilon}\delta(U_0, U_1)^2 = \delta(U_0, U_1), \tag{3.17}$$

where for settings satisfying (A1) or (A2), we can freely choose which of U_0 or U_1 we want to vary, while for settings satisfying (A3) the above holds true for at least one of $i \in \{0,1\}$.

For the resulting variations of type (V1) we then obtain

Lemma 3.7. *Let $U_{0,1}^*$ be as in Assumption 1, let $q_{0,1}^*$ be any maps in $\mathcal{R}$ and let $\mathcal{Z}_0$ be the corresponding set of singularity models defined in Section 2. Then*

$$dE(\mathfrak{z})(\partial_\varepsilon \mathfrak{z}) \geq \tfrac{\delta}{\log\mu} - C\big(\|\beta\|\bar{\nu} + \mathcal{T}(\log\mu)^{-1} f_\mu^{-1}\big) \tag{3.18}$$

for the variation $\mathfrak{z}_\varepsilon$ in $\mathcal{Z}_0$ that is induced by changing U_i as described in Lemma 3.6, and for δ, $\mathcal{T}$ and $\bar{\nu}$ as in (3.5) and f_μ as in (2.12).

Finally, in non-integrable settings where we have to deal with maps U_i which are not harmonic, we also want to consider variations of type (V3) which increase $E(U_i)$ and hence show

Lemma 3.8. *Let $U_{0,1}^* : S^2 \to N$ be harmonic maps into an analytic Riemannian manifold of any dimension with $U_0^*(0) = U_1^*(0)$. Then for any $\mathfrak{z} \in \mathcal{Z}_0$ there exists a variation $\mathfrak{z}_\varepsilon$ in $\mathcal{Z}_0$ induced by a variation of U_i in $\mathcal{H}_1^\sigma(U_i^*)$ which satisfies (3.8) so that*

$$dE(\mathfrak{z})(\partial_\varepsilon \mathfrak{z}) \geq \mathcal{T}_i - C\big[\|\beta\|\bar{\nu} + \tfrac{\delta}{\log\mu} + f_\mu^{-1}(\log\mu)^{-1}\mathcal{T}\big], \tag{3.19}$$

δ, $\mathcal{T}$. and $\bar{\nu}$ as in (3.5) and f_μ as in (2.12).

We note that while the previous Lemmas 3.5, 3.6 and 3.7 require Assumption 1, this Lemma 3.8 is valid for more general settings, though in the present paper will only be applied for settings satisfying (A3).

In the next section we will use these three lemmas to identify a unit direction $y_{\mathfrak{z}} = \frac{\partial_\varepsilon \mathfrak{z}_\varepsilon}{\|\partial_\varepsilon \mathfrak{z}_\varepsilon\|}$ in which $dE(\mathfrak{z})$ is not only guaranteed to be positive, but also so that it dominates all other terms that appear when we analyse the behaviour of E in a neighbourhood of $\mathcal{Z}$, compare (1.11).

4. Proof of the main theorems

Here we give the proof of our main results based on the lemmas stated in the previous section. So let (u_j) be a sequence of almost harmonic maps from S^2 to N that converges to a bubble tree with bubble $\omega_1(z) = U_1^*(q_1^*(\frac{1}{z}))$ and base $\omega_0 = U_0^* \circ q_0^*$ and let $\mathcal{Z} = \{\mathfrak{z} : S^2 \to N\}$ be the corresponding set of singularity models constructed in Section 2.

For sufficiently large j our construction and (1.4) ensure that the function $\mathfrak{z} \mapsto \|u_j - \mathfrak{z}\|_{\mathfrak{z}}$ achieves its minimum on $\mathcal{Z}$ and that the corresponding $\mathfrak{z}_j \in \mathcal{Z}$ with $\|u_j - \mathfrak{z}_j\|_{\mathfrak{z}_j} = \mathrm{dist}(u_j, \mathcal{Z}) := \inf_{z \in \mathcal{Z}} \|u_j - \mathfrak{z}\|_{\mathfrak{z}}$ are obtained from maps $U_{0,1}^{(j)} \to U_{0,1}^*$, rational functions $q_{0,1}^{(j)} \to q_{0,1}^*$ and scales $\mu(\mathfrak{z}_j) \to \infty$ that are comparable to the numbers μ_j from (1.4). Furthermore, (1.4) can be used to see that not only $\|u_j - \mathfrak{z}_j\|_{\mathfrak{z}_j} \to 0$ but also $\|u_j - \mathfrak{z}_j\|_{L^\infty(S^2)} \to 0$.

It hence suffices to prove that there exists a constant $\varepsilon > 0$ so that the claims of our main results hold true for all maps u for which

$$\|u - \mathfrak{z}\|_{\mathfrak{z}} = \mathrm{dist}(u, \mathcal{Z}) < \varepsilon \text{ and } \|u - \mathfrak{z}\|_{L^\infty(S^2)} < \varepsilon \text{ for some } \mathfrak{z} \in \mathcal{Z}, \tag{4.1}$$

where after a rotation of the domain we can assume without loss of generality that u is so that the above holds for a $\mathfrak{z} \in \mathcal{Z}_0$.

Here and in the following we continue to use the convention that all claims are to be understood to hold true for the set of singularity models $\mathcal{Z}$ obtained for sufficiently small $\sigma, \sigma_1 > 0$ and sufficiently large $\bar{\mu}$.

Remark 4.1. Given any u for which (4.1) holds we use the convention that

$$\|dE(u)\|_* := \sup\{dE(u)(w) : \|w\|_{\mathfrak{z}} = 1\} \tag{4.2}$$

is to be computed with respect to the weighted norm $\|\cdot\|_{\mathfrak{z}}$ defined in (3.1) for a $\mathfrak{z} \in \mathcal{Z}$ for which (4.1) holds.

We do not claim that $\mathfrak{z}$ is uniquely determined by this relation but observe that $\|\cdot\|_{\tilde{\mathfrak{z}}} \sim \|\cdot\|_{\mathfrak{z}}$ if (4.1) is satisfied for both $\mathfrak{z}$ and $\tilde{\mathfrak{z}}$ and hence that it does not matter which such element of $\mathcal{Z}$ we use to define $\|dE(u)\|_*$ in our main results.

We will combine the results from the previous section with the following lemma that describes the basic estimates that we can obtain from our method of proof, compare also [7, Section 2].

Lemma 4.2. *Let $U^*_{0,1} : S^2 \to N$ be any given harmonic maps into a smooth Riemannian manifold, let $q^*_{0,1} \in \mathcal{R}$ and let $\mathcal{Z}$ be the corresponding set of singularity models. Then there exist constants $\varepsilon > 0$, $\bar{\Lambda} < \infty$ and C so that the following holds true for any $u \in H^1(S^2, N)$ with $\|\tau_{g_{S^2}}(u)\|_{L^2(S^2,g_{S^2})} \leq 1$ for which there exists $\mathfrak{z} \in \mathcal{Z}$ so that (4.1) is satisfied.*

We can bound $w := u - \mathfrak{z}$ by

$$\|w\|_{\mathfrak{z}} \leq C\|dE(u) - dE(\mathfrak{z})\|_* \tag{4.3}$$

and furthermore get that

$$dE(\mathfrak{z})(y_{\mathfrak{z}}) \leq C|dE(u)(y_{\mathfrak{z}})| + C\|dE(u)\|_*^2 \tag{4.4}$$

for any $u, \mathfrak{z}$ for which there is a unit element $y_{\mathfrak{z}} \in T_{\mathfrak{z}}\mathcal{Z}$ with

$$Q(\mathfrak{z}, y_{\mathfrak{z}}) := \frac{\|dE(\mathfrak{z})\|_*^2 + \|d^2E(\mathfrak{z})(y_{\mathfrak{z}}, \cdot)\|_* \|dE(\mathfrak{z})\|_*}{dE(\mathfrak{z})(y_{\mathfrak{z}})} \leq \bar{\Lambda}^{-1}. \tag{4.5}$$

For C^2-energies that are defined on a fixed Hilbert-space, the analogue of the above lemma can be obtained from a simple argument that is based on the fundamental theorem of calculus, compare [7]. In more involved settings such as the present problem where we consider an energy on a suitable Hilbert-manifold of maps and where we work with distances that are calculated using a weighted norm that depends on the corresponding element of $\mathcal{Z}$ one needs to proceed with more care. For almost harmonic maps from higher genus surfaces which are close to a simple bubble tree a similar result was obtained by the author in Section 4 of [11], and most of the arguments from [11] are applicable also in the present situation. For the convenience of the reader we hence include a sketch of the proof of this lemma, but omit the technical details except to explain why the estimate (4.3) is simpler than the one obtained in [11, Lemma 4.4].

Sketch of the proof of Lemma 4.2. Given u, $\mathfrak{z}$ and $w = u - \mathfrak{z}$ as in the lemma we set $w_{\mathfrak{z}} := P_{\mathfrak{z}} w$, P_p the projection onto T_pN. We first note that (4.1) ensures that $\|w - w_{\mathfrak{z}}\|_{\mathfrak{z}} \leq C\|w\|_{L^\infty}\|w\|_{\mathfrak{z}}$ where and in the following all L^p norms are computed over (S^2, g_{S^2}) unless indicated otherwise.

We can easily check that the variation of the weight $\rho_{\mathfrak{z}}$ in (3.1) is controlled by $\|\partial_\varepsilon \rho_{\mathfrak{z}}\|_{L^2} \lesssim 1$ whenever $\|\partial_\varepsilon \mathfrak{z}\|_{\mathfrak{z}} \leq 1$. As $\mathfrak{z}$ minimises the distance to u and thus

$$0 = \tfrac{d}{d\varepsilon}|_{\varepsilon=0}\|u - \mathfrak{z}\|_{\mathfrak{z}}^2 = -2\langle u - \mathfrak{z}, \partial_\varepsilon \mathfrak{z}\rangle_{\mathfrak{z}} + 2\int \rho_{\mathfrak{z}} \partial_\varepsilon \rho_{\mathfrak{z}} |w|^2$$

we hence get that $\|P^{T_\mathfrak{z}\mathcal{Z}} w_\mathfrak{z}\|_\mathfrak{z} \leq C\|w\|_{L^\infty}\|w\|_\mathfrak{z}$, i.e. that $w_\mathfrak{z}$ is nearly orthogonal to $T_\mathfrak{z}\mathcal{Z}$.

We then set $\tilde{w}_\mathfrak{z} := P^{\mathcal{V}_\mathfrak{z}^+} w_\mathfrak{z} - P^{\mathcal{V}_\mathfrak{z}^-} w_\mathfrak{z}$, where $P^{\mathcal{V}_\mathfrak{z}^\pm} : \Gamma(\mathfrak{z}^*TN) \to \mathcal{V}_\mathfrak{z}^\pm$ denote the $\langle \cdot, \cdot \rangle_\mathfrak{z}$ orthogonal projections onto the subspaces $\mathcal{V}_\mathfrak{z}^\pm$ of $T_\mathfrak{z}^\perp \mathcal{Z}$ from Lemma 3.3. The uniform definiteness of d^2E on these spaces ensures that

$$(c_0 - C\|w\|_{L^\infty})\|w\|_\mathfrak{z}^2 \leq d^2E(\mathfrak{z})(w_\mathfrak{z}, \tilde{w}_\mathfrak{z}) \leq (dE(u) - dE(\mathfrak{z}))(\tilde{w}_\mathfrak{z}) + \mathrm{err}_1, \tag{4.6}$$

where we can analyse the resulting error term as explained in detail in the proof of Lemma 4.4 of [11], resulting in an estimate of

$$|\mathrm{err}_1| \lesssim \|w\|_{L^\infty}\|w\|_\mathfrak{z}^2 + \|w\|_{L^\infty}\|\tau_{g_{S^2}}(u)\|_{L^2}\|\tilde{w}_\mathfrak{z}\|_{L^4}^2. \tag{4.7}$$

The only difference to the proof in [11] is that in the present situation

$$\|\tilde{w}_\mathfrak{z}\|_{L^4} \lesssim \|\tilde{w}_\mathfrak{z}\|_{H^1} \lesssim \|\tilde{w}_\mathfrak{z}\|_\mathfrak{z} \leq \|w_\mathfrak{z}\|_\mathfrak{z} \lesssim \|w\|_\mathfrak{z}$$

and that we can hence bound the second term on the right hand side of (4.7) by a multiple of the first term. This was not possible in [11], as the correct weighted norm in that context does not control the H^1 norm in a uniform way.

For $\varepsilon > 0$ sufficiently small we hence conclude that

$$\|w\|_\mathfrak{z}^2 \leq C(dE(u) - dE(\mathfrak{z}))(\tilde{w}_\mathfrak{z}) \tag{4.8}$$

which yields the first claim of the lemma.

We can now use that for unit elements $y_\mathfrak{z}$ we have

$$dE(\mathfrak{z})(y_\mathfrak{z}) = dE(u)(y_\mathfrak{z}) + d^2E(\mathfrak{z})(y_\mathfrak{z}, w_\mathfrak{z}) + \mathrm{err}_2$$

for an error of order $|\mathrm{err}_2| \lesssim \|w\|_\mathfrak{z}^2$. If we can choose such a $y_\mathfrak{z}$ so that it satisfies (4.5) we can thus conclude that

$$\begin{aligned} dE(\mathfrak{z})(y_\mathfrak{z}) &\leq |dE(u)(y_\mathfrak{z})| + |d^2E(\mathfrak{z})(y_\mathfrak{z}, w_\mathfrak{z})| + C\|dE(u)\|_*^2 + C\|dE(\mathfrak{z})\|_*^2 \\ &\leq C|dE(u)(y_\mathfrak{z})| + C\|dE(u)\|_*^2 + C\bar{\Lambda}^{-1} dE(\mathfrak{z})(y_\mathfrak{z}) \end{aligned} \tag{4.9}$$

for constants C that only depend on U_i^* and q_i^*. For $\bar{\Lambda} = \bar{\Lambda}(U_i^*, q_i^*)$ sufficiently large we can absorb the last term into the left hand side, which yields the second claim of the lemma. □

To prove our main results we will now apply the above lemma in directions $y_\mathfrak{z}$ that correspond to variations $\mathfrak{z}_\varepsilon$ as considered in one of the three Lemmas 3.5, 3.7, 3.8.

4.1. Proofs of Theorem 1.3 and Corollary 1.4

We first consider settings in which the base and bubble are given by parametrisations of the same non-degenerate minimal sphere with the same orientation.

So let $U^*: \mathbb{R}^2 \to N$, $q_{0,1}^* \in \mathcal{R}$ and $\omega_{0,1}$ be as in Theorem 1.3, let $\mathcal{Z}$ be the corresponding set of singularity models and let u be a map for which (4.1) holds true for $\varepsilon > 0$ chosen as in Lemma 4.2 and $\mathfrak{z} \in \mathcal{Z}_0$.

In this case, which corresponds to (A1), we can always obtain the desired $y_\mathfrak{z}$ by considering variations $\mathfrak{z}_\varepsilon$ of type (V1) for which U_0, $q_{0,1}$ and μ remain fixed while U_1 varies according to $U_{1,\varepsilon} = U_1(\cdot + \varepsilon a)$ for $|a| \sim 1$ chosen as in Lemma 3.6. As $\|\partial_\varepsilon \mathfrak{z}_\varepsilon\|_\mathfrak{z} \sim 1$ we obtain from Lemma 3.7 that

$$dE(\mathfrak{z})(y_\mathfrak{z}) \geq 2c\tfrac{\delta}{\log\mu} - C\|\beta\|\bar{\nu} \tag{4.10}$$

for the corresponding unit element $y_\mathfrak{z} = \frac{\partial_\varepsilon \mathfrak{z}_\varepsilon}{\|\partial_\varepsilon \mathfrak{z}_\varepsilon\|_\mathfrak{z}}$ and for constants $C, c > 0$ that only depend on U^* and $q_{0,1}^*$. Here we use that $\mathcal{T} = 0$ as the maps U_i are given by $U_i(\cdot) = U^*(\cdot + c_i)$ for some $|c_i| < \sigma$ since U^* is assumed to be non-degenerate. As $dU^*(0) \neq 0$ this furthermore implies that $\delta \sim \|\beta\|$ as both of these quantities scale like $|c_0 - c_1|$. As $\bar{\nu} \lesssim \mu^{-1} \ll (\log\mu)^{-1}$ we hence deduce from (4.10) that

$$dE(\mathfrak{z})(y_\mathfrak{z}) \geq c\tfrac{\delta}{\log\mu}$$

while Lemmas 3.2 and 3.3 ensure that

$$\|dE(\mathfrak{z})\|_* \leq C\tfrac{\delta(\log f_\mu)^{\frac{1}{2}}}{\log\mu} \text{ and } \|d^2E(\mathfrak{z})(y_\mathfrak{z}, \cdot)\|_* \leq C\tfrac{(\log f_\mu)^{\frac{1}{2}}}{\log\mu}, \tag{4.11}$$

again for constants that only depend on the limiting configuration.

We hence deduce that

$$Q(\mathfrak{z}, y_\mathfrak{z}) \leq C\tfrac{\log f_\mu}{\log\mu} \leq C\sigma_1 \leq \bar{\Lambda}^{-1},$$

$\bar{\Lambda}$ as in Lemma 4.2 since f_μ satisfies (2.12) for a constant $\sigma_1 = \sigma(U_i^*, q_i^*) > 0$ that can still be reduced if necessary. This lemma hence implies that

$$\|w\|_\mathfrak{z} \lesssim \|dE(u)\|_* + \|dE(\mathfrak{z})\|_* \tag{4.12}$$

and

$$\tfrac{\delta}{\log\mu} \leq dE(u)(y_\mathfrak{z}) + C\|dE(u)\|_*^2. \tag{4.13}$$

If u is harmonic, then (4.13) ensures that $\delta(\mathfrak{z}) = 0$ which implies that $U_0 = U_1$ and thus that $\mathfrak{z}(z) = U^*(q_0(z) + q_1(\frac{1}{\mu z}) + c_0)$. In this case we hence have that $dE(\mathfrak{z}) = 0$ so

(4.12) ensures that indeed $u = \mathfrak{z} = U^*(q_0(z) + q_1(\frac{1}{\mu z}) + c_0)$, establishing the claim about *harmonic* maps made in the theorem.

To prove the claims about almost harmonic maps made in the theorem and the subsequent Corollary 1.4 we now show that

$$|E(u) - E^*| \lesssim \tfrac{\delta^2}{\log \mu} + \|dE(u)\|_*^2, \tag{4.14}$$

where we write for short $E^* = E(\omega_0) + E(\omega_1) = [\deg(q_0^*) + \deg(q_1^*)]E(U^*)$.

To see this we first use (4.11) and (4.12) to bound

$$\begin{aligned}|E(u) - E(\mathfrak{z})| &\lesssim |dE(\mathfrak{z})(w)| + \|w\|_{\mathfrak{z}}^2 \lesssim \|dE(\mathfrak{z})\|_*^2 + \|w\|_{\mathfrak{z}}^2 \lesssim \|dE(\mathfrak{z})\|_*^2 + \|dE(u)\|_*^2 \\ &\lesssim \tfrac{\delta^2 \log f_\mu}{(\log \mu)^2} + \|dE(u)\|_*^2 \lesssim \tfrac{\delta^2}{\log \mu} + \|dE(u)\|_*^2.\end{aligned}$$

To obtain (4.14) it hence suffices to check that

$$|E(\mathfrak{z}) - E^*| \lesssim \tfrac{\delta^2}{\log \mu} \text{ for all } \mathfrak{z} \in \mathcal{Z}_0. \tag{4.15}$$

If $\delta = 0$, i.e. if $U_{0,1} = U^*(\cdot + c_{0,1})$ for numbers $c_0 = c_1$, then this is trivially true as in this case $\mathfrak{z} = U^*(q_\mu(\cdot) + c_0)$ has energy E^* since $q_\mu(z) = q_0(z) + q_1(\frac{1}{\mu z})$ is a rational map of degree $\deg(q_0) + \deg(q_1) = \deg(q_0^*) + \deg(q_1^*)$.

Conversely, if $c_0 \neq c_1$ then we interpolate between $\mathfrak{z}$ and $\hat{\mathfrak{z}} = U^* \circ (q_\mu + c_0)$, which satisfies $\delta(\hat{\mathfrak{z}}) = 0$ and thus $E(\hat{\mathfrak{z}}) = E^*$, using the family $(\mathfrak{z}_t)_{t \in [0, |c_1 - c_0|]}$ obtained by changing U_1 according to $U_{1,t} = U^*(\cdot + c_1 + t\frac{c_0 - c_1}{|c_1 - c_0|})$, while keeping U_0, $q_{0,1}$ and μ fixed. This variation satisfies $\|\partial_t \mathfrak{z}_t\| \lesssim 1$ and $\|\beta(\mathfrak{z}_t)\| \lesssim \delta(\mathfrak{z}_t) \lesssim \delta := \delta(\mathfrak{z})$ so we can use (3.14) to bound $|\partial_t E(\mathfrak{z}_t)| \lesssim \frac{\delta}{\log \mu}$ which, when integrated over $0 \leq t \leq |c_0 - c_1| \sim \delta$, yields (4.15) and hence (4.14).

