Title
Min-max minimal disks with free boundary in riemannian manifolds

Journal
Geometry and Topology, 24(1)

ISSN
1465-3060

Authors
Lin, L
Sun, A
Zhou, X

Publication Date
2020

DOI
10.2140/gt.2020.24.470

Geometry & Topology 24 (2020) 471–532

Min-max minimal disks with free boundary in Riemannian manifolds

LONGZHI LIN
AO SUN
XIN ZHOU

We establish a min-max theory for constructing minimal disks with free boundary in any closed Riemannian manifold. The main result is an effective version of the partial Morse theory for minimal disks with free boundary established by Fraser. Our theory also includes as a special case the min-max theory for the Plateau problem of minimal disks, which can be used to generalize the famous work by Morse–Tompkins and Shiffman on minimal surfaces in $\mathbb{R}^n$ to the Riemannian setting.

More precisely, we generalize, to the free boundary setting, the min-max construction of minimal surfaces using harmonic replacement introduced by Colding–Minicozzi. As a key ingredient to this construction, we show an energy convexity for weakly harmonic maps with mixed Dirichlet and free boundaries from the half unit 2–disk in $\mathbb{R}^2$ into any closed Riemannian manifold, which in particular yields the uniqueness of such weakly harmonic maps. This is a free boundary analogue of the energy convexity and uniqueness for weakly harmonic maps with Dirichlet boundary on the unit 2–disk proved by Colding and Minicozzi.

35R35, 49J35, 49Q05, 53C43

Published: 25 March 2020 DOI: 10.2140/gt.2020.24.471

0 Introduction

Given a closed Riemannian manifold $\mathcal{N}^n$ (that is isometrically embedded in $\mathbb{R}^N$) and an embedded submanifold Γ of codimension $l \geq 1$, a map from the unit 2–disk $D \subset \mathbb{R}^2$ into $\mathcal{N}$ with boundary lying on Γ is said to be a *minimal disk with free boundary* if it minimizes area up to the first order among all such maps. One physical model of such surfaces is the soap film whose boundary is constrained (but allowed to move freely) on the boundary of some smooth domain in $\mathbb{R}^3$. When a soap film achieves the equilibrium state, it will minimize the area up to first order. Geometrically, the stationary soap film would have vanishing mean curvature in the interior and meet the boundary of the given domain orthogonally. Therefore, the orthogonality condition at the boundary is called the *free boundary* condition.

After earlier works of Gergonne in 1816 and H Schwarz in 1890, Courant first studied systematically the free boundary problems for minimal surfaces in a series of seminal papers; see Courant [8, Chapter VI] and also Courant and Davids [9]). In particular, Courant and Davids proved that given an embedded closed surface S in $\mathbb{R}^3$ other than the sphere, there exists a minimal disk Σ with free boundary on S under certain linking conditions; see [8, pages 213–218]. Since then there have been immense research activities on this topic. To remove the topological assumption on the constraint surface S, Smyth [51] showed that if S is the boundary of a tetrahedron (which is *nonsmooth*), then there must exist exactly three minimal disks embedded inside the tetrahedron satisfying the free boundary condition. When S is a smooth topological two-sphere, Struwe [52] used the mountain pass lemma to establish the existence of at least one *unstable* minimal disk with free boundary lying on S. In higher dimensions and codimensions (for $(\mathcal{N}, \Gamma)$), Ye [57] obtained the existence of an area-minimizing disk with free boundary when the kernel of $\pi_1(\Gamma) \to \pi_1(\mathcal{N})$ is nontrivial. About 20 years ago, Fraser [15] developed a partial Morse theory for finding minimal disks with free boundary in any codimension using the perturbed energy approach of Sacks–Uhlenbeck [45], and proved the existence of solutions with bounded Morse index when the relative homotopy group $\pi_k(\mathcal{N}, \Gamma)$ is nontrivial for some $k \in \mathbb{N}$. Recently, Fraser–Schoen [16; 17] established deep relations between the minimal surfaces with free boundary in round balls and the extremal eigenvalue problems.