From (4.13) and (4.14) we now immediately deduce that

$$|E(u) - E^*| \lesssim \min(\|dE(u)\|_*, \log \mu \|dE(u)\|_*^2)$$

which completes the proof of Theorem 1.3.

It remains to prove the L^2-Łojasiewicz-estimate (1.7) claimed in Corollary 1.4. To this end we first combine (4.13) with the fact that $\|\cdot\|_{L^2(S^2, g_{S^2})} \lesssim \|\cdot\|_{\mathfrak{z}}$, and hence $\|dE(u)\|_* \lesssim \|\tau_{g_{S^2}}(u)\|_{L^2(S^2, g_{S^2})}$, to bound

$$\tfrac{\delta}{\log \mu} \lesssim \|\tau_{g_{S^2}}(u)\|_{L^2(S^2, g_{S^2})} \|y_{\mathfrak{z}}\|_{L^2(S^2, g_{S^2})} + \|\tau_{g_{S^2}}(u)\|_{L^2(S^2, g_{S^2})}^2. \tag{4.16}$$

As $y_{\mathfrak{z}}$ is obtained from a variation of just the bubble U_1, it is mainly concentrated on a small ball around the point where the bubble is attached and hence its L^2 norm is small compared to $\|y_{\mathfrak{z}}\|_{\mathfrak{z}} = 1$. To be more precise, for variations satisfying (3.8) for which q_0 and U_0 are fixed we shall see in (5.19) that

$$\|\partial_\varepsilon \mathfrak{z}\|_{L^2(S^2, g_{S^2})} \lesssim \mu^{-1} (\log \mu)^{\frac{1}{2}} + \tfrac{\delta + \eta_0}{f_\mu \log \mu}. \tag{4.17}$$

In the present situation where $\eta_0 \sim 1$ we hence get that $\|y_{\mathfrak{z}}\|_{L^2(S^2,g_{S^2})} \lesssim (\log\mu)^{-1} f_\mu^{-1}$ so (4.16) allows us to conclude that

$$\frac{\delta^2}{\log\mu} \leq \|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2,g_{S^2})}. \tag{4.18}$$

Combined with (4.14) this immediately yields the claim $|E(u)-E^*| \lesssim \|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2,g_{S^2})}$ made in Corollary 1.4.

Remark 4.3. As noted in the introduction, the reason we cannot obtain these results from the existing theory is that the spectral gap at zero for the Jacobi-operator at $U^*(q_\mu)$ tends to zero as $\mu \to \infty$ and that the size of the neighbourhoods on which the existing theory yields Łojasiewicz-estimates scales like this spectral gap. To be more precise, as the energy defect scales like $\frac{\delta^2}{\log\mu}$ we can see that the Jacobi-operator must have an eigenvalue that scales like $\frac{1}{\log\mu}$. If we did not include these directions in our set of singularity models $\mathcal{Z}$, we would not get the uniform definiteness of d^2E orthogonal to $\mathcal{Z}$, which would mean that the argument of Lemma 4.2 would break down even for harmonic maps u unless we knew a priori that $\|w\|_{\mathfrak{z}}$ is bounded by a small enough multiple of $\frac{1}{\log\mu}$, compare (4.6).

4.2. Proofs of Theorems 1.5 and 1.7 and of Corollary 1.6

In this section we consider settings for which the assumptions of either Theorem 1.5 or of Theorem 1.7 are satisfied, i.e. for which either (A2) or (A3) of Assumption 1 hold. We note that in these settings $|d\beta(0)|$ and $\|\beta\|$ are of order 1 and first show that the energy defect of the singularity models $\mathfrak{z} \in \mathcal{Z}$ constructed in Section 2 is controlled by

Lemma 4.4. *For any $\mathfrak{z} \in \mathcal{Z}$ we have*

$$|E(\mathfrak{z}) - E^*| \lesssim \frac{\delta^2}{\log\mu} + \min(\nu_0,\nu_1) + \mu^{-\frac{1}{2}}\bar{\nu} + \mathcal{T}_1^{\alpha_1} + \mathcal{T}_0^{\alpha_0} + \mathcal{T}\mu^{-1}\log\mu, \tag{4.19}$$

where $\alpha_{0,1} \in (1,2]$ are the exponents for which the classical Łojasiewicz-Simon estimate (1.3) is valid near $U^*_{0,1}$ while $E^* := E(\omega_0) + E(\omega_1) = \deg(q_0)E(U_0^*) + \deg(q_1)E(U_1^*)$.

We note that the tension terms vanish for settings satisfying (A2), but that it is useful to work with this general formula as it will allow us to treat settings satisfying (A2) or (A3) simultaneously.

Proof of Lemma 4.4. It suffices to carry out the proof for elements of $\mathcal{Z}_0$ and using the symmetry described in Remark 2.2 we can assume without loss of generality that $\nu_0 \leq \nu_1$.

We will first prove the lemma in the special case where $\delta = \delta(\mathfrak{z}) = 0$. To this end, we consider the family $\mathfrak{z}(t)$, $t > 0$, which we obtain by varying the bubble scale according to $\mu(t) = \mu(\mathfrak{z})e^t$ while keeping the maps U_i and q_i fixed. From (3.15) and the subsequent remark we get that

$$|\tfrac{d}{dt}E(\mathfrak{z}(t))| = \mu(t)|\partial_\mu E(\mathfrak{z}(t))| \lesssim \nu_0(t) + \mu(t)^{-\frac{1}{2}}\nu_1(t) + \mathcal{T}\mu(t)^{-1}\log\mu(t) \text{ for } t \in [0,\infty),$$

where $\nu_i(t) = \max_{j\le n_i^*} |a_j(q_i)|\mu(t)^{-j} \lesssim e^{-t}\nu_i$, $\nu_i = \nu_i(\mathfrak{z})$. As $E(\mathfrak{z}(t)) \to \deg(q_0)E(U_0) + \deg(q_1)E(U_1)$ for $t \to \infty$ we hence deduce that

$$|E(\mathfrak{z}) - E^*| \le \sum_i \deg(q_i)|E(U_i) - E(U_i^*)| + C\big[\nu_0 + \mu^{-\frac{1}{2}}\nu_1 + \mathcal{T}\mu^{-1}\log\mu\big]$$

which, combined with (1.3), yields the claim in this special case where $\delta = 0$.

So suppose that $\delta = \delta(U_1, U_0) \neq 0$. In this case we choose $U_{i,t} = U_i(\cdot + c_i(t))$, $t \in [0,\delta]$ with $|\dot{c}_i(t)| \lesssim 1$ in a way that $\delta(U_{0,t}, U_{1,t}) = \delta - t$, compare Lemma 3.6. According to (3.14), the resulting family $\mathfrak{z}_t \in \mathcal{Z}$ satisfies

$$|\tfrac{d}{dt}E(z_t)| \lesssim \bar{\nu} + \tfrac{\delta}{\log\mu} + \tfrac{\mathcal{T}}{\log\mu},$$

where we note that the quantities on the right hand side are independent of t as we do not change the rational maps and only change U_i by a translation on the domain. Integrated over $[0,\delta]$, this ensures that the resulting element $\hat{\mathfrak{z}} = \mathfrak{z}_{t=\delta}$, which satisfies $\delta(\hat{\mathfrak{z}}) = 0$, must be so that

$$|E(\mathfrak{z}) - E(\hat{\mathfrak{z}})| \lesssim \tfrac{\delta^2}{\log\mu} + \delta\bar{\nu} + \tfrac{\delta}{\log\mu}\mathcal{T} \lesssim \tfrac{\delta^2}{\log\mu} + \bar{\nu}^2\log\mu + \mathcal{T}^2.$$

As $\bar{\nu}\log\mu \lesssim \mu^{-1}\log\mu \lesssim \mu^{-\frac{1}{2}}$ and as we have already shown that (4.19) holds for $\hat{\mathfrak{z}}$ we hence obtain the claimed bound on the energy defect also for general elements of $\mathcal{Z}_0$. □

We now explain how this lemma, combined with the results obtained in the previous section allows us to complete the

Proofs of Theorems 1.5 and 1.7. We recall that all previous considerations are applicable for singularity models obtained for any choice of function f_μ satisfying (2.12). For the present proof it is convenient to choose $f_\mu := \log\mu$, which of course satisfies this assumption (2.12) (provided $\bar{\mu}$ is sufficiently large).

As we are dealing with almost harmonic maps which converge to a bubble tree, we only need to consider maps u with $\|\tau_{g_{S^2}}(u)\|_{L^2(S^2,g_{S^2})} \le 1$ which satisfy (4.1) for some $\mathfrak{z} \in \mathcal{Z}$ and for $\varepsilon > 0$ as in Lemma 4.2.

We first note that

$$|E(u) - E(\mathfrak{z})| \lesssim \|dE(\mathfrak{z})\|_*^2 + \|w\|_{\mathfrak{z}}^2 \lesssim \bar{\nu}^2\log\mu + \tfrac{\delta^2(\log f_\mu)^2}{(\log\mu)^2} + \mathcal{T}^2 + \|dE(u)\|_*^2,$$

thanks to (4.3) and Lemma 3.2, and that combining this with Lemma 4.4 implies that

$$|E(u) - E^*| \lesssim \tfrac{\delta^2}{\log\mu} + \min(\nu_0,\nu_1) + \mu^{-\frac{1}{2}}\bar{\nu} + \mathcal{T}^{\bar{\alpha}} + \mathcal{T}\mu^{-1}\log\mu + \|dE(u)\|_*^2 \tag{4.20}$$

for any such u and $\mathfrak{z}$, and for $\bar{\alpha} := \min(\alpha_1, \alpha_2) \in (1,2]$.

In order to prove the bounds on $|E(u)-E^*|$ claimed in Theorems 1.5 and 1.7 we hence need to show that all these quantities are controlled by $\|dE(u)\|_*$. To prove this, and the related claim on the bubble scale made in the theorems, we will split our analysis into three cases, depending on whether the influence of δ, $\mathcal{T}$ or $\bar{\nu}$ dominates the behaviour of the energy and its relevant variations at $\mathfrak{z}$.

Situations where δ dominates, and where we will hence consider a variation of type (V1) that increases δ as done in Lemma 3.7 and the above proof of Theorem 1.3, correspond to

Case 1: $\mathfrak{z} \in \mathcal{Z}$ is so that

$$\bar{\nu}(\log \mu)^{\frac{1}{2}} \leq \tfrac{\delta}{\log \mu} \text{ and } \mathcal{T}(\log \log \mu)^{-1/2} \leq \tfrac{\delta}{\log \mu}. \tag{4.21}$$

In this case we let $y_{\mathfrak{z}}$ be the unit element in the direction of a variation $\partial_\varepsilon \mathfrak{z}_\varepsilon$ as obtained in Lemma 3.7. Inserting (4.21) into the estimates obtained in Lemmas 3.2, 3.3 and 3.7 gives

$$\|dE(\mathfrak{z})\|_* \lesssim \tfrac{\delta(\log \log \mu)^{\frac{1}{2}}}{\log \mu}, \quad \|d^2E(\mathfrak{z})(y_{\mathfrak{z}})\|_* \lesssim \tfrac{(\log \log \mu)^{\frac{1}{2}}}{\log \mu} \text{ and } dE(\mathfrak{z})(y_{\mathfrak{z}}) \geq c\tfrac{\delta}{\log \mu}. \tag{4.22}$$

This ensures that $Q(\mathfrak{z}, y_{\mathfrak{z}}) \lesssim \frac{\log \log \mu}{\log \mu}$ is small. Hence Lemma 4.2 is applicable and yields

$$\tfrac{\delta}{\log \mu} \lesssim |dE(u)(y_{\mathfrak{z}})| + \|dE(u)\|_*^2 \lesssim \|dE(u)\|_*.$$

Combined with (4.21) this immediately implies the claimed bound $\mu^{-\min(n_0^*, n_1^*)} \lesssim \bar{\nu} \lesssim \|dE(u)\|_*$ on the bubble scale, and inserting these estimates into (4.20) furthermore gives

$$|E(u) - E^*| \lesssim \tfrac{\delta}{\log \mu} + \left(\tfrac{\delta}{\log \mu}\right)^{\alpha} (\log \log \mu)^{\alpha/2} + \|dE(u)\|_*^2 \lesssim \tfrac{\delta}{\log \mu} + \|dE(u)\|_*^2 \lesssim \|dE(u)\|_*.$$

This completes the proofs of Theorems 1.5 and 1.7 in this first case where δ dominates.

Next, we consider the situation where $\mathcal{T}$ dominates and where we hence want to consider variations of type (V3). We recall that this is only relevant for settings satisfying the assumption (A3). This corresponds to

Case 2: $\mathfrak{z}$ is so that

$$\tfrac{\delta(\log \log \mu)^{1/2}}{\log \mu} \leq \mathcal{T} \text{ and } \bar{\nu}(\log \mu)^{\frac{1}{2}} \leq \mathcal{T}. \tag{4.23}$$

In this situation (4.20), and the fact that $\bar{\alpha} = \min(\alpha_0, \alpha_1) > 1$, ensures that

$$|E(u) - E^*| \lesssim \mathcal{T}. \tag{4.24}$$

To obtain a bound on $\mathcal{T} = \max(\mathcal{T}_0, \mathcal{T}_1)$ we fix i so that $\mathcal{T}$ is given by $\mathcal{T}_i$ and use a variation $\mathfrak{z}_\varepsilon$ as considered in Lemma 3.8. This lemma, as well as Lemma 3.2 and (4.23) ensure that

$$dE(\mathfrak{z})(y_{\mathfrak{z}}) \gtrsim \mathcal{T} \text{ and } \|dE(\mathfrak{z})\|_* \lesssim \mathcal{T} \text{ for } y_{\mathfrak{z}} = \|\partial_\varepsilon \mathfrak{z}_\varepsilon\|^{-1}\partial_\varepsilon \mathfrak{z}_\varepsilon. \tag{4.25}$$

At the same time, Lemma 3.3 and the subsequent remark ensure that we can make $\|d^2E(\mathfrak{z})(y_{\mathfrak{z}},\cdot)\|_*$ smaller than any given constant by reducing σ accordingly. Hence we can ensure that $Q(\mathfrak{z}, y_{\mathfrak{z}}) \lesssim \bar{\Lambda}^{-1}$ for $\bar{\Lambda}$ as in Lemma 4.2 and apply this lemma to get

$$\mathcal{T} \lesssim \|dE(u)(y_{\mathfrak{z}})\| + \|dE(u)\|_*^2 \lesssim \|dE(u)\|_*.$$

Combined with (4.20) and (4.23) this immediately yields the claimed bound on $|E(u) - E^*|$, while the claim on the bubble scale follows from $\mu^{-\min(n_0^*, n_1^*)} \lesssim \bar{\nu} \le \mathcal{T} \lesssim \|dE(u)\|_*$.

It remains to consider pairs u and $\mathfrak{z}$ for which both (4.21) and (4.23) are violated, and for which we hence know that

Case 3: $\mathfrak{z}$ is so that

$$\tfrac{\delta}{\log \mu} \le \bar{\nu}(\log \mu)^{\frac{1}{2}} \text{ and } \mathcal{T} \le \bar{\nu}(\log \mu). \tag{4.26}$$

In this case we will use variations of type (V2). To this end, we select i so that $\nu_{i^*} = \bar{\nu}$ (for $0^* = 1$ and $1^* = 0$) and use a variation of the rational map q_i as described in Lemma 3.5 to obtain a unit element $y_{\mathfrak{z}}$ with

$$dE(\mathfrak{z})(y_{\mathfrak{z}}) \gtrsim \bar{\nu} \text{ and } \|d^2E(\mathfrak{z})(y_{\mathfrak{z}},\cdot)\|_* \lesssim \mu^{-1}\log \mu.$$

As Lemma 3.2 and (4.26) allow us to bound $\|dE(\mathfrak{z})\|_* \le \bar{\nu}\log \mu$ this ensures that $Q(\mathfrak{z}, y_{\mathfrak{z}}) \lesssim \mu^{-1}(\log \mu)^2$ will be small, allowing us to apply Lemma 4.2 to deduce that $\bar{\nu} \lesssim \|dE(u)\|_*$. This immediately yields the claimed bound on the bubble scale and, as (4.26) ensures that all terms that involve δ or $\mathcal{T}$ in (4.20) are small compared to $\bar{\nu}$, also that $|E(u) - E^*| \lesssim \bar{\nu} \lesssim \|dE(u)\|_*$. This completes the proofs of both Theorems 1.5 and 1.7. □

To prove Corollary 1.6, i.e. to show that L^2-Łojasiewicz-estimates with an exponent $\alpha > 1$ hold, we need to modify the above argument in a way that favours variations for which only U_1 or q_1 change, as such variations lead to $\|\cdot\|_{\mathfrak{z}}$-unit elements $y_{\mathfrak{z}}$ with very small L^2 norm, compare (5.19).

Proof of Corollary 1.6. For the proof of this result we consider singularity models which are defined using $f_\mu := \mu^{\sigma_1}$, for $\sigma_1 > 0$ small enough so that all results in Section 3 hold for this σ_1 and for all sufficiently small σ and large $\bar{\mu}$. This will allow us to consider variations of the bubble instead of the base also in some situations where a variation of the base would result in a slightly larger change of the energy, which will be crucial to obtain estimates with exponent strictly larger than 1.

Before we split into cases that identify suitable variations based on the relative size of δ, ν_0 and ν_1 we recall that Corollary 1.6 corresponds to a setting in which the non-degeneracy condition (A2) holds. This means that all tension terms vanish and that (4.20) reduces to

$$|E(u) - E^*| \lesssim \tfrac{\delta^2}{\log \mu} + \min(\nu_0, \nu_1) + \mu^{-\frac{1}{2}}\bar{\nu} + (\|dE(u)\|_*)^2. \tag{4.27}$$

As $\|dE(u)\|_* \lesssim \|\tau_{g_{S^2}}(u)\|_{L^2(S^2, g_{S^2})}$, it hence suffices to show that the first three terms on the right hand side are bounded by $C\|\tau_{g_{S^2}}(u)\|^{\alpha}_{L^2(S^2, g_{S^2})}$ for some exponent $\alpha > 1$.

In this context we will use variations of type (V1) in

Case 1: $\mu^{c_1}\bar{\nu} \le \frac{\delta}{\log \mu}$, for $c_1 := \frac{\sigma_1}{4} \in (0, \frac{1}{16})$.

Here the influence of δ dominates and we can argue exactly as in the proof of Corollary 1.4 to see that

$$\tfrac{\delta}{(\log \mu)^2} \lesssim \|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2, g_{S^2})},$$

compare (4.18). As $\bar{\nu} \gtrsim \mu^{-n^*}$, $n^* := \min(n_1^*, n_0^*)$, we furthermore know that in this case

$$\bar{\nu}^{1-\frac{c_1}{n^*}} \lesssim \mu^{c_1}\bar{\nu} \le \tfrac{\delta}{\log \mu} \lesssim \|\tau_{g_{S^2}}(u)\|_{L^2(S^2, g_{S^2})}.$$

Inserted into (4.27) this immediately implies that

$$|E(u) - E^*| \lesssim \|\tau_{g_{S^2}}(u)\|^{\alpha}_{L^2(S^2, g_{S^2})} \text{ for any } \alpha \le (1 - \tfrac{c_1}{n^*})^{-1}.$$

It remains to consider

Case 2: $\mu^{c_1}\bar{\nu} \ge \frac{\delta}{\log \mu}$, for $c_1 := \frac{\sigma_1}{4} \in (0, \frac{1}{16})$.

As this condition ensures that $\frac{\delta^2}{\log \mu} \lesssim \bar{\nu}\mu^{-1+2c_1}\log \mu \ll \mu^{-1/2}\bar{\nu}$ we obtain from (4.27) and Lemma 3.2 that

$$|E(u) - E^*| \lesssim \min(\nu_0, \nu_1) + \mu^{-\frac{1}{2}}\bar{\nu} + \|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2, g_{S^2})} \text{ and } \|dE(z)\|_* \lesssim \mu^{c_1}(\log \mu)^{\frac{1}{2}}\bar{\nu}, \tag{4.28}$$

while Lemmas 3.5 and 3.3 ensure that

$$dE(\mathfrak{z})(y_{\mathfrak{z}}) \ge c\nu_{i^*} - C\mu^{-2c_1}\bar{\nu} \text{ and } \|d^2E(\mathfrak{z})(y_{\mathfrak{z}}, \cdot)\|_* \lesssim \mu^{-1+c_1}(\log \mu)^{\frac{1}{2}} \tag{4.29}$$

for the unit element $y_{\mathfrak{z}} = \|\partial_\varepsilon z\|_{\mathfrak{z}}^{-1} y_{\mathfrak{z}}$ that is obtained from a variation of type (V2) for which q_i changes as described in Lemma 3.5.

If $\mu^{c_1}\nu_0 \ge \nu_1$ we apply this for $i = 1$, and hence obtain that the resulting $y_{\mathfrak{z}}$ is so that

$$dE(\mathfrak{z})(y_{\mathfrak{z}}) \ge c\nu_0 - C\mu^{-2c_1}\nu_1 \gtrsim \nu_0,$$

and so that $Q(\mathfrak{z})(y_{\mathfrak{z}}) \lesssim \mu^{-1+3c_1}\log \mu$ is small. Hence Lemma 4.2 is applicable and yields

$$\begin{aligned}\nu_0 &\lesssim |dE(u)(y_{\mathfrak{z}})| \leqslant \|\tau_{g_{S^2}}(u)\|_{L^2(S^2, g_{S^2})}\|y_{\mathfrak{z}}\|_{L^2(S^2, g_{S^2})} + \|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2, g_{S^2})} \\ &\lesssim \mu^{-1}(\log \mu)^{\frac{1}{2}}\|\tau_{g_{S^2}}(u)\|_{L^2(S^2, g_{S^2})} + \|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2, g_{S^2})}\end{aligned}$$

since such a variation of the rational map q_1 of the bubble leads to an element $y_{\mathfrak{z}}$ whose L^2 norm is of order $\mu^{-1}(\log\mu)^{\frac{1}{2}}$, compare (5.19). In this case where $\mu^{c_1}\bar\nu \geq \frac{\delta}{\log\mu}$ and $\mu^{c_1}\nu_0 \geq \nu_1$ we can thus bound

$$\mu^{-\frac{1}{2}}\bar\nu + \nu_0 \lesssim \nu_0 \lesssim \|\tau_{g_{S^2}}(u)\|^{\alpha}_{L^2(S^2,g_{S^2})} \text{ for any } \alpha \in (1,2] \text{ with } \alpha < (1-\tfrac{1}{n^*})^{-1}$$

and (4.28) implies that the claim of the corollary holds true for any such exponent.

It remains to consider the case where $\nu_0 \leq \mu^{-c_1}\nu_1$ and $\frac{\delta}{\log\mu} \leq \mu^{c_1}\bar\nu$. Here we instead use a variation of q_0, resulting in an element $y_{\mathfrak{z}}$ with $\|y_{\mathfrak{z}}\|_{L^2(S^2,g_{S^2})} \sim 1$ and $dE(\mathfrak{z})(y_{\mathfrak{z}}) \gtrsim \bar\nu$, compare (4.29). As $Q(\mathfrak{z})(y_{\mathfrak{z}}) \lesssim \mu^{-1+2c_1}\log\mu$ is small, we can apply Lemma 4.2 to deduce that

$$\bar\nu = \nu_1 \lesssim \|dE(u)\|_* \lesssim \|\tau_{g_{S^2}}(u)\|_{L^2(S^2,g_{S^2})}.$$

Inserted into (4.28) this yields that

$$\begin{aligned}|E(u)-E^*| &\lesssim \mu^{-c_1}\bar\nu + \|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2,g_{S^2})}\\ &\lesssim \bar\nu^{\alpha} + \|\tau_{g_{S^2}}(u)\|^2_{L^2(S^2,g_{S^2})} \lesssim \|\tau_{g_{S^2}}(u)\|^{\alpha}_{L^2(S^2,g_{S^2})}\end{aligned}$$

provided $\alpha \in (1,2]$ is chosen so that $\alpha \leq 1 + \frac{c_1}{n^*}$. □

5. Variations of the energy along $\mathcal{Z}_0$

In this section we carry out the proofs of the key lemmas on the behaviour of the energy on the manifold $\mathcal{Z}_0$ of singularity models, i.e. of Lemmas 3.4, 3.5, 3.7 and 3.8, that we stated in Section 3.

Before we turn to these proofs we collect a number of estimates, which will be used both throughout this section as well as in the subsequent Section 6 where we will prove the properties of $dE(\mathfrak{z})$ and $d^2E(\mathfrak{z})$ claimed in Lemmas 3.1, 3.2 and 3.3.

5.1. Technical estimates

Let $\mathcal{Z}_0$ be the set of singularity models constructed in Section 2. Given any $\mathfrak{z} = \mathfrak{z}(U_0 \circ q_0, U_1 \circ q_1(\frac{1}{\mu}\cdot)) \in \mathcal{Z}_0$ we work in fixed stereographic coordinates which we can assume to be scaled in a way that $|a_{n_0^*}(q_0)| = 1$. As we represent $q_1(\frac{1}{\mu}\cdot)$ for any fixed $\mathfrak{z}$ by a function q_1 satisfying $|a_{n_1^*}(q_1)| = 1$ we hence know that $\mu_1(\mathfrak{z}) = \mu$, $\mu_0(\mathfrak{z}) = 1$, compare (2.9), and hence that the radii of the annuli defined in (2.10) are given by

$$r_1(\mathfrak{z}) = f_\mu \mu^{-1} \ll \hat r(\mathfrak{z}) = \mu^{-\frac{1}{2}} \ll r_0(\mathfrak{z}) = f_\mu{}^{-1}. \tag{5.1}$$

To simplify the notation we will in the following write for short $\phi_\mu = \phi_{r_0,r_1}$ for the cut-off function (2.15) used in the definition of $\mathfrak{z}$ and also use the slight abuse of notation of writing $\partial_\varepsilon\phi_\mu := \partial_\varepsilon|_{\varepsilon=0}\phi_{r_0(\mathfrak{z}_\varepsilon),r_1(\mathfrak{z}_\varepsilon)}$ even though (5.1) is only applicable at $\varepsilon = 0$.

For most of the proofs it will be more convenient to view the rational maps q_i as maps into the sphere $S^2_{p^*} := S^2 - (0,0,1)$, which is shifted in a way that elements of $\mathcal{R}$ map 0 to the origin in $\mathbb{R}^3$. We hence associate to any given $U : \hat{\mathbb{C}} \to N$ and any rational map $q : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ the maps $\hat{U} := U \circ \pi_{p^*}^{-1} : S^2_{p^*} \to N$ and $\hat{q} = \pi_{p^*} \circ q : \hat{\mathbb{C}} \to S^2_{p^*}$ for $\pi_{p^*} := \pi - (0,0,1) : \mathbb{R}^2 \to S^2_{p^*}$.

As at least one of $q_0(z)$ or $q_{1,\mu}(z) := q_1(\frac{1}{\mu z})$ will be small for every $z \in \mathbb{C}$ we can easily check that the map $\hat{q}_\mu = \pi_{p^*} \circ (\hat{q}_0 \circ \pi_{p^*}^{-1} + \hat{q}_{1,\mu} \circ \pi_{p^*}^{-1})$ that represents $q_\mu = q_0 + q_{1,\mu}$ is so that

$$|\hat{q}_\mu| \lesssim |\hat{q}_{1,\mu}| + |\hat{q}_0| \text{ and } |\nabla \hat{q}_\mu| \lesssim |\nabla \hat{q}_0| + |\nabla \hat{q}_{1,\mu}|. \tag{5.2}$$

Here and in the following we use the convention that unless specified otherwise all quantities are computed with respect to the euclidean metric on $\mathbb{R}^2$, rather than with respect to the metric that is induced by g_{S^2}.

We can hence estimate the maps $\tilde{u}_i := u_i - U_i(0) = \hat{U}_i \circ \hat{q}_\mu - U_i(0)$ by

$$|\tilde{u}_i| \lesssim |\hat{q}_{1,\mu}| + |\hat{q}_0|, \quad |\nabla u_i| \lesssim |\nabla \hat{q}_0| + |\nabla \hat{q}_{1,\mu}| \text{ and } |\Delta u_i| \lesssim |\nabla \hat{q}_{1,\mu}|^2 + |\nabla \hat{q}_0|^2, \tag{5.3}$$

where the last estimate follows as the maps $\hat{q}_i$ are harmonic maps into the sphere.

Since the maps $\hat{q}_i \circ \pi_{p^*}^{-1} : S^2_{p^*} \to S^2_{p^*}$ and their variations are uniformly bounded in C^k, we have

$$|\hat{q}_0(z)| + |\partial_\varepsilon \hat{q}_0| \lesssim |\pi(z) - \pi(0)| \lesssim \tfrac{|z|}{1+|z|}, \quad |\nabla \hat{q}_0| + |\partial_\varepsilon \nabla \hat{q}_0| \lesssim |\nabla \pi(z)| \lesssim \tfrac{1}{1+|z|^2} \tag{5.4}$$

and

$$|\hat{q}_{1,\mu}| + |\partial_\varepsilon \nabla \hat{q}_0| \lesssim |\pi(\tfrac{1}{\mu z}) - \pi(0)| \lesssim \tfrac{1}{1+|\mu z|}, \quad |\nabla \hat{q}_{1,\mu}| + |\partial_\varepsilon \nabla \hat{q}_{1,\mu}| \lesssim |\nabla(\pi(\tfrac{1}{\mu z}))| \lesssim \tfrac{\mu}{1+|\mu z|^2}. \tag{5.5}$$

In particular we can always estimate $|\nabla \hat{q}_\mu| \lesssim \rho_{\mathfrak{z}}$, though will later need more refined bounds in settings where we deal with rational maps $q^*_{0,1}$ that are branched at $z = 0$, compare (5.45) and (5.46) below.

The above estimates imply in particular that

$$\| |\hat{q}_{1,\mu}| + |\hat{q}_0| + |\partial_\varepsilon \hat{q}_{1,\mu}| + |\partial_\varepsilon \hat{q}_0| \|_{L^\infty(A)} + \| |\nabla \hat{q}_{1,\mu}| + |\nabla \hat{q}_0| \|_{L^2(A)} \lesssim f_\mu{}^{-1}, \tag{5.6}$$

where we recall that the annuli A, $\hat{A}$ and A^* are as defined in (2.10). Similarly we have

$$\int_{\Omega_1} \phi_\mu{}^2 \rho_{\mathfrak{z}}^2 \lesssim f_\mu{}^{-2} (\log \mu)^{-2}, \int_{\Omega_1} \phi_\mu \rho_{\mathfrak{z}}^2 \lesssim f_\mu{}^{-2} (\log \mu)^{-1}, \int_A |\nabla \phi_\mu| \rho_{\mathfrak{z}} \lesssim f_\mu{}^{-1} (\log \mu)^{-1} \tag{5.7}$$

for the cut-off function ϕ_μ that is defined in (2.15) and that satisfies

$$|\nabla \phi_\mu| \lesssim c_\mu r^{-1} \mathbb{1}_A, \quad |\Delta \phi_\mu| \lesssim c_\mu r^{-2} \mathbb{1}_{A^*} \text{ for } c_\mu := \log(\tfrac{r_0}{r_1}). \tag{5.8}$$

As $\hat{\gamma}$ is a geodesic with $\hat{\gamma}(0) = U_1(0)$ and $|\hat{\gamma}'| = \delta$, δ as in (3.5), we have

$$|\tilde{\gamma}_{U_1}| \leq \delta\phi_\mu \text{ with } |\nabla\tilde{\gamma}_{U_1}| = |\nabla\gamma| \lesssim \delta|\nabla\phi_\mu| \lesssim c_\mu\delta r^{-1}\mathbb{1}_A \tag{5.9}$$

for $\tilde{\gamma}_\cdot$ defined as in (2.16) and

$$\tau(\gamma) = \Delta\phi_\mu \cdot \hat{\gamma}' \circ \phi_\mu \text{ while } P_\gamma^\perp(\Delta\gamma) = |\nabla\phi_\mu|^2 \cdot A(\hat{\gamma})(\hat{\gamma}', \hat{\gamma}') \circ \phi_\mu \tag{5.10}$$

so

$$|\tau(\gamma)| \lesssim \delta|\Delta\phi_\mu| \lesssim \delta c_\mu r^{-2}\mathbb{1}_{A^*} \quad \text{and} \quad |P_\gamma^\perp(\Delta\gamma)| \lesssim \delta^2|\nabla\phi_\mu|^2 \leq c_\mu^2 r^{-2}\delta^2\mathbb{1}_A. \tag{5.11}$$

As $c_\mu \sim (\log\mu)^{-1}$ we hence get

$$\|\tau(\gamma)\|_{L^1} \lesssim \delta(\log\mu)^{-1}, \quad \|P_\gamma^\perp(\Delta\gamma)\|_{L^1} \lesssim \delta^2(\log\mu)^{-1} \text{ and } \|\Delta\gamma\|_{L^1} \lesssim \delta(\log\mu)^{-1}. \tag{5.12}$$

To obtain suitable bounds on the variations of ϕ_μ we recall that $|\partial_\varepsilon\mu_i| \lesssim \mu_i$ and that $|\partial_\mu f_\mu|\mu \leq f_\mu$. This ensures that

$$|\partial_\varepsilon r_i| \lesssim r_i, \quad |\partial_\varepsilon\hat{r}| \lesssim \hat{r} \text{ and } |\partial_\varepsilon c_{r_0,r_1}| \lesssim {c_\mu}^2 \lesssim (\log\mu)^{-2} \tag{5.13}$$

which in turn imply that

$$|\partial_\varepsilon\phi_\mu| \leqslant c_\mu\mathbb{1}_A, \quad |\partial_\varepsilon\nabla\phi_\mu| \leqslant c_\mu^2 r^{-1}\mathbb{1}_A + c_\mu r^{-1}\mathbb{1}_{A^*}, \quad |\partial_\varepsilon\Delta\phi_\mu| \leqslant c_\mu r^{-2}\mathbb{1}_{A^*}. \tag{5.14}$$

We also observe that the definition of ϕ_μ and the above estimates ensure that

$$\|\nabla\phi_\mu\|_{L^2}^2 = 2\pi c_\mu + O((\log\mu)^{-2}) \text{ and that } |\partial_\varepsilon\|\nabla\phi_\mu\|_{L^2(\mathbb{R}^2)}^2| \lesssim (\log\mu)^{-2}. \tag{5.15}$$

As $|\partial_\varepsilon\hat{\gamma}| + |\partial_\varepsilon\nabla\hat{\gamma}| \lesssim \eta_0$, η_0 defined in (3.9), we can furthermore bound

$$|\partial_\varepsilon\gamma| \lesssim \eta_0 + c_\mu\delta\mathbb{1}_A, \quad |\partial_\varepsilon\tilde{\gamma}_{U_1}| \lesssim \eta_0\phi_\mu + c_\mu\delta\mathbb{1}_A, \quad |\partial_\varepsilon\tilde{\gamma}_{U_0}| \lesssim \eta_0(1-\phi_\mu) + c_\mu\delta\mathbb{1}_A \tag{5.16}$$

and

$$|\partial_\varepsilon\nabla\gamma| \lesssim (\eta_0 + \delta c_\mu)c_\mu r^{-1}\mathbb{1}_A + \delta c_\mu r^{-1}\mathbb{1}_{A^*} \text{ so } \|\partial_\varepsilon\nabla\gamma\|_{L^2} \lesssim \eta_0(\log\mu)^{-\frac{1}{2}} + \delta(\log\mu)^{-1}. \tag{5.17}$$

We also note that $|q_0(z)| \lesssim |z|$ is small on the set $\{z : |z| \leq 2\mu^{-\frac{1}{2}}\}$ where we consider the function v_1 (defined in (2.18)) and hence $j_1 = d\beta(0)(q_0(z))$. On this set we can hence bound $|q_0(z)| \lesssim |\hat{q}_0(z)|$ and thus have

$$|j_1| \lesssim |d\beta(0)||\hat{q}_0|, \quad |\nabla j_1| \lesssim |d\beta(0)||\nabla\hat{q}_0|, \quad \Delta j_1 = 0$$

as well as

$$|\partial_\varepsilon j_1| \lesssim |d\beta(0)||\partial_\varepsilon \hat{q}_0| + |\partial_\varepsilon d\beta(0)||\hat{q}_0|, \ |\partial_\varepsilon \nabla j_1| \lesssim |d\beta(0)||\partial_\varepsilon \nabla \hat{q}_0| + |\partial_\varepsilon d\beta(0)||\nabla \hat{q}_0|. \quad (5.18)$$

Analogue estimates hold for j_0 on the set $|z| \geq \frac{1}{2}\mu^{-\frac{1}{2}}$ where v_0, and hence j_0, are considered.

We also observe that we can combine the above estimates with (5.57) to see that if $\mathfrak{z}_\varepsilon$ is a variation along which q_0 and U_0 are fixed then

$$\begin{aligned} \|\partial_\varepsilon \mathfrak{z}\|_{L^2(S^2,g_{S^2})} &\lesssim \|\partial_\varepsilon v_1|\nabla\pi|\|_{L^2(\mathbb{D}_{2\hat{r}})} + \|\partial_\varepsilon v_0|\nabla\pi|\|_{L^2(\mathbb{R}^2\setminus\mathbb{D}_{\frac{\hat{r}}{2}})} + |\hat{A}|^{\frac{1}{2}}\|v_1 - v_0\|_{L^\infty} \\ &\lesssim \|z\|_{L^2(\mathbb{D}_{2\hat{r}})} + \|\tfrac{1}{1+\mu|z|}|\nabla\pi|\|_{L^2(\mathbb{R}^2)} + \tfrac{\delta+\eta_0}{\log\mu}(\|\phi_\mu\|_{L^2(\mathbb{R}^2)} + |\mathbb{D}_{2\hat{r}}|^{\frac{1}{2}}) + \mu^{-1} \\ &\lesssim \mu^{-1}(\log\mu)^{\frac{1}{2}} + \tfrac{\delta+\eta_0}{f_\mu \log\mu} \end{aligned} \quad (5.19)$$

as claimed in the proofs of Corollaries 1.4 and 1.6.

5.2. General variations of the energy along $\mathcal{Z}$

The purpose of this section is to derive the expression for $\partial_\varepsilon E(\mathfrak{z}_\varepsilon)$ claimed in Lemma 3.4.

Thanks to the symmetry of the construction and the fact that on $\Omega_1 = \{z : |z| \leq \mu^{-\frac{1}{2}}\}$ our map $\mathfrak{z}$ is essentially described by $\mathfrak{z}_1 = \pi_N(v_1)$ the main step in this proof is to show

Lemma 5.1. *If $\mathfrak{z}_\varepsilon$ is a variation in $\mathcal{Z}_0$ so that (3.8) is satisfied then the corresponding $\mathfrak{z}_1 = \pi_N(v_1)$, v_1 as defined in (2.18), is so that*

$$\int_{\Omega_1} \partial_\varepsilon \mathfrak{z}_1 \Delta \mathfrak{z}_1 = -\int_{\Omega_1} j_1 \partial_\varepsilon \Delta u_1 - \partial_\varepsilon u_1 \tau(u_1) + \int_{A\cap\Omega_1} \Delta\gamma \partial_\varepsilon \gamma + err_1 \quad (5.20)$$

for u_i, j_i and γ as in (2.13), (2.17) and (2.14) and an error term which is bounded by $|err_1| \leq R_1$ for

$$\begin{aligned} R_1 := &\delta(\log\mu)^{-1}{f_\mu}^{-2} + \delta^2(\log\mu)^{-1} + \eta_0[|d\beta(0)|I_3 + \delta(\log\mu)^{-1}{f_\mu}^{-1}] \\ &+ |d\beta(0)|^2(I_1 + I_2) + |d\beta(0)||\partial_\varepsilon d\beta(0)|I_1 + |d\beta(0)|I_{\partial\Omega_1} \\ &+ \mathcal{T}[|d\beta(0)|(I_3 + I_4) + \eta_0(\log\mu)^{-1}{f_\mu}^{-2} + |\partial_\varepsilon d\beta(0)|I_3] + \eta_\tau|d\beta(0)|I_3. \end{aligned} \quad (5.21)$$

Here δ, τ, η_0 and η_τ are as in (3.5) and (3.9) and the integrals $I_\cdot = I_\cdot(q_0, q_{1,\mu})$ are as in

Lemma 5.2. *For any variations $q_{0,1}^{(\varepsilon)}$ in $\mathcal{R}^\sigma(q_{0,1}^*)$, $\mathcal{R}^\sigma(\cdot)$ as defined in (2.4), which satisfy (3.8), we can bound*

$$I_1 := \int_{\Omega_1} |\hat{q}_0|^2\rho_q^2 + |\hat{q}_{1,\mu}|^2|\nabla\hat{q}_0|^2 \leq \bar{\nu}^2 \log\mu \quad (5.22)$$

$$I_2 := \int_{\Omega_1} |\hat{q}_0||\partial_\varepsilon \hat{q}_0|\rho_q^2 + |\partial_\varepsilon \hat{q}_0||\hat{q}_{1,\mu}||\nabla \hat{q}_0||\nabla q_{1,\mu}| + |\hat{q}_{1,\mu}||\nabla \hat{q}_0|^2 + |\hat{q}_0|\rho_q|\nabla \hat{q}_0| \le \mu^{-\frac{1}{2}}\bar{\nu} \tag{5.23}$$

$$I_3 := \int_{\Omega_1} |\hat{q}_0|\rho_q^2 + |\nabla \hat{q}_0||\hat{q}_{1,\mu}|\rho_q \le \bar{\nu} \tag{5.24}$$

$$I_4 := \int_{\Omega_1} |\partial_\varepsilon \hat{q}_0|\rho_q^2 + |\hat{q}_0||\partial_\varepsilon \nabla \hat{q}_\mu|\rho_q \lesssim \mu^{-1} \tag{5.25}$$

$$I_{\partial\Omega} := \int_{\partial\Omega_1} |\hat{q}_0|^2 + |\hat{q}_{1,\mu}|^2 dS \le \mu^{-\frac{1}{2}}\bar{\nu} \tag{5.26}$$

where we write for short for $\rho_q := |\nabla \hat{q}_0| + |\nabla \hat{q}_{1,\mu}|$.

This lemma ensures that R_1 defined by (5.21) is bounded by the quantity R_0 defined in (3.12) and hence that err_1 can be included in the error term in Lemma 3.4.

Similarly the contributions of $\mathfrak{z} - \mathfrak{z}_1$ to $\partial_\varepsilon E(\mathfrak{z}_\varepsilon)$ are of lower order, namely

Lemma 5.3. *For any variation in* $\mathcal{Z}_0$ *as considered above we can bound*

$$\left| \int_{\Omega_1} \partial_\varepsilon \mathfrak{z} \Delta \mathfrak{z} - \partial_\varepsilon \mathfrak{z}_1 \Delta \mathfrak{z}_1 \right| \le \|\beta\|\bar{\nu}(\mu^{-\frac{1}{2}} + \eta_0) + \delta\mu^{-1} + \mathcal{T}\bar{\nu}\mu^{-1}(\|\beta\| + \|\partial_\varepsilon \beta\|). \tag{5.27}$$

Combining these three lemmas with their analogues on Ω_0, which follow by symmetry, we get

$$\begin{aligned} \partial_\varepsilon E(\mathfrak{z}) &= -\int_{\Omega_1} \partial_\varepsilon \mathfrak{z} \Delta \mathfrak{z} - \int_{\Omega_0} \partial_\varepsilon \mathfrak{z} \Delta \mathfrak{z} \\ &= \int_{\Omega_1} j_1 \partial_\varepsilon \Delta u_1 - \partial_\varepsilon u_1 \tau(u_1) + \int_{\Omega_0} j_0 \partial_\varepsilon \Delta u_0 - \partial_\varepsilon u_0 \tau(u_0) - \int_A \Delta\gamma \partial_\varepsilon \gamma + \mathrm{err}_2 \end{aligned} \tag{5.28}$$

for an error term that satisfies the required bound $|\mathrm{err}_2| \lesssim R_0$.

As $|\hat{\gamma}'| = \delta$ we can use (5.15) to see that

$$-\int_A \partial_\varepsilon \gamma \Delta \gamma = \tfrac{1}{2}\tfrac{d}{d\varepsilon}\|\nabla\gamma\|_{L^2}^2 = \tfrac{1}{2}\tfrac{d}{d\varepsilon}\big(|\hat{\gamma}'|^2\|\nabla\phi_\mu\|_{L^2}^2\big) = \pi c_\mu \tfrac{d}{d\varepsilon}\delta(U_0, U_1)^2 + \mathrm{err}_3 \tag{5.29}$$

for an error term that is bounded by $|\mathrm{err}_3| \lesssim \frac{\delta(\delta+\eta_0)}{(\log\mu)^2} \lesssim R_0$.

Finally, we can use that

$$\deg(q_1)\tfrac{d}{d\varepsilon}E(\hat{U}_1) = \tfrac{d}{d\varepsilon}E(\hat{U}_1\circ\hat{q}_{1,\mu}) = -\int_{\mathbb{R}^2}\tau(\hat{U}_1\circ\hat{q}_{1,\mu})\cdot\partial_\varepsilon(\hat{U}_1\circ\hat{q}_{1,\mu})$$
$$= -\int_{\Omega_1}\tau(u_1)\cdot\partial_\varepsilon u_1 + \mathrm{err}_3 + \mathrm{err}_4. \tag{5.30}$$

where we use (5.5) to bound

$$|\mathrm{err}_3| \lesssim |\int_{\Omega_0}\tau(\hat{U}_1\circ\hat{q}_{1,\mu})\cdot\partial_\varepsilon(\hat{U}_1\circ\hat{q}_{1,\mu})| \lesssim \mathcal{T}\int_{\Omega_0}|\nabla\hat{q}_{1,\mu}|^2 \lesssim \mu^{-1}\mathcal{T}\lesssim R_0$$

and set $\mathrm{err}_4 := \int_{\Omega_1}\partial_\varepsilon u_1\tau(u_1) - \partial_\varepsilon(\hat{U}_1\circ\hat{q}_{1,\mu})\tau(\hat{U}_1\circ\hat{q}_{1,\mu})$. To bound this term we can use that $|\partial_\varepsilon(u_1-\hat{U}_1\circ\hat{q}_{1,\mu})|\lesssim|\hat{q}_0|+|\partial_\varepsilon\hat{q}_0|$, while (1.10) ensures that

$$\begin{aligned}|\tau(u_1)-\tau(U_1\circ\hat{q}_{1,\mu})| &= |\tau(\hat{U}_1)\circ\hat{q}_\mu|\nabla\hat{q}_\mu|^2 - \tau(\hat{U}_1)\circ\hat{q}_{1,\mu}|\nabla\hat{q}_{1,\mu}|^2| \\ &\lesssim \mathcal{T}\cdot(|\hat{q}_0|\rho_q^2+|\nabla\hat{q}_0|\rho_q).\end{aligned}$$

Combined with (5.4) and (5.5) this yields

$$|\mathrm{err}_4| \lesssim \mathcal{T}\int_{\Omega_1}(|\hat{q}_0|+|\partial_\varepsilon\hat{q}_0|)\rho_q^2+|\nabla\hat{q}_0|\rho_q \lesssim \mu^{-1}\log\mu\mathcal{T}\lesssim R_0.$$

This reduces the proof of Lemma 3.4 to the proofs of the three lemmas stated above, which we carry out in the remainder of this section.

To prove Lemma 5.1, and later on also Lemmas 3.2 and 3.3, it is useful to expand

$$\mathfrak{z}_1 = \pi_N(v_1) = u_1 + P_{u_1}(\tilde{\gamma}_{U_1}+j_1) + \mathrm{err}_{u_1} = u_1 + \tilde{\gamma}_{U_1} + j_1 - P_{u_1}^{\perp}(\tilde{\gamma}_{U_1}+j_1) + \mathrm{err}_{u_1} \tag{5.31}$$

where we recall that u_1, $\tilde{\gamma}_{U_1}$ and j_1 are defined by (2.13), (2.16) and (2.17). We note that the resulting error term can be written as

$$\mathrm{err}_{u_1} = \int_0^1 \big(d\pi_N(u_1+t(\tilde{\gamma}_{U_1}+j_1)) - d\pi_N(u_1)\big)(\tilde{\gamma}_{U_1}+j_1)dt \tag{5.32}$$

and hence bounded by

$$|\mathrm{err}_{u_1}| \lesssim |j_1|^2 + |\tilde{\gamma}_{U_1}|^2 \lesssim |d\beta(0)|^2|\hat{q}_0|^2 + \delta^2\phi_\mu \tag{5.33}$$

since $d\pi_N(p)$ agrees with the projection P_p onto the tangent space T_pN whenever $p\in N$.

Proof of Lemma 5.1. Writing $\mathfrak{z}_1$ as in (5.31) and using that $\Delta j_1 = 0$ we find that

$$\begin{aligned}\int_{\Omega_1} \partial_\varepsilon \mathfrak{z}_1 \Delta \mathfrak{z}_1 &= \int_{\Omega_1} \partial_\varepsilon u_1 \, \Delta u_1 + \partial_\varepsilon (P_{u_1} j_1) \, \Delta u_1 - \partial_\varepsilon u_1 \, \Delta(P_{u_1}^\perp j_1) + \partial_\varepsilon \gamma \, \Delta\gamma + T_1 \\ &= \int_{\Omega_1} \partial_\varepsilon u_1 \, \tau(u_1) - P_{u_1} j_1 \partial_\varepsilon \Delta u_1 - \partial_\varepsilon u_1 \, \Delta(P_{u_1}^\perp j_1) + \partial_\varepsilon \gamma \, \Delta\gamma + T_1 + T_2\end{aligned} \tag{5.34}$$

for $T_2 := \int_{\Omega_1} \partial_\varepsilon (P_{u_1} j_1 \Delta u_1) = \int_{\Omega_1} \partial_\varepsilon (j_1 \tau(u_1))$ and

$$\begin{aligned}|T_1| \leq& \int_{\Omega_1} |\partial_\varepsilon (P_{u_1} \tilde{\gamma}_{U_1}) \, \Delta u_1| + |\partial_\varepsilon \mathrm{err}_{u_1}| \, |\Delta u_1| + |\partial_\varepsilon (\mathfrak{z}_1 - u_1) \, \Delta(P_{u_1}^\perp j_1)| \\ &+ \int_{\Omega_1} |\partial_\varepsilon \mathfrak{z}_1 \, \Delta \mathrm{err}_{u_1}| + |\partial_\varepsilon \mathfrak{z}_1 \Delta(P_{u_1}^\perp \tilde{\gamma}_{U_1})| + |\int_{\Omega_1} \partial_\varepsilon (\mathfrak{z}_1 - \gamma) \Delta\gamma| \\ =:& \, T_1^{(1)} + ... + T_1^{(6)}.\end{aligned} \tag{5.35}$$

We can rewrite the third term in (5.34) using integration by parts, which results in a boundary term of the form

$$B_1 := -\int_{\partial\Omega_1} \partial_n (P_{u_1}^\perp j_1) \cdot \partial_\varepsilon u_1 - P_{u_1}^\perp j_1 \cdot \partial_n \partial_\varepsilon u_1 dS, \tag{5.36}$$

to obtain the desired expression

$$\int_{\Omega_1} \partial_\varepsilon \mathfrak{z}_1 \Delta \mathfrak{z}_1 = \int_{\Omega_1} \partial_\varepsilon u_1 \, \tau(u_1) - \int_{\Omega_1} j_1 \partial_\varepsilon \Delta u_1 + \int_{\Omega_1} \partial_\varepsilon \gamma \Delta\gamma + T_1 + T_2 + B_1. \tag{5.37}$$

This reduces the proof of the lemma to showing that all error terms $T_1^{(i)}$, T_2 and B_1 obtained above are controlled by the quantity R_1 defined in (5.21).

To bound $T_1^{(1)}$ we consider $A(p)$ as a bilinear form $A(p) : \mathbb{R}^N \times \mathbb{R}^N \to T_p^\perp N$ on the whole space which vanishes in normal directions and write

$$\begin{aligned}\partial_\varepsilon (P_{u_1} \tilde{\gamma}_{U_1}) A(u_1)(\nabla u_1, \nabla u_1) &= -P_{u_1}(\tilde{\gamma}_{U_1}) \partial_\varepsilon \big(A(u_1)(\nabla u_1, \nabla u_1) \big) \\ &= -P_{u_1}(\tilde{\gamma}_{U_1}) \cdot (\partial_\varepsilon A(u_1))(\nabla u_1, \nabla u_1).\end{aligned}$$

Combined with (1.10), (5.9) and (5.16) we can hence bound

$$\begin{aligned}T_1^{(1)} &\lesssim \int |\partial_\varepsilon (P_{u_1}(\tilde{\gamma}_{U_1})| |\tau(u_1)| + \int |P_{u_1}(\tilde{\gamma}_{U_1})| |(\partial_\varepsilon A(u_1))(\nabla u_1, \nabla u_1)| \\ &\lesssim \big[(\delta + \eta_0)\mathcal{T} + \delta \big] \int_{A \cap \Omega_1} (\phi_\mu + c_\mu) |\nabla \hat{q}_\mu|^2 \lesssim (\eta_0 \mathcal{T} + \delta)(\log\mu)^{-1} f_\mu{}^{-2},\end{aligned}$$

where we use (5.6) and (5.7) in the last step. Hence $T_1^{(1)} \lesssim R_1$ as required.

Next we use that the formula (5.32) for err_{u_1}, combined with (5.16) and (5.18), gives

$$\begin{aligned}|\partial_\varepsilon \mathrm{err}_{u_1}| &\lesssim |j_1|^2 + |\tilde\gamma_{U_1}|^2 + (|j_1| + |\tilde\gamma_{U_1}|)(|\partial_\varepsilon j_1| + |\partial_\varepsilon \tilde\gamma_{U_1}|) \\ &\lesssim |d\beta(0)|^2|\hat q_0|^2 + \delta^2\phi_\mu + (|d\beta(0)||\hat q_0| + \delta\phi_\mu)\big[|d\beta(0)||\partial_\varepsilon \hat q_0| \\ &\quad + |\partial_\varepsilon d\beta(0)||\hat q_0| + \eta_0\phi_\mu + c_\mu\delta\mathbb{1}_A\big].\end{aligned} \tag{5.38}$$

Using additionally (5.3), (5.6) and (5.7) we hence see that

$$\begin{aligned}T_1^{(2)} \lesssim &|d\beta(0)|^2(I_1+I_2) + |d\beta(0)||\partial_\varepsilon d\beta(0)|I_1 + (\log\mu)^{-1}{f_\mu}^{-3}\delta + \eta_0|d\beta(0)|I_3 \\ &+ \eta_0\delta(\log\mu)^{-2}{f_\mu}^{-2} + \delta^2(\log\mu)^{-1}{f_\mu}^{-2}\end{aligned}$$

and thus that $T_1^{(2)} \lesssim R_1$.

We then split

$$T_1^{(3)} \le \int_{\Omega_1} |P_{u_1}^\perp(\partial_\varepsilon(\mathfrak{z}_1 - u_1))||\Delta P_{u_1}^\perp j_1| + \int_{\Omega_1} |\partial_\varepsilon(\mathfrak{z}_1 - u_1)||P_{u_1}(\Delta P_{u_1}^\perp j_1)|.$$

As $|\Delta u_1| + |\nabla u_1|^2 \lesssim |\nabla \hat q_\mu|^2 \lesssim \rho_q^2$ and as $\Delta j_1 = 0$ we can certainly estimate

$$|\Delta(P_{u_1}^\perp j_1)| \lesssim |j_1|(|\nabla u_1|^2 + |\Delta u_1|) + |\nabla j_1||\nabla u_1| \lesssim |d\beta(0)|(|\hat q_0|\rho_q + |\nabla \hat q_0|)\rho_q.$$

To obtain an improved bound on the tangential part of this quantity we write $P_{u_1}^\perp j_1 = \sum_j \langle j_1, \nu_{u_1}^j\rangle \nu_{u_1}^j$ to see that $|P_{u_1}(\Delta(P_{u_1}^\perp j_1)| \lesssim |j_1|\rho_q^2 + \rho_q|P_{u_1}^\perp(\nabla j_1)|$, and write

$$j_1 = dU_0(0)(q_0) - dU_1(0)(q_0) = P_{U_1(0)}(j_1) + (P_{U_0(0)} - P_{U_1(0)})(dU_0(0)(q_0)) \tag{5.39}$$

to see that

$$\begin{aligned}|P_{u_1}^\perp(\nabla j_1)| &\lesssim |u_1 - U_1(0)||\nabla j_1| + |U_0(0) - U_1(0)||\nabla \hat q_0| \\ &\lesssim |d\beta(0)|(|\hat q_0| + |\hat q_{1,\mu}|)|\nabla \hat q_0| + \delta|\nabla \hat q_0|.\end{aligned} \tag{5.40}$$

Combined we hence get

$$|P_{u_1}(\Delta(P_{u_1}^\perp j_1))| \lesssim |d\beta(0)||\hat q_0|\rho_q^2 + |d\beta(0)||\hat q_{1,\mu}||\nabla \hat q_0|\rho_q + \delta|\nabla \hat q_0|\rho_q.$$

On the other hand, while we can only bound

$$\begin{aligned}|\partial_\varepsilon(\mathfrak{z}_1 - u_1)| &= |\partial_\varepsilon(P_{u_1}(\tilde\gamma_{U_1} + j_1) + \mathrm{err}_{u_1})| \lesssim |j_1| + |\partial_\varepsilon j_1| + |\tilde\gamma_{U_1}| + |\partial_\varepsilon \tilde\gamma_{U_1}| \\ &\lesssim (|d\beta(0)| + |\partial_\varepsilon d\beta(0)|)|\hat q_0| + |d\beta(0)||\partial_\varepsilon \hat q_0| + \delta(\phi_\mu + c_\mu\mathbb{1}_A) + \eta_0\phi_\mu,\end{aligned}$$

compare (5.38) and (5.16), we get a stronger bound of

$$|P_{u_1}^\perp(\partial_\varepsilon(\mathfrak{z}_1 - u_1)| = |(P_{u_1}^\perp - P_{\mathfrak{z}_1}^\perp)(\partial_\varepsilon \mathfrak{z}_1)| \leq |u_1 - \mathfrak{z}_1||\partial_\varepsilon \mathfrak{z}_1|$$
$$\lesssim |j_1| + |\tilde{\gamma}_{U_1}| \lesssim |d\beta(0)||\hat{q}_0| + \delta\phi_\mu.$$

Combined we hence get

$$\begin{aligned}
T_1^{(3)} &\lesssim \int_{\Omega_1} |d\beta(0)||\hat{q}_0||\Delta P_{u_1}^\perp j_1| + \delta \int_{A\cap\Omega 1} (\phi_\mu + c_\mu)|\Delta P_{u_1}^\perp j_1| \\
&\quad + \int_{\Omega_1} (|\partial_\varepsilon d\beta(0)||\hat{q}_0| + |d\beta(0)||\partial_\varepsilon \hat{q}_0|)|P_{u_1}(\Delta P_{u_1}^\perp j_1)| + \eta_0 \int_{A\cap\Omega_1} \phi_\mu |P_{u_1}(\Delta P_{u_1}^\perp j_1)| \\
&\lesssim |d\beta(0)|^2(I_1 + I_2) + |d\beta(0)||\partial_\varepsilon d\beta(0)|I_1 \\
&\quad + \delta(|\partial_\varepsilon d\beta(0)| + |d\beta(0)| + \eta_0)\mu^{-\frac{1}{2}} + \eta_0|d\beta(0)|I_3
\end{aligned}$$

and thus $T_1^{(3)} \lesssim R_1$. Here we simply use that $\|\hat{q}_0\|_{L^\infty(\Omega_1)} + \|\nabla \hat{q}_0\|_{L^2(\Omega_1)} \lesssim \mu^{-\frac{1}{2}}$ to deal with the terms involving δ as these only give lower order contributions.

Next we estimate

$$T_1^{(4)} \leq \frac{1}{2} \int_{\Omega_1} |P_{\mathfrak{z}_1}(d^2\pi_N(u_1)(\nabla j_1, \nabla j_1))| + \int_{\Omega_1} |\Delta \mathrm{err}_{u_1} - \tfrac{1}{2} d^2\pi_N(u_1)(\nabla j_1, \nabla j_1)|.$$

To deal with the first term we exploit that $A(x) = -dP_x$ and thus that

$$d^2\pi(x)(v, w) = -A(x)(v, w) \in T_x^\perp N \text{ for all } x \in N, v, w \in T_x N. \tag{5.41}$$

Combined with (5.40) this allows us to bound

$$\begin{aligned}
|P_{\mathfrak{z}_1}(d^2\pi_N(u_1)(\nabla j_1, \nabla j_1))| &\lesssim |\mathfrak{z}_1 - u_1||\nabla j_1|^2 + |P_{u_1}^\perp \nabla j_1||\nabla j_1| \\
&\lesssim |d\beta(0)|^2(|\hat{q}_0| + |\hat{q}_{1,\mu}| + |\tilde{\gamma}_{U_1}|))|\nabla \hat{q}_0|^2
\end{aligned}$$

while the formula (5.32) for err_{u_1} allows us to check that

$$\begin{aligned}
&|\Delta \mathrm{err}_{u_1} - \tfrac{1}{2} d^2\pi_N(u_1)(\nabla j_1, \nabla j_1)| \\
&\quad \lesssim (|j_1| + |\tilde{\gamma}_{U_1}|)\big[|\nabla u_1|(|\nabla j_1| + |\nabla \gamma|) + |\nabla j_1|^2 + |\nabla \gamma|^2 + |\tau(\gamma)|\big] \\
&\qquad + (|j_1|^2 + |\tilde{\gamma}_{U_1}|^2)\rho_q^2 + |\nabla \gamma|^2 + |\nabla \gamma||\nabla j_1|.
\end{aligned}$$

Combined we hence get

$$\begin{aligned}
T_1^{(4)} &\lesssim |d\beta(0)|^2(I_1 + I_2) + \delta|d\beta(0)|\mu^{-\frac{1}{2}} + \delta^2(\|{\phi_\mu}^2\rho_q^2\|_{L^1(\Omega_1)} + \|\nabla \phi_\mu\|_{L^2}^2) \\
&\quad + (\|j_1\|_{L^\infty(\Omega_1)} + \delta)\|\tau(\gamma)\|_{L^1} \\
&\lesssim |d\beta(0)|^2(I_1 + I_2) + \delta|d\beta(0)|\mu^{-\frac{1}{2}} + \delta^2(\log \mu)^{-1}
\end{aligned}$$

where we also use (5.6), (5.7), (5.8) and (5.12) in the second step. Thus also $T_1^{(4)} \lesssim R_1$.

Next we estimate

$$\begin{aligned} T_1^{(5)} &\lesssim \int_{A\cap\Omega_1} |P_{\mathfrak{z}_1}(\Delta(P_{u_1}^{\perp}\tilde{\gamma}_{U_1}))| \\ &\lesssim \int_{A\cap\Omega_1} |u_1-\mathfrak{z}_1|(|\Delta\gamma|+|\nabla\gamma|\rho_q)+|P_{u_1}^{\perp}(\nabla\gamma)|\rho_q+|\tilde{\gamma}_{U_1}|\rho_q^2 \\ &\lesssim (\mu^{-\frac{1}{2}}+\delta)\delta(\log\mu)^{-1}+\int_{A\cap\Omega_1} |u_1-\gamma||\nabla\gamma|\rho_q+\delta(\log\mu)^{-1}f_\mu{}^{-2} \end{aligned}$$

where we used (5.12), $|u_1-\mathfrak{z}_1|\le \mu^{-\frac{1}{2}}+\delta$ and (5.7) in the last step. Using furthermore that $|u_1-\gamma|\lesssim |u_1-U_1(0)|+|\tilde{\gamma}_{U_1}|\le f_\mu{}^{-1}+\delta\phi_\mu$ on A we hence get

$$T_1^{(5)} \lesssim \mu^{-\frac{1}{2}}\delta+\delta^2(\log\mu)^{-1}+\delta(\log\mu)^{-1}f_\mu{}^{-2}\lesssim R_1. \tag{5.42}$$

Finally, to bound $T_1^{(6)}$ we write $\mathfrak{z}_1$ instead as

$$\mathfrak{z}_1=\gamma+P_\gamma(\tilde{u}_1+j_1)+\mathrm{err}_\gamma \text{ for } \mathrm{err}_\gamma=\int_0^1 \big(d\pi_N(\gamma+t(j_1+\tilde{u_1}))-d\pi_N(\gamma)\big)(\tilde{u_1}+j_1)$$

and use (5.6) and (5.16) to see that

$$|\partial_\varepsilon(\mathfrak{z}_1-\gamma)|\lesssim |\partial_\varepsilon\gamma|f_\mu{}^{-1}+f_\mu{}^{-2}\lesssim(\eta_0+\delta(\log\mu)^{-1})f_\mu{}^{-1}+f_\mu{}^{-2} \text{ on } A.$$

As $\|\Delta\gamma\|_{L^1}\lesssim\delta(\log\mu)^{-1}$ we hence immediately deduce that

$$T_1^{(6)}\le\delta(\log\mu)^{-1}f_\mu{}^{-2}+\delta^2(\log\mu)^{-2}f_\mu{}^{-1}+\eta_0\delta(\log\mu)^{-1}f_\mu{}^{-1}\lesssim R_1.$$

This completes the proof that the term T_1 obtained in (5.34) is controlled by R_1 and it hence remains to show that the term T_2 obtained in this estimate (5.34) and the boundary term B_1 defined in (5.36) also have this property.

Since T_2 can be written as $T_2=\int_{\Omega_1}\partial_\varepsilon(j_1\tau(\hat{U}_1)\circ\hat{q}_\mu|\nabla\hat{q}_\mu|^2)$, compare (1.10), we furthermore have

$$\begin{aligned} |T_2| &\lesssim \mathcal{T}\int|\partial_\varepsilon j_1|\rho_q^2+|j_1|(|\partial_\varepsilon\hat{q}_\mu|\rho_q^2+\rho_q|\partial_\varepsilon\nabla\hat{q}_\mu|)+\eta_{\mathcal{T}}\int|j_1||\nabla\hat{q}_{1,\mu}|^2 \\ &\lesssim \mathcal{T}|d\beta(0)|(I_3+I_4)+\mathcal{T}|\partial_\varepsilon d\beta(0)|I_3+\eta_{\mathcal{T}}|d\beta(0)|I_3\lesssim R_1. \end{aligned}$$

Finally, to estimate B_1 we use (5.39) to see that

$$|P_{u_1}^{\perp}j_1|\lesssim|\hat{q}_\mu||j_1|+\delta|\hat{q}_0|\lesssim|d\beta(0)||\hat{q}_\mu||\hat{q}_0|+\delta|\hat{q}_0|. \tag{5.43}$$

Since $|\nabla u_1|+|\partial_\varepsilon \nabla u_1|\lesssim 1$ on $\partial\Omega_1$ and hence $|P_{u_1}(\partial_n P_{u_1}^\perp j_1)|\lesssim |\partial_n u_1||P_{u_1}^\perp j_1|\lesssim |P_{u_1}^\perp j_1|$ we can hence bound

$$|B_1|\lesssim \int_{\partial\Omega_1} |P_{u_1}^\perp j_1| dS \lesssim |d\beta(0)| I_{\partial\Omega}+\mu^{-1}\delta, \tag{5.44}$$

where the last step follows since $|\hat{q}_0|$ and $|\partial\Omega_1|$ are of order $O(\mu^{-\frac{1}{2}})$.

We have thus shown that all error terms are controlled by R_1 which completes the proof of Lemma 5.1. $\square$

Proof of Lemma 5.2. To prove this lemma we will use that

$$|\hat{q}_0(z)|+|z||\nabla\hat{q}_0(z)|\lesssim \sum_{j=1}^{n_0^*}|a_j(q_0)||z|^j \text{ for every } z\in\mathbb{C} \tag{5.45}$$

and for $q_0\in\mathcal{R}^\sigma(q_0^*)$, $\hat{q}_0=\pi_{p^*}^{-1}\circ q_0$. To see this we note that our choice of σ ensures that the right hand side is at least $\frac{1}{2}|z|^{n_0^*}$ and that (5.45) is hence trivially satisfied outside of any fixed size disc $\mathbb{D}_{\bar{r}}(0)$ since the maps $\hat{q}_0$ satisfy uniform C^k bounds. As $q_0^*(0)=0$ and as the maps $\hat{q}_0$ are uniformly close to $\hat{q}_0^*$ we can fix $\bar{r}=\bar{r}(q_0^*)>0$ small enough so that $|q_0(z)|\leq 1$ for all $|z|\leq 2\bar{r}$. Standard estimates from complex analysis then ensure that $|a_j(q_0)|\lesssim (2\bar{r})^{-j}$ for all j and hence that $\sum_{j\geq n_0^*+1} j|a_j(q_0)||z|^j\lesssim |z|^{n_0^*}$ for all $|z|\leq\bar{r}$ which ensures that (5.45) also holds on $\mathbb{D}_{\bar{r}}(0)$.

Similarly, given $q_1\in\mathcal{R}^\sigma(q_i^*)$ and $\mu>\bar{\mu}$ we can use that $\hat{q}_{1,\mu}=\pi_{p^*}\circ q_1(\frac{1}{\mu\cdot})$ is so that

$$|\hat{q}_{1,\mu}(z)|\lesssim \sum_{k=1}^{n_1^*}\tfrac{|a_k(q_1)|}{1+|\mu z|^k} \text{ and } |\nabla\hat{q}_{1,\mu}(z)|\lesssim \mu\sum_{k=1}^{n_1^*}\tfrac{|a_k(q_1)|}{1+|\mu z|^{k+1}}. \tag{5.46}$$

The proofs of all claims made in the lemma now follow from short explicit calculation based these two estimates (5.45) and (5.46). To begin with we note that

$$\begin{aligned}\int_{\Omega_1}|\hat{q}_0|^2|\nabla\hat{q}_{1,\mu}|^2+|\hat{q}_{1,\mu}|^2|\nabla\hat{q}_0|^2 &\lesssim \sum|a_j(q_0)|^2|a_k(q_1)|^2\mu^{-2j}\mu^2\int_{\Omega_1}\tfrac{|\mu z|^{2j}}{1+|\mu z|^{2k+2}}+\tfrac{|\mu z|^{2j-2}}{1+|\mu z|^{2k}}\\ &\lesssim \sum|a_j(q_0)|^2|a_k(q_1)|^2\mu^{-2j}\int_0^{\sqrt{\mu}}\tfrac{t^{2j-1}}{1+t^{2k}}dt\\ &\lesssim \sum|a_j(q_0)|^2|a_k(q_1)|^2(\mu^{-2j}+\mu^{-j}\mu^{-k}\log\mu)\leq\bar{\nu}^2\log\mu\end{aligned} \tag{5.47}$$

where here and in the following we only need to sum over $j=1,\dots,n_0^*$ and $k=1,\dots,n_1^*$ and can use that the corresponding coefficients $a_j(q_i)$ are uniformly bounded as we only consider elements of $\mathcal{R}^\sigma(q_i^*)$. As we furthermore have

$$\int_{\Omega_1} |\hat{q}_0|^{2\alpha}|\nabla \hat{q}_0|^2 \lesssim \sum_j |a_j(q_0)|^{2(1+\alpha)} \int_{\Omega_1} |z|^{2(1+\alpha)j-2} \lesssim \bar{\nu}^{1+\alpha} \text{ for any } \alpha \geq 0, \tag{5.48}$$

so in particular for $\alpha = 1$, we obtain the claimed estimate (5.22) on I_1.

Using (5.45), (5.46) and (5.4) we then get

$$\begin{aligned}\int_{\Omega_1} |\hat{q}_0||\partial_\varepsilon \hat{q}_0||\nabla \hat{q}_{1,\mu}|^2 + |\partial_\varepsilon \hat{q}_0||\hat{q}_{1,\mu}||\nabla \hat{q}_{1,\mu}||\nabla q_{1,\mu}| &\lesssim \sum |a_j(q_0)||a_k(q_1)|^2 \mu^{-j-1} \int_0^{\sqrt{\mu}} \tfrac{t^{j+1}}{1+t^{2k+1}} dt \\ &\lesssim \bar{\nu}\mu^{-1} \log \mu.\end{aligned}$$

As the other contributions to I_2 are bounded by

$$\int_{\Omega_1} |\hat{q}_{1,\mu}||\nabla \hat{q}_0|^2 \lesssim \sum |a_k(q_1)||a_j(q_0)|^2 \mu^{-2j} \int_0^{\sqrt{\mu}} \tfrac{t^{2j-1}}{1+t^k} \lesssim \bar{\nu}^{3/2} \lesssim \mu^{-\frac{1}{2}} \bar{\nu}$$

and

$$\int |\hat{q}_0|\rho_q|\nabla \hat{q}_0| \leq (\bar{\nu}^2 \log \mu)^{\frac{1}{2}} \bar{\nu}^{\frac{1}{2}} \lesssim \mu^{-1} (\log \mu)^{\frac{1}{2}} \bar{\nu},$$

compare (5.47) and (5.48), we hence obtain the claimed bound (5.23) on I_2.

Combining (5.48) with

$$\begin{aligned}\int_{\Omega_1} |\hat{q}_0||\nabla \hat{q}_{1,\mu}|^2 + |\nabla \hat{q}_0||\hat{q}_{1,\mu}||\nabla \hat{q}_{1,\mu}| &\lesssim \sum |a_j(q_0)||a_k(q_1)|^2 \mu^{-j} \int_0^{\sqrt{\mu}} \tfrac{t^{j+1}}{1+t^{2k+2}} \\ &\lesssim \sum |a_j(q_0)||a_k(q_1)|^2 (\mu^{-j} + \mu^{-j/2}\mu^{-k} \log \mu) \lesssim \bar{\nu}\end{aligned}$$

then shows that I_3 satisfies (5.24), while (5.4) and (5.5) give the claimed bound on

$$I_4 \leq \int_{\Omega_1} |z|(1 + \tfrac{\mu^2}{1+|\mu z|^4}) \lesssim \mu^{-\frac{3}{2}} + \mu^{-1} \int_0^{\sqrt{\mu}} \tfrac{t^2}{1+t^4} dt \lesssim \mu^{-1}.$$

Finally, as $|\hat{q}_0|^2 + |\hat{q}_{1,\mu}|^2 \lesssim \bar{\nu}$ on $\partial\Omega_1 = \partial\mathbb{D}_{\mu^{-\frac{1}{2}}}$, we get that $|I_{\partial\Omega_1}| \leq |\partial\Omega_1|\bar{\nu} \lesssim \bar{\nu}\mu^{-\frac{1}{2}}$. $\square$

In Section 6 we will also use that a very similar calculation gives

$$\int |\partial_\varepsilon \hat{q}_0|^2 \rho_q^2 + |\hat{q}_0|^2 (|\partial_\varepsilon \nabla \hat{q}_0|^2 + |\partial_\varepsilon \nabla \hat{q}_{1,\mu}|^2) \lesssim \eta_{rat} (\mu^{-j^*} + \bar{\nu})^2 \log \mu. \tag{5.49}$$

Having hence completed the proofs of both Lemmas 5.1 and 5.2, we finally turn to the proof that the contribution of the error terms $\mathrm{err}_\varphi = \mathfrak{z} - \mathfrak{z}_1$ to $\partial_\varepsilon E(\mathfrak{z}_\varepsilon)$ is controlled as described in Lemma 5.3.

As err_φ is supported on $\hat{A}$ (defined in (2.10)) we first collect some estimates that are valid on this set where $|z| \sim \mu^{-\frac{1}{2}}$ and where we can hence bound

$$|q_0| + |q_{1,\mu}| + \mu^{-\frac{1}{2}}(|\nabla q_0| + |\nabla q_{1,\mu}|) \lesssim \bar{\nu}^{\frac{1}{2}} \tag{5.50}$$

and

$$|\partial_\varepsilon q_0| + |\partial_\varepsilon q_{\mu,1}| + \mu^{-\frac{1}{2}}|\partial_\varepsilon \nabla q_0| + |\partial_\varepsilon \nabla q_{\mu,1}| \lesssim \mu^{-\frac{1}{2}}. \tag{5.51}$$

Since $v_0 - v_1 = \beta(q_\mu) - \beta(0) - d\beta(0)(q_\mu)$ we have

$$|v_0 - v_1| + \mu^{-\frac{1}{2}}|\nabla(v_0 - v_1)| + \mu^{-1}|\Delta(v_0 - v_1)| \lesssim \|\beta\|(|q_\mu|^2 + \mu^{-1}|\nabla q_\mu|^2) \leq \|\beta\|\bar{\nu} \tag{5.52}$$

as well as

$$|\partial_\varepsilon(v_0 - v_1)| + \mu^{-\frac{1}{2}}|\partial_\varepsilon \nabla(v_0 - v_1)| + \mu^{-1}|\partial_\varepsilon \Delta(v_0 - v_1)| \leq (\|\beta\| + \|\partial_\varepsilon \beta\|)\bar{\nu} + \eta_{rat}\|\beta\|\mu^{-j^*}. \tag{5.53}$$

We furthermore note that γ is harmonic on $\hat{A}$ so we can bound

$$|\nabla v_i|^2 + |\Delta v_i| \lesssim \rho_q^2 + |\nabla \gamma|^2 \lesssim \mu(\bar{\nu} + \tfrac{\delta^2}{(\log \mu)^2}). \tag{5.54}$$

Since the error term is given by

$$\mathrm{err}_\varphi = \pi_N(v_1 + t\varphi(v_0 - v_1)) - \pi_N(v_1) = \int_0^1 d\pi_N(v_1 + t\varphi(v_0 - v_1))(\varphi(v_0 - v_1))dt \tag{5.55}$$

for a function φ that vanishes for $|z| \leq \frac{1}{2}\mu^{-\frac{1}{2}}$ and that satisfies $|\nabla \varphi|^2 + |\Delta \varphi| \lesssim \mu$ we have

$$|\mathrm{err}_\varphi| \lesssim \|\beta\|\bar{\nu} \qquad \text{and} \qquad |\Delta \mathrm{err}_\varphi| \lesssim \|\beta\|\mu\bar{\nu}. \tag{5.56}$$

As $|\partial_\varepsilon \mu_i| \lesssim \mu_i$ we have

$$|\partial_\varepsilon(\varphi_\mu(z))| \lesssim \mu^{-\frac{1}{2}}|\partial_\varepsilon \tfrac{\mu_1}{\mu_0}||\phi'(\mu^{\frac{1}{2}}|z|)||z| \lesssim \mathbb{1}_{\hat{A}} \tag{5.57}$$

allowing us to bound

$$|\partial_\varepsilon \mathrm{err}_\varphi| \lesssim |v_0 - v_1| + |\partial_\varepsilon(v_0 - v_1)| \lesssim (\|\beta\| + \|\partial_\varepsilon \beta\|)\bar{\nu} + \eta_{rat}\|\beta\|\mu^{-j^*}. \tag{5.58}$$

We will later also use that the above estimates also allow us to bound

$$\mu^{-1}|\partial_\varepsilon \Delta \mathrm{err}_\varphi| \lesssim (\|\beta\| + \|\partial_\varepsilon \beta\|)\bar{\nu} + \eta_{rat}\|\beta\|\mu^{-j^*}. \tag{5.59}$$

Based on these estimates we can now complete the

Proof of Lemma 5.3. Since $|\partial_\varepsilon \mathfrak{z}_1| \lesssim \mu^{-\frac{1}{2}} + \delta + \eta_0$ on $\hat{A}$ and since $|\hat{A}| \lesssim \mu^{-1}$ we can bound

$$\begin{aligned} |\int_{\Omega_1} \partial_\varepsilon \mathfrak{z} \Delta \mathfrak{z} - \partial_\varepsilon \mathfrak{z}_1 \Delta \mathfrak{z}_1| &\leq \int_{\hat{A}\cap\Omega_1} (|\partial_\varepsilon \mathfrak{z}_1| + |\partial_\varepsilon \mathrm{err}_\varphi|)|\Delta \mathrm{err}_\varphi| + \int_{\hat{A}\cap\Omega_1} |\partial_\varepsilon \mathrm{err}_\varphi \Delta \mathfrak{z}_1| \\ &\lesssim \|\beta\|(\mu^{-\frac{1}{2}} + \delta + \eta_0)\bar{\nu} + \int_{\hat{A}\cap\Omega_1} |\partial_\varepsilon \mathrm{err}_\varphi \cdot \Delta \mathfrak{z}_1|. \end{aligned}$$

To bound this last term we split $\partial_\varepsilon \mathrm{err}_\varphi \Delta \mathfrak{z}_1$ into the contributions of the tangential and the normal parts. Since $P_{\mathfrak{z}_1} = d\pi(\mathfrak{z}_1)$ and since $|v_1 - \mathfrak{z}_1| \leq \delta + \|\beta\|\bar{\nu}^{\frac{1}{2}}$ and $|\partial_\varepsilon v_1| \lesssim \eta_0 + \mu^{-\frac{1}{2}}$ on $\hat{A}$ we get

$$\begin{aligned} |P_{\mathfrak{z}_1}^\perp(\partial_\varepsilon \mathrm{err}_\varphi)| &\lesssim (|v_1 - \mathfrak{z}_1| + |v_0 - v_1|)|\partial_\varepsilon(\varphi(v_0 - v_1))| + |\partial_\varepsilon v_1||\varphi(v_0 - v_1)| \\ &\leq (\delta + \|\beta\|\bar{\nu}^{\frac{1}{2}})(\|\beta\| + \|\partial_\varepsilon \beta\|)\mu^{-1} + (\eta_0 + \mu^{-\frac{1}{2}})\bar{\nu}\|\beta\| \\ &\lesssim \|\beta\|\mu^{-3/2} + \delta(\|\partial_\varepsilon \beta\| + \|\beta\|)\mu^{-1} + \eta_0\|\beta\|\bar{\nu} \end{aligned} \tag{5.60}$$

and can thus in particular bound $|P_{\mathfrak{z}_1}^\perp(\partial_\varepsilon \mathrm{err}_\varphi)| \lesssim \mu^{-1}(\|\beta\| + \delta)$.

Since (5.54) ensures that $|\Delta \mathfrak{z}_1| = |\Delta(\pi_N(v_1))| \lesssim \mu(\bar{\nu} + \delta^2)$ on $\hat{A}$ and since $|\hat{A}| \lesssim \mu^{-1}$ this suffices to estimate

$$\int |P_{\mathfrak{z}_1}^\perp \partial_\varepsilon \mathrm{err}_\varphi||\Delta \mathfrak{z}_1| \leq (\bar{\nu} + \delta^2)\mu^{-1}\left[\|\beta\| + \delta\right]$$

and hence to see that this term is bounded by the right hand side of (5.27).

From (5.31) we furthermore get that

$$\begin{aligned} |P_{u_1}(\Delta \mathfrak{z}_1)| &\lesssim |\tau(u_1)| + |\nabla \gamma|^2 + |\nabla j_1|^2 + \rho_q(|\nabla \gamma| + |\nabla j_1|) \\ &\leq \mathcal{T}\bar{\nu}\mu + \delta^2\mu + \delta\bar{\nu}^{\frac{1}{2}}\mu + \|\beta\|\bar{\nu}\mu. \end{aligned}$$

As (5.58) ensures that $|\partial_\varepsilon \mathrm{err}_\varphi| \lesssim (\|\beta\| + \|\partial_\varepsilon \beta\|)\mu^{-1}$ and as $|\hat{A}| \lesssim \mu^{-1}$ we immediately conclude that also $\int |\partial_\varepsilon \mathrm{err}_\varphi||P_{u_1}\Delta \mathfrak{z}_1|$ is controlled by the right hand side of (5.27), which completes the proof of the lemma. □

5.3. Dominating terms in the energy expansions for variations of the rational maps

In this section we explain how the general energy expansion proven in the previous section can be used to identify a variation of type (V2) for which Lemma 3.5 holds true.

By symmetry it suffices to prove the claim of the lemma for $i = 1$, i.e. to identify a variation of q_1 so that (3.16) holds for the resulting family $\mathfrak{z}_\varepsilon$ for which we keep $U_{0.1}$, q_0 and μ fixed.

For such variations we know from Lemma 3.4 that

$$dE(\mathfrak{z})(\partial_\varepsilon \mathfrak{z}) = \int\limits_{\Omega_1} j_1 \partial_\varepsilon \Delta u_1 + \int\limits_{\Omega_0} j_0 \partial_\varepsilon \Delta u_0 + \mathrm{err}_1 \tag{5.61}$$

now for an error that is bounded by $|\mathrm{err}_1| \lesssim R_2$ for

$$R_2 := \bar{\nu}\|\beta\|\mu^{-\frac{1}{2}} + \tfrac{\delta^2}{\log\mu} + \tfrac{\delta}{f_\mu^2 \log\mu} + \mathcal{T}\mu^{-1}\log\mu \tag{5.62}$$

as $\eta_0 = \eta_\mathcal{T} = 0$. As $\partial_\varepsilon U_0 = 0$ and as $|\partial_\varepsilon \hat{q}_\mu| \lesssim |\partial_\varepsilon q_{1,\mu}|$ is small on Ω_0 we can easily check that

$$|\int\limits_{\Omega_0} j_0 \partial_\varepsilon \Delta u_0| \lesssim \bar{\nu}\|\beta\|\mu^{-\frac{1}{2}} \leq R_2. \tag{5.63}$$

To prove the lemma it hence suffices to show that we can find a variation of q_1 for which

$$\int\limits_{\Omega_1} j_1 \partial_\varepsilon \Delta u_1 \geq c_1 |d\beta(0)|^2 \nu_0 - \mathrm{err}_2 \tag{5.64}$$

for an error term that is bounded by $|\mathrm{err}_2| \lesssim R_2$.

To this end we will use that away from $z = 0$ the functions $2(\mu z)^{-j}$ are well approximated by the first two components $h_{j,\mu}(z) := \frac{2(\mu\bar{z})^j}{1+|\mu\bar{z}|^{2j}}$ of the harmonic map $\bar{h}_{j,\mu}(z) := \pi((\mu\bar{z})^j) : \hat{\mathbb{C}} \to S^2$, where we continue to identify $(x, y) \in \mathbb{R}^2$ with $x + iy \in \mathbb{C}$ whenever convenient. We hence write $u_1(z) = U_1(q_0(z) + q_{1,\mu}(z))$ as

$$u_1(z) = U_1(0) + dU_1(0)\Big(\tfrac{1}{2}\sum_{j=1}^{n_1^*} a_j(q_1) h_{j,\mu}(z) + q_0(z)\Big) + f_1(z) + f_2(z) + f_3(z)$$

for

$$\begin{aligned}
f_1(z) &:= u_1(z) - (U_1(0) + dU_1(0))(q_\mu) = \int\limits_0^1 (dU_1(tq_\mu) - dU_1(0))(q_\mu)dt \\
f_2(z) &:= dU_1(0))\big(\sum_{1\leq j\leq n_0^*} a_j(q_1)((\mu z)^{-j} - \tfrac{1}{2}h_{j,\mu}(z)) \\
&= \tfrac{1}{2}dU_1(0)\big(\sum_{1\leq j\leq n_0^*} a_j(q_1)\tfrac{(\mu\bar{z})^j}{|\mu z|^{2j}(1+|\mu z|^{2j})}\big) \\
f_3(z) &:= dU_1(0)\big(\sum_{j\geq n_0^*+1} a_j(q_1)(\mu z)^{-j}\big).
\end{aligned}$$

As $\Delta j_1 = 0$ we get that $T_B^i := \int_{\Omega_1} j_1 \Delta \partial_\varepsilon f_i = \int_{\partial\Omega_1} j_1 \partial_n \partial_\varepsilon f_i - \partial_n j_1 \partial_\varepsilon f_i dS$ and will use this to show that all these terms are controlled by R_2.

We can use that $|\partial\Omega_1| \sim \mu^{-\frac{1}{2}}$ and that the rational functions are bounded by (5.50) and their variations by (5.51) at points with $|z| \sim \mu^{-\frac{1}{2}}$. This immediately implies that

$$|T_B^1| \lesssim |d\beta(0)| \int\limits_{\partial\Omega_1} |q_0|(|\nabla q_\mu||\partial_\varepsilon q_\mu| + |q_\mu||\nabla \partial_\varepsilon q_\mu|) + |\nabla q_0||\partial_\varepsilon q_\mu||q_\mu| \lesssim |d\beta(0)|\mu^{-\frac{1}{2}}\bar{\nu} \lesssim R_2,$$

while $|T_B^3| \lesssim |d\beta(0)|\mu^{-\frac{1}{2}}\nu_{\hat{q}_0} \lesssim R_2$ follows since $|\partial_\varepsilon f_3| + \mu^{-\frac{1}{2}}|\partial_\varepsilon \nabla f_3| \lesssim \mu^{-(n_0^*+1)/2} \lesssim \mu^{-\frac{1}{2}}\bar{\nu}^{\frac{1}{2}}$.

We can now use that terms of the form $d\beta(0)(a_j z^j)$, $a_j \in \mathbb{C}$, and their radial derivatives are $L^2(\partial\mathbb{D}_r)$ orthogonal to terms of the form $dU_1(0)(b_k(\bar{z})^k)$, $b_k \in \mathbb{C}$, and their radial derivatives on any circle $\partial\mathbb{D}_r = \partial\mathbb{D}_r(0)$ if $j \neq k$. This implies that

$$\begin{aligned}|T_B^2| &\lesssim |d\beta(0)| \sum_{j=1}^{n_0^*} |a_j(q_0)||a_j(\partial_\varepsilon q_1)| \int\limits_{\partial\Omega_1} \left[|z^j||\partial_n \tfrac{(\mu\bar{z})^j}{|\mu z|^{2j}(1+|\mu z|^{2j})}| + |\partial_n z^j||\tfrac{(\mu\bar{z})^j}{|\mu z|^{2j}(1+|\mu z|^{2j})}|\right] \\ &\lesssim |d\beta(0)| \sum_{j \le n_0^*} |a_j(q_0)||\partial_\varepsilon a_j(q_1)|\mu^{-2j} \lesssim |d\beta(0)|\mu^{-1}\bar{\nu} \lesssim R_2.\end{aligned}$$

At this stage we hence know that

$$\int\limits_{\Omega_1} j_1 \partial_\varepsilon \Delta u_1 = \tfrac{1}{2}\sum_{j=1}^{n_1^*} \int\limits_{\Omega_1} j_1 \cdot dU_1(0)(\partial_\varepsilon a_j(q_1)\Delta h_{j,\mu}) + \mathrm{err}_3 \tag{5.65}$$

for $|\mathrm{err}_3| \lesssim R_2$ since we know that $\Delta q_0 = 0$.

As $\bar{h}_{j,\mu}$ is a harmonic map into S^2 and as

$$|\nabla \bar{h}_{j,\mu}|^2(z) = |\nabla \pi((\mu\bar{z})^j)|^2 \cdot (j\mu)^2|\mu z|^{2(j-1)} = \mu^2 \tfrac{8j^2|\mu z|^{2j-2}}{(1+|\mu z|^{2j})^2}$$

we can write

$$\Delta h_{j,\mu} = -|\nabla \bar{h}_{j,\mu}|^2 h_{j,\mu} = -2\mu^2 H_j(\mu|z|) \cdot (\mu\bar{z})^j$$

for $H_j(t) := \frac{4j^2 t^{2j-2}}{(1+t^{2j})^3}$. The orthogonality of the different Fourier modes on circles hence allows us to write the main terms in (5.65) as

$$\begin{aligned}\tfrac{1}{2}\int\limits_{\Omega_1} j_1 \cdot dU_1(0)(\partial_\varepsilon a_j(q_1)\Delta h_{j,\mu}) &= -I_j^H(\mu)\int\limits_{S^1} d\beta(0)(a_j(q_0)e^{ij\theta}) \cdot dU_1(0)(a_j(\partial_\varepsilon q_1)e^{-ij\theta}) \\ &= -|a_j(q_0)||a_1(\partial_\varepsilon q_1)|I_j^H(\mu)I_j^\theta(\alpha_j, U_0, U_1)\end{aligned} \tag{5.66}$$

for $\alpha_j = \alpha_j(\partial_\varepsilon q_1, q_0) := j^{-1}\big[\mathrm{Arg}(a_j(\partial_\varepsilon q_1)) - \mathrm{Arg}(a_j(q_0))\big]$,

$$I_j^\theta(\alpha, U_0, U_1) := \int_{S^1} d(U_0 - U_1)(0)(e^{ij\theta})dU_1(0)(e^{-ij(\theta-\alpha)})d\theta \tag{5.67}$$

and

$$I_j^H(\mu) := \mu^2 \int_0^{\mu^{-\frac{1}{2}}} r^j H_j(\mu r)(\mu r)^j r dr = 4j^2\mu^{-j} \int_0^{\sqrt{\mu}} \tfrac{t^{4j-1}}{(1+t^{2j})^3} dt \sim \mu^{-j}. \tag{5.68}$$

We note that the above calculation applies for any variation of q_i that satisfies (3.8) and hence shows that $\partial_\varepsilon E(\mathfrak{z})$ is always controlled by (3.15).

To find a specific variation of q_1 for which Lemma 3.5 holds we now use the following lemma, a proof of which is included below.

Lemma 5.4. *Let $U_{0,1}^* : \hat{\mathbb{C}} \to N$ be distinct harmonic maps which satisfy Assumption 1. Then there exists $\alpha^* \in [0, 2\pi)$ and $c^* > 0$ so that $-I_j^\theta(\alpha^*, U_0^*, U_1^*) > 2c^*$ for all $j \in \mathbb{N}$ and hence so that*

$$I_j^\theta(\alpha^*, U_0, U_1) \geq c^* \text{ for all } U_i \in \mathcal{R}^\sigma(U_i^*),\ j \in \mathbb{N}.$$

Given a rational map q_0 we will apply this lemma for $j_0 \in \{1, \dots, n_0^*\}$ chosen so that

$$|a_j(q_0)|\mu^{-j}\Lambda_1^j \leq |a_{j_0}(q_0)|\mu^{-j_0}\Lambda_1^{j_0} \text{ for } j = 1, \dots, n_0^* \tag{5.69}$$

where $\Lambda_1 \geq 1$ is a large number that only depends on $U_{0,1}^*$ and $q_{0,1}^*$ and that is fixed below.

Writing q_1 in the form $q_1 = \frac{r_1}{r_2}$ for polynomials $r_{1,2}$ which are normalised by $r_2(0) = 1$, we finally define the desired variation of q_1 by

$$q_1^{(\varepsilon)}(z) := \frac{r_1(z) + \varepsilon(e^{i\theta^*}z)^{j_0}}{r_2(z)}$$

where we choose $\theta^* := \alpha^* + \frac{1}{j_0}\mathrm{Arg}(a_{j_0})$ so that $\alpha_{j_0}(\partial_\varepsilon q_1, q_0)$ is given by the number α^* from Lemma 5.4.

As $a_j(\partial_\varepsilon q_1) = 0$ for $j < j_0$ while $|a_{j_0}(\partial_\varepsilon q_1)| = 1$ we hence get from (5.65) and (5.66) that

$$\int_{\Omega_1} j_1 \partial_\varepsilon \Delta u_1 \geq c^* |a_{j_0}(q_0)| I^H_{j_0}(\mu) - C \sum_{j=j_0+1}^{n_0^*} |a_j(q_0)| I^H_j(\mu) - CR_2$$
$$\geq 2c|a_{j_0}(q_0)|\mu^{-j_0} - C \sum_{j=j_0+1}^{n_0^*} |a_j(q_0)|\mu^{-j} - CR_2$$
$$\geq (2c - C\Lambda_1^{-1})|a_{j_0}(q_0)|\mu^{-j_0} - CR_2 \geq c|a_{j_0}(q_0)|\mu^{-j_0} - CR_2$$

for a constant $c > 0$ that only depends on U_i^* and $q_{0,1}^*$, where the last step holds true provided $\Lambda_1 = \Lambda_1(U_i^*, q_i^*)$ in (5.69) is chosen sufficiently large.

Since (5.69) ensures that $\nu_{q_0} \lesssim |a_{j_0}(q_0)|\mu^{-j_0}$ this completes the proof of (5.64) and hence of Lemma 3.5 up to the

Proof of Lemma 5.4. We use that harmonic spheres are weakly conformal and that $dU_i^*(0) \neq 0$.

If $U_1^*(z) = U_0^*(\bar{z})$ we can hence choose coordinates on the target so that $dU_{0,1}^*(0)$ are given by the matrices with columns $\frac{1}{\sqrt{2}}|dU_0^*(0)|(1,0,0_{N-2})^T$ and $\frac{1}{\sqrt{2}}|dU_0^*(0)|(0,\pm 1, 0_{N-2})^T$ when viewed as maps from $\mathbb{R}^2$ to $\mathbb{R}^N$. Setting $\alpha^* := \pi$ hence gives

$$-I_j^\theta(\alpha^*, U_0^*, U_1^*) = \tfrac{1}{2}|dU_0^*(0)|^2 \int_{S^1} \sin^2(j\theta) + \cos^2(j\theta) = \pi |dU_0^*(0)|^2$$

which establishes the claim in this case.

So suppose instead that N is three dimensional and that $U_{0,1}^*$ are minimal spheres whose tangent spaces at $U_0^*(0) = U_1^*(0)$ intersect transversally. In this case we choose the coordinates on the domain so that $\partial_{x_1} U_1^*(0)$ is contained in the intersection of these tangent spaces and then use the conformality of U_1^* to choose coordinates on the target so that $\partial_{x_{1,2}} U_1^* = c_{U_1^*} e_{1,2}$ and so that $T_{U_0^*(0)}N = \mathbb{R}^3 \times \{0\}$. Here e_i is the standard basis of $\mathbb{R}^N$ and $c_{U_1^*} := \frac{1}{\sqrt{2}}|dU_1^*(0)|$.

The matrix $A = c_{U_0^*}^{-1} dU_0{}^*(0) : \mathbb{R}^2 \to \mathbb{R}^N$, whose columns are orthonormal, is hence so that $A_{31} = 0$, $A_{32} \neq 0$ and so that $A_{ij} = 0$ for $i \geq 4$. As

$$\int dU_1^*(0)(e^{ij\theta}) dU_1^*(0)(e^{-ij(\theta-\alpha^*)}) d\theta = \int \cos(j\theta)\cos(j\theta + \theta^*) - \sin(j\theta)\sin(j\theta + \alpha^*) d\theta$$
$$= 0,$$

for any α^* we hence get that

$$I_j^\theta(\alpha^*, U_0^*, U_1^*) = c_{U_1^*} c_{U_0^*} \int [A_{11}\cos(j\theta) + A_{12}\sin(j\theta)] \cos(j\theta + \alpha^*) d\theta$$
$$- c_{U_1^*} c_{U_0^*} \int [A_{21}\cos(j\theta) + A_{22}\sin(j\theta)] \sin(j\theta + \alpha^*) d\theta$$
$$= \pi c_{U_1^*} c_{U_0^*} [(A_{11} - A_{22})\cos\alpha^* - (A_{12} + A_{21})\sin\alpha^*].$$

If $A_{11} \neq A_{22}$ we can hence choose α^* to be either 0 or π. Conversely, if $A_{11} = A_{22}$ then we cannot have that also $A_{12} = -A_{21}$ as both columns of A have unit length and as $A_{31} = 0$ while $A_{32} \neq 0$. In this second case we can thus choose α^* as either $\frac{\pi}{2}$ or $\frac{3\pi}{2}$. $\square$

5.4. Dominating terms in the energy expansion for variations of the underlying maps U_i

We now want to prove that we can always find variations of the maps U_i in $\mathcal{H}_1(U_i^*)$ so that the induced variation in $\mathcal{Z}_0$ is as described in Lemmas 3.7 respectively 3.8.

For such a variation we can always bound $|\partial_\varepsilon \Delta u_1| \lesssim \rho_q^2$ and hence have

$$|\int_{\Omega_1} j_1 \partial_\varepsilon \Delta u_1| \lesssim |d\beta(0)| \int_{\Omega_1} |\hat{q}_0| \rho_q^2 \lesssim |d\beta(0)| I_3 \lesssim \|\beta\| \bar{\nu},$$

compare Lemma 5.2, while the analogue estimate for $|\int_{\Omega_0} j_0 \partial_\varepsilon \Delta u_0|$ follows by symmetry.

The general formula (3.11) for the variation of the energy on $\mathcal{Z}_0$ hence tells us that

$$dE(\mathfrak{z})(\partial_\varepsilon \mathfrak{z}) = \pi c_\mu \tfrac{d}{d\varepsilon} \delta(U_0, U_1)^2 + \deg(q_1) \tfrac{d}{d\varepsilon} E(U_1) + \deg(q_0) \tfrac{d}{d\varepsilon} E(U_0) + \mathrm{err}_4 \tag{5.70}$$

for $|\mathrm{err}_4| \lesssim \|\beta\| \bar{\nu} + \frac{\delta}{\log \mu} [\delta + f_\mu{}^{-1} + (\log \mu)^{-1}] + \mathcal{T} (\log \mu)^{-1} f_\mu{}^{-1}$.

As $c_\mu \sim (\log \mu)^{-1}$ we immediately get that Lemma 3.8 holds true if we choose a variation $U_{i,\varepsilon}$ in $\mathcal{H}_1^\sigma(U_i^*)$ for which (2.8) holds.

On the other hand, if we consider variations of U_i that are induced by translations on the domain then we know that $\frac{d}{d\varepsilon}|_{\varepsilon=0} E(U_i) = 0$ as the energy is conformally invariant. To prove Lemma 3.7 it hence suffices to prove Lemma 3.6, i.e. that there always exist variations of the form $U_{i,\varepsilon} = U_i(\cdot + \varepsilon e^{i\theta_0})$ so that

$$\tfrac{d}{d\varepsilon} \delta^2(U_0, U_1) \gtrsim \delta(U_0, U_1). \tag{5.71}$$

If we are in a non-degenerate setting for which $U_0^* = U_1^*$ then we know that $U_i(z) = U_0^*(z + c_i)$ for some $|c_i| \leq \sigma$. If $c_0 = c_1$ then $\delta = 0$ so there is nothing to prove, while for $c_0 \neq c_1$ we can consider $U_{1,\varepsilon}(z) = U_0^*(z + c_1 + \varepsilon \frac{c_0 - c_1}{|c_0 - c_1|}) = U_1(z + \varepsilon \frac{c_0 - c_1}{|c_0 - c_1|})$ and use that $dU_0^*(0) \neq 0$ to get a variation for which (5.71) holds. The analogue argument also applies in settings where $U_0^*(z) = U_1^*(\bar{z})$ and in both these cases we could just as well have constructed a variation of U_0.

So suppose that the maps $U_{0,1}^*$ are instead as in (A3) of Assumption 1. In this case we can exploit that the tangent spaces $T_{U_i^*(0)} U_i^*(\hat{\mathbb{C}})$, which are 2 dimensional subspaces of the same 3 dimensional space $T_{U_i^*(0)} N$, intersect transversally and that the length δ of the geodesic $\hat{\gamma}_{U_1, U_0}$ that connects the points $U_i(0)$ can be made smaller than any given constant by reducing σ.

We hence obtain that at least one of the angles $\alpha_i \in [0, \frac{\pi}{2}]$ at which $\hat{\gamma}$ intersects $T_{U_i(0)} U_i(\hat{\mathbb{C}})$ must be so that $|\frac{\pi}{2} - \alpha_i| \geq \frac{1}{3} \bar{\alpha}$, where $\bar{\alpha} \in (0, \frac{\pi}{2}]$ denotes the angle between the spaces $T_{U_i^*(0)}U_i^*(\hat{\mathbb{C}})$. For this i we hence know that $\|P^{T_{U_i(0)}U_i(\hat{\mathbb{C}})}(\frac{\hat{\gamma}'(i^*)}{\|\hat{\gamma}'(i^*)\|})\| = \cos(\alpha_i) \geq \sin(\bar{\alpha}/3)$ is bounded away from zero, where we recall that $\hat{\gamma}(i^*) = U_i(0)$, $0^* = 1, 1^* = 0$.

As U_i^* is weakly conformal and $dU_i^*(0) \neq 0$ and as the maps U_i are C^k close to U_i^* we hence deduce that $U_i(\cdot + \varepsilon e^{i\theta_0})$ has the desired property (5.71) if θ_0 is chosen so that $\partial_\varepsilon U_i(0)(e^{i\theta_0})$ points in the direction of $\pm P^{T_{U_i(0)}U_i(\hat{\mathbb{C}})}(\frac{\hat{\gamma}'(i^*)}{\|\hat{\gamma}'(i^*)\|})$.

This completes the proof of Lemma 3.6 and hence the proof of Lemma 3.7.

6. Estimates on the first and second variation of the energy

We conclude this paper with the proofs of the claims on the first and second variation of the energy at points $\mathfrak{z} \in \mathcal{Z}_0$ made in Lemmas 3.1-3.3. We will continue to work in stereographic coordinates which are scaled so that $|a_{n_0^*}(q_0)| = 1$ and hence so that $\mu_0 = 1$, but we now need to consider variations in general directions $w \in \Gamma(\mathfrak{z}^*TN)$. It is hence useful to observe that for any such w we can bound

$$\int_{S^1} |w(re^{i\theta})| d\theta \lesssim [\log(\min(r^{-1}, \mu r))]^{\frac{1}{2}} \|w\|_{\mathfrak{z}} \text{ for any } r \in [r_1, r_0] = [\mu^{-1} f_\mu, {f_\mu}^{-1}], \tag{6.1}$$

since both $\int_{S^1} |w(e^{i\theta})|$ and $\int_{S^1} |w(\mu^{-1}e^{i\theta})|$ are controlled by $\|w\|_{\mathfrak{z}}$ and since the integral of $\partial_r \int_{S^1} |w(re^{i\theta})| d\theta$ over any interval I can be bounded by $\|\nabla w\|_{L^2}\big(\int_I r^{-1}\big)^{\frac{1}{2}}$. As $\rho_{\mathfrak{z}} \lesssim 1 + \frac{\mu}{1+\mu^2 r^2}$ we can thus bound

$$\int_A \rho_{\mathfrak{z}}^2 |w|^2 dx \lesssim {f_\mu}^{-2} \log f_\mu \|w\|_{\mathfrak{z}}^2 \text{ and } \int_A \tfrac{1}{r^2} |w|^2 dx \leq C(\log \mu)^2 \|w\|_{\mathfrak{z}}^2 \tag{6.2}$$

which in particular implies that

$$\int [\rho_z^2 + |\partial_\varepsilon \nabla \gamma|^2 + |\partial_\varepsilon \nabla v_i|^2] |w|^2 \lesssim \|w\|_{\mathfrak{z}}^2 \text{ while } \int_A |\nabla \gamma|^2 |w|^2 \lesssim \delta^2 \|w\|_{\mathfrak{z}}^2. \tag{6.3}$$

We note that combined with the formula for the second variation (3.2), this immediately gives (3.3).

In the proofs of Lemmas 3.2 and 3.3 below we will furthermore use that (6.1) ensures that

$$\int_{A^*} \tfrac{1}{r^2} |w| dx \leq C(\log f_\mu)^{\frac{1}{2}} \|w\|_{\mathfrak{z}} \text{ and that } \|w\|_{L^1(\hat{A})} \leq C\mu^{-1}(\log \mu)^{\frac{1}{2}}. \tag{6.4}$$

6.1. Uniform definiteness of the second variation orthogonal to $\mathcal{Z}$

In this section we want to prove that the second variation is uniformly definite orthogonal to our set of singularity models as claimed in Lemma 3.1.

The main step in the proof of Lemma 3.1 is to show that there exists a constant $c_0 > 0$ so that for any $w \in T_{\mathfrak{z}}^{\perp}\mathcal{Z}$ with $\|w\|_{\mathfrak{z}} = 1$ there is an element $v \in T_{\mathfrak{z}}^{\perp}\mathcal{Z}$ with $\|v\|_{\mathfrak{z}} = 1$ so that

$$d^2E(\mathfrak{z})(w, v) \geq c_0. \tag{6.5}$$

Once we have shown this we can define $\mathcal{V}_{\mathfrak{z}}^{\pm}$ as the span of the eigenfunctions to positive respectively negative eigenvalues of the corresponding Jacobi-operator

$$\tilde{L}_{\mathfrak{z}} : T_{\mathfrak{z}}^{\perp}\mathcal{Z} \to T_{\mathfrak{z}}^{\perp}\mathcal{Z},$$

which is characterised by

$$d^2E(\mathfrak{z})(v, w) = \langle \tilde{L}_{\mathfrak{z}} v, w\rangle_{\mathfrak{z}} \text{ for } v, w \in T_{\mathfrak{z}}^{\perp}\mathcal{Z}$$

and obtain the claim of the lemma from the spectral theorem for selfadjoint Fredholm operators.

To prove this claim (6.5) we will relate elements $\mathfrak{z}$ of $\mathcal{Z}_0$ and vector fields v and w along $\mathfrak{z}$ to the corresponding elements $\hat{\mathfrak{z}} = U_0 \circ q_0$ of $\mathcal{H}^{\sigma}(\omega_0)$ and to vector fields $\hat{v}$ and $\hat{w}$ along $\hat{\mathfrak{z}}$.

For such vector fields we will always work with respect to the inner product

$$\langle \hat{w}, \hat{v}\rangle_{\pi} := \int_{\mathbb{R}^2} |\nabla\pi|^2 \hat{w}\hat{v} + \nabla\hat{w}\nabla\hat{v} \tag{6.6}$$

which (upto a factor 2) corresponds to the H^1 inner product on the sphere and which approximates $\langle\cdot,\rangle_{\mathfrak{z}}$ well away from the origin.

We note that $d^2E(\omega_0)$ is uniformly definite on $T_{\omega_0}^{\perp}\mathcal{H}^{\sigma}(\omega_0)$ since $T_{\omega_0}\mathcal{H}^{\sigma}(\omega_0)$ coincides with the kernel of the corresponding Jacobi-operator $L_{\omega_0} : \Gamma(\omega_0^*TN) \to \Gamma(\omega_0^*TN)$ and since the eigenvalues of this self-adjoint Fredholm operator tend to infinity. As $\mathcal{H}^{\sigma}(\omega_0)$ is a finite dimensional manifold which is contained in a small C^k neighbourhood of ω_0 we can then use the continuity of the Jacobi-operator to deduce that the analogue statement is also true for general elements of $\mathcal{H}^{\sigma}(\omega_0)$ (provided $\sigma > 0$ is chosen sufficiently small).

I.e. we can use that there exists $c > 0$ so that for any $\hat{\mathfrak{z}} \in \mathcal{H}^{\sigma}(\omega_0)$ and any $\hat{w} \in T_{\hat{\mathfrak{z}}}^{\perp}\mathcal{H}^{\sigma}(\omega_0)$ with $\|\hat{w}\|_{\pi} = 1$ there exists a unit element $\hat{v} \in T_{\hat{\mathfrak{z}}}^{\perp}\mathcal{H}(\omega_0)$ so that

$$d^2E(\hat{\mathfrak{z}})(\hat{w}, \hat{v}) \geq c. \tag{6.7}$$

In the proof below we will furthermore use that variations $\mathfrak{z}_{\varepsilon}$ in $\mathcal{Z}_0$ induced by variations of only U_0 and q_0 are well approximated by the corresponding variations $\hat{\mathfrak{z}}_{\varepsilon} = U_0^{(\varepsilon)} \circ q_0^{(\varepsilon)}$ in $\mathcal{H}^{\sigma}(\omega_0)$. To be more precise we can easily check that

$$\| |\mathfrak{z} - \hat{\mathfrak{z}}|^2 + |\partial_{\varepsilon}(\mathfrak{z} - \hat{\mathfrak{z}})|^2 \|_{L^{\infty}(\mathbb{R}^2 \setminus \mathbb{D}_{r_0^2})} + \int_{\mathbb{R}^2} |\nabla(\mathfrak{z} - \hat{\mathfrak{z}})|^2 + |\nabla\partial_{\varepsilon}(\mathfrak{z} - \hat{\mathfrak{z}})|^2 = o(1) \tag{6.8}$$

where here and in the following we write that a quantity is given by $o(1)$ if we can ensure that it is smaller than any given positive number by choosing $\bar{\mu}$ sufficiently large and σ and σ_1 sufficiently small. Here we use that $|1-\phi_\mu| \lesssim \frac{\log f_\mu}{\log \mu} = o(1)$ for $|z| \geq r_0^2 = f_\mu{}^{-2}$ to bound the L^∞ norm of $\partial_\varepsilon(\mathfrak{z}-\hat{\mathfrak{z}})$ and we note that we furthermore have

$$\|\partial_\varepsilon \mathfrak{z}\|_{\mathfrak{z},\mathbb{D}_{r_0}} + \|\partial_\varepsilon \hat{\mathfrak{z}}\|_{\pi,\mathbb{D}_{r_0}} = o(1) \tag{6.9}$$

where $\|\cdot\|_{\pi,\Omega}$ and $\|\cdot\|_{\mathfrak{z},\Omega}$ are defined as in (3.1) and (6.6) but with the integrals taken over Ω.

With these preparations in place we now turn to the proof of (6.5).

So let $\mathfrak{z} \in \mathcal{Z}_0$ and let $w \in T_{\mathfrak{z}}^{\perp}\mathcal{Z}$ be so that $\|w\|_{\mathfrak{z}} = 1$. We first remark that (6.2) and (6.3) ensure that $\|\rho_{\mathfrak{z}} w\|_{L^2(A)} + \||\nabla \mathfrak{z}| w\|_{L^2(A)} \leq \|(\rho_{\mathfrak{z}} + |\nabla\gamma|)w\|_{L^2(A)} = o(1)$. From the expression (3.2) for the second variation and the fact that outside of A we have $\nabla\gamma = 0$ and thus $|\nabla \mathfrak{z}| \lesssim \rho_{\mathfrak{z}}$, we hence obtain that

$$\begin{aligned} d^2E(\mathfrak{z})(w,w) &\geq \int_{\mathbb{R}^2} |\nabla w|^2 - C\int_{\mathbb{R}^2} |\nabla \mathfrak{z}|^2|w|^2 \qquad = \|w\|_{\mathfrak{z}}^2 - \int_{\mathbb{R}^2} (\rho_{\mathfrak{z}}^2 + C|\nabla \mathfrak{z}|^2)|w|^2 \\ &\geq \|w\|_{\mathfrak{z}}^2 - C\|w\|_{\mathfrak{z},\mathbb{R}^2\setminus A}^2 - o(1). \end{aligned}$$

The claim (6.5) is thus trivially true (for $v = w$) if $\|w\|_{\mathfrak{z},\mathbb{R}^2\setminus A}$ is small and by symmetry we can hence assume that $\|w\|_{\mathfrak{z},\mathbb{R}^2\setminus\mathbb{D}_{r_0}}^2 \geq c_4$ for a small, but fixed constant $c_4 > 0$.

Given such a w we first want to construct a $\hat{w} \in T_{\hat{\mathfrak{z}}}^{\perp}\mathcal{H}^\sigma(\omega_0)$, so that

$$\|w - \hat{w}\|_{\pi,\mathbb{R}^2\setminus\mathbb{D}_{s_0}} = o(1) \text{ while } \|\hat{w}\|_{\pi,\mathbb{D}_{2s_0}} = o(1) \text{ and } \|w\|_{\mathbb{D}_{2s_0}\setminus\mathbb{D}_{s_0/2}} = o(1) \tag{6.10}$$

for some $s_0 \in [2r_0^2, r_0]$. To this end we initially construct a function w_1 which vanishes on $\Omega_1 = \{z : |z| \leq \mu^{-\frac{1}{2}}\}$ and for which there is a radius $s_0 \in [2r_0^2, r_0]$ so that

$$w_1 \equiv w \text{ on } \mathbb{R}^2 \setminus \mathbb{D}_{s_0} \text{ while } \|w_1\|_{\mathfrak{z},\mathbb{D}_{2s_0}} = o(1).$$

To obtain such a function w_1 we use that $\mathbb{D}_{r_0} \setminus \mathbb{D}_{r_0^2}$ contains $J \gtrsim \log(r_0^{-1}) = \log f_\mu$ disjoint annuli of the form $\mathbb{D}_{4r} \setminus \mathbb{D}_{\frac{1}{4}r}$ to select $s_0 \in [2r_0^2, \frac{1}{2}r_0]$ for which

$$\int_{\mathbb{D}_{2s_0}\setminus\mathbb{D}_{s_0/2}} |\nabla w|^2 dx + \int_{S^1} |\partial_\theta w(s_0 e^{i\theta})|^2 \lesssim (\log f_\mu)^{-1} = o(1). \tag{6.11}$$

On $\mathbb{D}_{s_0} \setminus \mathbb{D}_{s_0/2}$ we then define w_1 using a standard interpolation between $w|_{\partial\mathbb{D}_{s_0}}$ and the mean value $\bar{w}$ over this circle which, thanks to (6.1), is bounded by $|\bar{w}| \lesssim (\log f_\mu)^{\frac{1}{2}}$. Combined with (6.11) and $\int_A \rho_{\mathfrak{z}}^2 \lesssim f_\mu{}^{-2}$ this ensures that $\|w_1\|_{\mathfrak{z},\mathbb{D}_{2s_0}\setminus\mathbb{D}_{s_0^2}} = o(1)$. On $\mathbb{D}_{s_0/2} \setminus \mathbb{D}_{\mu^{-\frac{1}{2}}}$ we can then choose w_1 as the harmonic function which transitions between

0 and $\bar{w}$ and note that this map has energy of order $O(\frac{\log f_\mu}{\log \mu}) = o(1)$ and weighted L^2 norm of order $O((\log f_\mu)^{\frac{1}{2}} f_\mu{}^{-1}) = o(1)$.

Given such a function w_1 we then set $w_2 := P_{\hat{\mathfrak{z}}} w_1$ to obtain a vector field along $\hat{\mathfrak{z}}$. We note that (6.8) ensures that $\|w_1 - w_2\|_{\pi,\mathbb{R}^2\setminus\mathbb{D}_{s_0}} \sim \|w_1 - w_2\|_{\mathfrak{z},\mathbb{R}^2\setminus\mathbb{D}_{s_0}} = o(1)$ and that we still have $\|w_2\|_{\mathfrak{z},\mathbb{D}_{s_0}} = o(1)$.

Given a variation $\hat{\mathfrak{z}}_\varepsilon$ in $\mathcal{H}^\sigma(\omega_0)$ with $\|\partial_\varepsilon \hat{\mathfrak{z}}\|_\pi = 1$ we can hence consider the corresponding variation $\mathfrak{z}_\varepsilon$ of $\mathfrak{z}$ in $\mathcal{Z}$ and use (6.8) and (6.9) as well as that $w \in P_{\mathfrak{z}}^\perp \mathcal{Z}$ to see that

$$|\langle w_2, \partial_\varepsilon \hat{\mathfrak{z}}\rangle_\pi| = |\langle w, \partial_\varepsilon \mathfrak{z}\rangle_{\mathfrak{z},\mathbb{R}^2\setminus\mathbb{D}_{s_0}}| + o(1) \le |\langle w, \partial_\varepsilon \mathfrak{z}\rangle_{\mathfrak{z}}| + o(1) = o(1).$$

We thus deduce that $\|P^{T_{\hat{\mathfrak{z}}}\hat{\mathcal{Z}}} w_2\|_\pi = o(1)$ and hence that $\hat{w} := P^{T_{\hat{\mathfrak{z}}}^\perp \hat{\mathcal{Z}}} w_2$ satisfies (6.10).

As $\|\hat{w}\|_\pi^2 \ge c_4 - o(1)$ is bounded away from zero we can hence apply (6.7) to obtain a vector-field $\hat{v} \in T_{\hat{\mathfrak{z}}} \mathcal{H}^\sigma(\omega_0)$ with $\|\hat{v}\|_\pi \lesssim 1$ so that

$$d^2E(\hat{\mathfrak{z}})(\hat{w}, \hat{v}) \ge 1 \tag{6.12}$$

and we will obtain the required element v of $T_{\mathfrak{z}}^\perp \mathcal{Z}$ by modifying this vector field $\hat{v}$.

To this end we repeat the argument from above to obtain a v_1 with

$$v_1 \equiv \hat{v} \text{ on } \mathbb{R}^2 \setminus \mathbb{D}_{2s_0}, \quad v_1 \equiv 0 \text{ on } \mathbb{D}_{\mu^{-\frac{1}{2}}}, \quad \|v_1\|_{\mathfrak{z},\mathbb{D}_{2s_0}\setminus\mathbb{D}_{s_0/2}} \lesssim 1 \text{ and } \|v_1\|_{\mathfrak{z},\mathbb{D}_{s_0/2}} = o(1), \tag{6.13}$$

except that we work with a radius $\hat{s}_0 \in [s_0, 2s_0]$ with $\int_{S^1} |\partial_\theta \hat{v}(\hat{s}_0 e^{i\theta})|^2 \lesssim 1$ rather than with s_0.

As $\hat{v}$ is tangential to N along $\hat{\mathfrak{z}}$ and supported on Ω_0 where $\rho_{\mathfrak{z}} \sim |\nabla\pi|$ we can combine (6.8) and (6.13) to see that

$$\|P_{\mathfrak{z}} v_1 - v_1\|_{\mathfrak{z}} \lesssim \|v_1\|_{\mathfrak{z},\mathbb{D}s_0/2} + \|\mathfrak{z} - \hat{\mathfrak{z}}\|_{L^\infty(\mathbb{R}^2\setminus\mathbb{D}_{r_0^2})} \|v_1\|_{\mathfrak{z}} + \||\nabla\mathfrak{z} - \nabla\hat{\mathfrak{z}}| v_1\|_{L^2(\mathbb{R}^2\setminus\mathbb{D}_{r_0^2})} = o(1)$$

where we use that $|\nabla\mathfrak{z} - \nabla\hat{\mathfrak{z}}| \lesssim o(1)(\rho_{\mathfrak{z}} + |\nabla\phi_\mu|)$ on $\mathbb{R}^2 \setminus \mathbb{D}_{r_0^2}$ to deal with the last term.

As $\hat{v}$ is orthogonal to $T_{\hat{\mathfrak{z}}}\mathcal{H}^\sigma(\omega_0)$ we can combine this bound with (6.8) and (6.9) to see that $\langle \partial_\varepsilon \mathfrak{z}_\varepsilon, P_{\mathfrak{z}} v_1\rangle_{\mathfrak{z}} = o(1)$ for all variations $\mathfrak{z}_\varepsilon$ in $\mathcal{Z}$ that are induced by variations of only U_0 and q_0. Conversely, variations of U_1 and q_1 result in variations $\mathfrak{z}_\varepsilon$ for which $\|\partial_\varepsilon \mathfrak{z}_\varepsilon\|_{\mathfrak{z},\Omega_0} = o(1)$ and for which we hence trivially know that $\langle \partial_\varepsilon \mathfrak{z}_\varepsilon, P_{\mathfrak{z}} v_1\rangle_{\mathfrak{z}} = o(1)$.

All in all this ensures that $v := P^{T_{\mathfrak{z}}^\perp \mathcal{Z}}(P_{\mathfrak{z}} v_1) \in T_{\mathfrak{z}}^\perp \mathcal{Z}$ is so that

$$\|v\|_{\mathfrak{z},\mathbb{D}_{s_0/2}} = o(1), \quad \|v\|_{\mathfrak{z},\mathbb{D}_{2s_0}\setminus\mathbb{D}_{s_0/2}} \lesssim 1 \text{ and } \|v - \hat{v}\|_{\mathfrak{z},\mathbb{R}^2\setminus\mathbb{D}_{2s_0}} = o(1).$$

We now write for short $II_{\mathfrak{z}}(w,v) = \nabla w \nabla v - A(\mathfrak{z})(\nabla\mathfrak{z}, \nabla\mathfrak{z}) A(\mathfrak{z})(w,v)$ for the integrand that appears in the formula (3.2) for d^2E and note that $\int_\Omega |II_{\mathfrak{z}}(w,v)| \lesssim \|w\|_{\mathfrak{z},\Omega}\|v\|_{\mathfrak{z},\Omega} + o(1)$ for every $\Omega \subset \mathbb{R}^2$ where the $o(1)$ term comes from the contribution of $\nabla\gamma$ to $\nabla\mathfrak{z}$, compare (6.3). We hence get

$$
\begin{aligned}
d^2E(\mathfrak{z})(v,w) &\geq \int\limits_{\mathbb{R}^2\setminus\mathbb{D}_{2s_0}} II_{\mathfrak{z}}(w,v) - C\big[\|v\|_{\mathfrak{z},\mathbb{D}_{s_0/2}} + \|w\|_{\mathfrak{z},\mathbb{D}_{2s_0}\setminus\mathbb{D}_{s_0/2}}\big] - o(1)\\
&\geq \int\limits_{\mathbb{R}^2\setminus\mathbb{D}_{2s_0}} II_{\hat{\mathfrak{z}}}(\hat{w},\hat{v}) - C\big[\|v-\hat{v}\|_{\mathfrak{z},\mathbb{R}^2\setminus\mathbb{D}_{2s_0}} + \|w-\hat{w}\|_{\mathfrak{z},\mathbb{R}^2\setminus\mathbb{D}_{2s_0}}\\
&\quad + \|\mathfrak{z}-\hat{\mathfrak{z}}\|_{L^\infty(\mathbb{R}^2\setminus\mathbb{D}_{2s_0})}\big] - o(1)\\
&\geq \int\limits_{\mathbb{R}^2} II_{\hat{\mathfrak{z}}}(\hat{w},\hat{v}) - C\|\hat{w}\|_{\pi,\mathbb{D}_{2s_0}} - o(1)\\
&\geq d^2E(\hat{\mathfrak{z}})(\hat{w},\hat{v}) - o(1) \geq 1 - o(1) \geq \tfrac{1}{2},
\end{aligned}
$$

where the last inequality holds true provided σ and σ_1 are sufficiently small and $\bar{\mu}$ is sufficiently large. As $\|v\|_{\mathfrak{z}} \lesssim 1$ we hence obtain that (6.5) holds true for the corresponding unit element of $T_{\mathfrak{z}}^{\perp}\mathcal{Z}$ which completes the proof of Lemma 3.1.

6.2. Proofs of Lemmas 3.2 and 3.3

We finally turn to the proofs of the estimates on the norms of the first and second variation of E claimed in Lemmas 3.2 and 3.3. To simplify the notation we write for short

$$R_3 := \|\beta\|\bar{\nu}(\log\mu)^{\frac{1}{2}} + \tfrac{\delta}{f_\mu \log\mu} \tag{6.14}$$

and

$$R_4 := (\|\beta\| + \|\partial_\varepsilon\beta\|)\bar{\nu}\log(\mu)^{\frac{1}{2}} + \eta_{rat}\|\beta\|\mu^{-j^*}\log(\mu)^{\frac{1}{2}} + \tfrac{(\delta+\eta_0)(\log f_\mu)^{\frac{1}{2}}}{\log\mu} \tag{6.15}$$

and recall that to prove Lemma 3.2 we need to show that

$$|dE(\mathfrak{z})(w)| = |\int \tau(\mathfrak{z})w| \lesssim R_3 + \mathcal{T} \text{ for every } w \text{ with } \|w\|_{\mathfrak{z}} = 1$$

while Lemma 3.3 asserts that

$$|d^2E(\mathfrak{z})(\partial_\varepsilon\mathfrak{z},w)| = |\int \partial_\varepsilon(\tau(\mathfrak{z}))w| \lesssim R_4 + \mathcal{T} + \eta_{\mathcal{T}} \text{ for every } w \in \Gamma(\mathfrak{z}^*TN) \text{ with } \|w\|_{\mathfrak{z}} = 1.$$

By symmetry it suffices to bound the corresponding integrals over Ω_1 where $\mathfrak{z} = \mathfrak{z}_1 + \mathrm{err}_\varphi$ for $\mathfrak{z}_1 = \pi_N(v_1)$ and for err_φ which is defined by (5.55) and supported on $\hat{A}$.

To prove both lemmas we write the tension of $\mathfrak{z}$ on Ω_1 as

$$\tau(\mathfrak{z}) = P_{\mathfrak{z}}(\Delta\mathfrak{z}_1 + \Delta\mathrm{err}_\varphi) = \tau(\mathfrak{z}_1) + (P_{\mathfrak{z}} - P_{\mathfrak{z}_1})(\Delta\mathfrak{z}_1) + P_{\mathfrak{z}}(\Delta\mathrm{err}_\varphi), \tag{6.16}$$

and use that $\Delta j_1 = 0$ to write

$$\tau(\mathfrak{z}_1) = P_{\mathfrak{z}_1}(d\pi_N(v_1)\Delta v_1) + P_{\mathfrak{z}_1}(d^2\pi_N(v_1)(\nabla v_1, \nabla v_1)) = T^\tau_{u_1} + T^A_{u_1} + T_\gamma + T_\pi \tag{6.17}$$

where we split the contribution of Δu_1 into the terms

$$T^\tau_{u_1} := P_{\mathfrak{z}_1}(d\pi_N(v_1)(\tau(u_1)) \text{ and } T^A_{u_1} := -P_{\mathfrak{z}_1}(d\pi_N(v_1))(A(u_1)(\nabla u_1, \nabla u_1))) \tag{6.18}$$

and write for short

$$T_\gamma := P_{\mathfrak{z}_1}(\Delta\gamma) \text{ and } T_\pi := P_{\mathfrak{z}_1}(d^2\pi_N(v_1)(\nabla v_1, \nabla v_1)). \tag{6.19}$$

In the analysis of the resulting integrals we will use that

$$|\mathfrak{z}_1 - v_1| \le |u_1 - v_1| \le |\tilde{\gamma}_{U_1}| + |j_1| \lesssim \delta\phi_\mu + \|\beta\||\hat{q}_0| \tag{6.20}$$

which, thanks to (5.47), (5.48) and (5.7), ensures that

$$\||v_1 - \mathfrak{z}_1|\rho_q\|_{L^2(\Omega_1)} \le \||u_1 - v_1|\rho_q\|_{L^2(\Omega_1)} \lesssim R_3. \tag{6.21}$$

Similarly, since we can bound

$$\begin{aligned} |\partial_\varepsilon(u_1 - v_1)| &\lesssim |\partial_\varepsilon j_1| + |\partial_\varepsilon\tilde{\gamma}_{U_1}| + |\tilde{\gamma}_{U_1}| + |j_1| \\ &\lesssim (\|\beta\| + \|\partial_\varepsilon\beta\|)|\hat{q}_0| + \|\beta\||\partial_\varepsilon\hat{q}_0| + (\delta + \eta_0)\phi_\mu + \delta(\log\mu)^{-1}\mathbb{1}_A, \end{aligned} \tag{6.22}$$

compare also (5.16), we can use these estimates (5.47), (5.48) and (5.7) together with (5.49) to see that

$$\||\partial_\varepsilon(u_1 - v_1)|\rho_q\|_{L^2} \lesssim R_4. \tag{6.23}$$

To analyse terms that involve γ we instead write $v_1 = \gamma + \tilde{u}_1 + j_1$ to see that

$$|v_1 - \gamma| + |\partial_\varepsilon(v_1 - \gamma)| \lesssim |\hat{q}_0| + |\hat{q}_{1,\mu}| + |\partial_\varepsilon\hat{q}_0| + |\partial_\varepsilon\hat{q}_{1,\mu}| \lesssim |z| + (1 + \mu^2|z|^2)^{-1}$$

and hence that also

$$\||v_1 - \gamma|\nabla\gamma\|_{L^2} + \||\partial_\varepsilon(v_1 - \gamma)|\nabla\gamma\|_{L^2} \le \tfrac{\delta}{f_\mu \log\mu} \lesssim R_3. \tag{6.24}$$

Finally, since $|P_{\mathfrak{z}_1}\nabla v_1 - \nabla v_1| \lesssim |u_1 - \mathfrak{z}_1|\rho_q + \|\beta\||\nabla\hat{q}_0| + |v_1 - \gamma||\nabla\gamma|$, we also get that

$$\|P_{\mathfrak{z}_1}\nabla v_1 - \nabla v_1\|_{L^2} \lesssim R_3. \tag{6.25}$$

With these estimates in place we can now show that all contributions to $dE(\mathfrak{z})(w)$ are controlled by $R_3 + \mathcal{T}$ and all contributions to $d^2E(\mathfrak{z})(\partial_\varepsilon\mathfrak{z}, w)$ are controlled by $R_4 + \mathcal{T} + \eta_{\mathcal{T}}$.

To begin with we use that $|\tau(u_1)| \lesssim \mathcal{T}\rho_q^2 \lesssim \mathcal{T}\rho_{\mathfrak{z}}^2$ to see that

$$\int_{\Omega_1} |T_{u_1}^\tau w| \lesssim \mathcal{T}\|w\|_{\mathfrak{z}} = \mathcal{T}$$

while we can use that $|\partial_\varepsilon T_{u_1}^\tau| \lesssim |\tau(u_1)| + |\partial_\varepsilon \tau(u_1)| \lesssim \mathcal{T}\rho_q\rho_{\mathfrak{z}} + \eta_\tau \rho_q^2 \lesssim (\mathcal{T} + \eta_\tau)\rho_{\mathfrak{z}}^2$ to bound

$$|\int T_{u_1}^\tau w| \lesssim \mathcal{T} + \eta_\tau.$$

As $A(u_1)(\nabla u_1, \nabla u_1) \in T_{u_1}^\perp N$ we can then write

$$T_{u_1}^A = P_{\mathfrak{z}_1}\big((d\pi_N(u_1) - d\pi_N(v_1))(A(u_1)(\nabla u_1, \nabla u_1))\big) \tag{6.26}$$

and hence use (6.21) to bound

$$\int_{\Omega_1} |T_{u_1} w| \lesssim \||u_1 - z_1|\rho_q\|_{L^2} \lesssim R_3.$$

From (6.26) we see that $|\partial_\varepsilon T_{u_1}^A| \lesssim \big[|u_1 - v_1| + |\partial_\varepsilon(u_1 - v_1)|\big]\rho_q \cdot \rho_{\mathfrak{z}}$ which we can combine with (6.21) and (6.23) to see that also

$$|\int \partial_\varepsilon T_{u_1}^A w| \lesssim \||u_1 - v_1|\rho_q\|_{L^2} + \||\partial_\varepsilon(u_1 - v_1)|\rho_q\|_{L^2} \lesssim R_3 + R_4 \lesssim R_4.$$

Similarly, we can write

$$T_\gamma = P_{\mathfrak{z}_1}((d\pi_N(\gamma) - d\pi_N(v_1))(A(\gamma)(\nabla\gamma, \nabla\gamma))) + P_{\mathfrak{z}_1}(d\pi_N(v_1)\tau(\gamma))) \tag{6.27}$$

and use (5.11), (6.3), (6.4) and (6.24) to see that

$$|\int T_\gamma w| \leq \delta\||\gamma - v_1||\nabla\gamma|\|_{L^2} + \delta c_\mu \int_{A^*} |w| r^{-2} \lesssim \tfrac{\delta^2}{f_\mu \log\mu} + \tfrac{\delta(\log f_\mu)^{\frac{1}{2}}}{\log\mu} \lesssim R_3. \tag{6.28}$$

As

$$|\partial_\varepsilon \tau(\gamma)| \lesssim (\eta_0 + \delta^2 c_\mu)|\Delta\phi_\mu| + \delta|\partial_\varepsilon \Delta\phi_\mu| \lesssim (\eta_0 + \delta)c_\mu r^{-2}\mathbb{1}_{A^*} \tag{6.29}$$

we can furthermore use (6.3), (6.4), (6.24) and (5.11) to see that

$$\begin{aligned}|\int \partial_\varepsilon T_\gamma w| &\lesssim \||\gamma - v_1||\nabla\gamma|)\|_{L^2} + \||\partial_\varepsilon(\gamma - v_1)||\nabla\gamma|\|_{L^2} + \int |\tau(\gamma)||w| + |\partial_\varepsilon\tau(\gamma)||w| \\ &\lesssim \tfrac{\delta+\eta_0}{(\log\mu)f_\mu} + \tfrac{\delta+\eta_0}{(\log\mu)} \int_{A^*} r^{-2}|w| \lesssim \tfrac{\delta+\eta_0}{(\log\mu)}(\log f_\mu)^{\frac{1}{2}} \lesssim R_4\end{aligned}$$

where the penultimate step follows from (6.4).

Finally we use (5.41) to write

$$T_\pi = P_{\mathfrak{z}_1}(d^2\pi_N(v_1)(\nabla v_1, \nabla v_1) - d^2\pi_N(\mathfrak{z}_1)(P_{\mathfrak{z}_1}\nabla v_1, P_{\mathfrak{z}_1}\nabla v_1)). \tag{6.30}$$

From (6.3), (6.21), (6.24) and (6.25) we hence get that also

$$\int |T_\pi w| \lesssim \| |v_1 - \mathfrak{z}_1|\rho_q\|_{L^2} + \||v_1 - \mathfrak{z}_1||\nabla\gamma|\|_{L^2} + \|P_{\mathfrak{z}_1}\nabla v_1 - \nabla v_q\|_{L^2} \lesssim R_3.$$

From (6.30) we see that

$$|\int \partial_\varepsilon T_\pi w| \lesssim \||v_1 - \mathfrak{z}_1|\nabla v_1\|_{L^2} + \||\partial_\varepsilon(v_1 - \mathfrak{z}_1)|\nabla v_1\|_{L^2} + \|P_{\mathfrak{z}_1}\nabla v_1 - \nabla v_1\|_{L^2}$$
$$+ \int |\partial_\varepsilon(P_{\mathfrak{z}_1}^\perp \nabla v_1)||\nabla v_1||w|.$$

As we can write $P_{\mathfrak{z}_1}^\perp \nabla v_1 = (P_{\mathfrak{z}_1}^\perp - P_{u_1}^\perp)(\nabla u_1) + (P_{\mathfrak{z}_1}^\perp - P_\gamma^\perp)(\nabla\gamma) + (P_{\mathfrak{z}_1}^\perp - P_{U_1(0)}^\perp)\nabla j_1$ we hence also get that $|\int \partial_\varepsilon T_\pi w| \lesssim R_4$.

All in all we hence get that

$$\int_{\Omega_1} |\tau(\mathfrak{z}_1)w| \lesssim R_3 + \mathcal{T} \text{ and } \int_{\Omega_1} |\partial_\varepsilon\tau(\mathfrak{z}_1)w| \lesssim R_4 + \mathcal{T} + \eta_{\mathcal{T}}.$$

To complete the proofs of Lemmas 3.2 and 3.3 it remains to bound the contributions of the last two terms in (6.16). These are supported on $\hat{A}$ where $\rho_{\mathfrak{z}} \lesssim 1$, $\tau(\gamma) = 0$ and hence

$$|\Delta\mathfrak{z}_1| \lesssim \mu(\bar{\nu} + \delta^2(\log\mu)^{-2}) \tag{6.31}$$

so in particular $|\Delta\mathfrak{z}_1| \lesssim 1 + \frac{\delta^2}{(\log\mu)^2}\mu$. Combined with (5.56) and (6.4) this shows that

$$|\int (P_{\mathfrak{z}} - P_{\mathfrak{z}_1})(\Delta\mathfrak{z}_1)w| \lesssim \|\beta\|\bar{\nu}\big[1 + \tfrac{\delta^2}{(\log\mu)^2}\mu\big]\|w\|_{L^1(\hat{A})} \lesssim \|\beta\|\bar{\nu}\mu^{-1}(\log\mu)^{\frac{1}{2}} + \tfrac{\delta^2}{(\log\mu)^{3/2}}$$

only gives a lower order contribution, while (5.56) and (6.4) also allow us to bound

$$\int |\Delta \mathrm{err}_\varphi||w| \lesssim \|\beta\|\bar{\nu}\int_{\hat{A}} |w| \lesssim \|\beta\|\bar{\nu}(\log\mu)^{\frac{1}{2}} \lesssim R_3.$$

This completes the proof of Lemma 3.2

Combining (6.31) with (5.58) gives a pointwise bound of $|\partial_\varepsilon \mathrm{err}_\varphi||\Delta\mathfrak{z}_1| \lesssim R_4$ while combining $|\partial_\varepsilon\Delta\mathfrak{z}_1| \lesssim \mu(\bar{\nu} + \eta_{rat}\mu^{-j^*}) + \mu\frac{\delta^2 + \delta\eta_0}{(\log\mu)^2}$ with (5.56) ensures that also $|\mathrm{err}_\varphi||\partial_\varepsilon\Delta\mathfrak{z}_1| \lesssim R_4$ on $\hat{A}$. Combined with (6.4) we can hence obtain that

$$|\int \partial_\varepsilon((P_{\mathfrak{z}_1}-P_{\mathfrak{z}})\Delta\mathfrak{z}_1)w| \lesssim R_4\mu^{-1}(\log\mu)^{\frac{1}{2}} \ll R_4$$

just gives a lower order contribution to $d^2E(\mathfrak{z})(\partial_\varepsilon\mathfrak{z}, w)$.

Finally (5.56), (5.59) and (6.4) allow us to also bound

$$\begin{aligned}\int |\partial_\varepsilon(P_{\mathfrak{z}}\Delta\mathrm{err}_\varphi)||w| &\lesssim (\|\partial_\varepsilon\Delta\mathrm{err}_\varphi\|_{L^\infty(\hat{A})} + \|\Delta\mathrm{err}_\varphi\|_{L^\infty(\hat{A})})\|w\|_{L^1(\hat{A})} \\ &\lesssim \left[(\|\beta\| + \|\partial_\varepsilon\beta\|)\bar{\nu} + \eta_{rat}\|\beta\|\mu^{-j^*}\right](\log\mu)^{\frac{1}{2}} \lesssim R_4,\end{aligned} \tag{6.32}$$

which completes the proof of Lemma 3.3.

Data availability

No data was used for the research described in the article.

References

[1] W.Y. Ding, G. Tian, Energy identity for a class of approximate harmonic maps from surfaces, Commun. Anal. Geom. 3 (1995) 543–554.
[2] R.D. Gulliver, R. Osserman, H.L. Royden, A theory of branched immersions of surfaces, Am. J. Math. 95 (1973) 750–812.
[3] R.D. Gulliver, B. White, The rate of convergence of a harmonic map at a singular point, Math. Ann. 283 (4) (1989) 539–549.
[4] F.-H. Lin, C.-Y. Wang, Energy identity of harmonic map flows from surfaces at finite singular time, Calc. Var. Partial Differ. Equ. 6 (1998) 369–380.
[5] F.-H. Lin, Mapping problems, fundamental groups and defect measures, Acta Math. Sin. Engl. Ser. 15 (1999) 25–52.
[6] Q. Liu, Y. Yang, Rigidity of the harmonic map heat flow from the sphere to compact Kähler manifolds, Ark. Mat. 48 (1) (2010) 121–130.
[7] A. Malchiodi, M. Rupflin, B. Sharp, Łojasiewicz inequalities near simple bubble trees, Am. J. Math. 146 (2024) 1361–1397.
[8] J.D. Moore, Bumpy metrics and closed parametrized minimal surfaces in Riemannian manifolds, Trans. Am. Math. Soc. 358 (12) (2006) 5193–5256.
[9] J.D. Moore, Introduction to Global Analysis: Minimal Surfaces in Riemannian Manifolds, AMS Graduate Studies in Mathematics, vol. 187, American Mathematical Society, Providence, 2017.
[10] J. Qing, G. Tian, Bubbling of the heat flows for harmonic maps from surfaces, Commun. Pure Appl. Math. 50 (1997) 295–310.
[11] M. Rupflin, Lojasiewicz inequalities for almost harmonic maps near simple bubble trees, Ann. Inst. H. Poincaré C Anal. Non Linéaire (2024), https://doi.org/10.4171/AIHPC/131, published online first.
[12] J. Sacks, K. Uhlenbeck, The existence of minimal immersions of 2-spheres, Ann. Math. 113 (1981) 1–24.
[13] L. Simon, Asymptotics for a class of non-linear evolution equations, with applications to geometric problems, Ann. Math. 118 (1983) 525–571.
[14] M. Struwe, On the evolution of harmonic mappings of Riemannian surfaces, Comment. Math. Helv. 60 (1985) 558–581.
[15] P.M. Topping, Rigidity in the harmonic map heat flow, J. Differ. Geom. 45 (1997) 593–610.
[16] P.M. Topping, Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic map flow, Ann. Math. 159 (2004) 465–534.
[17] P.M. Topping, Bubbling of almost-harmonic maps between 2-spheres at points of zero energy density, in: Variational Problems in Riemannian Geometry: Bubbles, Scans and Geometric Flows, in: Progress in Nonlinear Differential Equations and Their Applications, vol. 59, 2004, pp. 33–42.

[18] A. Waldron, Lojasiewicz inequalities for maps of the 2-sphere, preprint, arXiv:2312.16686, 2023.
[19] B. White, The space of minimal submanifolds for varying Riemannian metrics, Indiana Univ. Math. J. 40 (1) (1991) 161–200.
[20] B. White, On the bumpy metrics theorem for minimal submanifolds, Am. J. Math. 139 (4) (2017) 1149–1155.
[21] B. White, Generic transversality of minimal submanifolds and generic regularity of two-dimensional area-minimizing integral currents, arXiv:1901.05148.