In this paper, we develop a direct variational theory for constructing minimal disks with free boundary for any pair $(\mathcal{N}, \Gamma)$, using the min-max method. In particular, given a k–parameter family of mappings from the unit disk D into $\mathcal{N}$ such that ∂D is mapped into Γ, one can associate to it a min-max value analogous to the classical Morse theory. We prove that if the min-max value is nontrivial, then there exists a minimal disk together with (possibly empty) several minimal spheres (possibly with a puncture), which is usually called a *bubble tree* and whose areas sum up to the min-max value (the so-called *energy identity*). Moreover, we prove that every approximate sequence of maps converges to a bubble tree such that the energy identity holds true. By reflecting on this strong convergence property, our results can be viewed as an effective version of Fraser [15]; see more discussion in Remark 0.2(2).

Our theory can be also used to generalize the famous work of Morse–Tompkins [34] and Shiffman [49] on minimal surfaces in $\mathbb{R}^n$ to the Riemannian setting (Theorem 0.3). In particular, if we take Γ to be a Jordan curve in $\mathcal{N}$ and assume that Γ bounds two different strictly minimizing minimal disks $\overline{v}_i\colon D \to \mathcal{N}$, $i = 1, 2$, such that $\overline{v}_i|_{\partial D}\colon \partial D \to \Gamma$ is a monotone parametrization, then our theory produces finitely many harmonic disks u_k (and minimal spheres) such that when restricted to ∂D, only one will have degree 1 and all others will have degree 0. If additionally one has that the $u_k|_{\partial D}\colon \partial D \to \Gamma$ are monotone parametrizations, then in the special case when $\mathcal{N}$ has nonpositive curvature (so that there exist no punctured minimal spheres), our theory produces a third *nonminimizing* minimal disk, and therefore it provides a direct generalization of the work of Morse–Tompkins [34] and Shiffman [49]; see Section 8 for more details.

Another novelty of this paper is reflected by our constructive method. The *Schwarz alternating method* introduced by H Schwarz goes back to the late 1860s, and later it was generalized to an iterative method for finding the solution of an elliptic PDE on a domain which is the union of two overlapping subdomains. In [6], Colding and Minicozzi adapted this method and used the *harmonic replacement* to construct min-max minimal surfaces. During this repeated replacement procedure, at each step one replaces a map u by a map $\widetilde{u}$ that coincides with u outside a disk and inside the disk is equal to an energy-minimizing map with the same boundary values as u. A key ingredient to this construction is a version of energy convexity for weakly harmonic maps with Dirichlet boundary and small energy on the unit 2–disk D, which also yields the (quantitative) uniqueness for such weakly harmonic maps; see also Lamm and the first author [27]. In this paper, we generalize this min-max construction of Colding–Minicozzi using harmonic replacement to construct minimal disks with free boundary in any closed Riemannian manifold. To this end, we will show an energy convexity for weakly harmonic maps with mixed Dirichlet and free boundaries from the half unit 2–disk into any closed Riemannian manifold (Theorem 2.2), which is the free boundary analogue of Colding–Minicozzi's energy convexity and uniqueness for weakly harmonic maps. We shall remark that a priori it is not at all clear if such energy convexity should hold, due to the complication of the free boundary component of the map. The key to this is an ϵ–regularity (gradient estimate) for weakly harmonic maps with mixed Dirichlet and free boundaries on the half unit 2–disk (Theorem 2.5).

Now we proceed to present the precise mathematical statements of our main results. To make the presentation simpler, we will focus on 1–parameter min-max constructions, though our results extend in a straightforward manner to k parameters. We will use $[0,1]$ as the parameter space. Consider the total variational space

$$\Omega = \left\{ \sigma\colon D\times[0,1]\to\mathcal{N} \;\middle|\; \begin{array}{l} \sigma\colon [0,1]\to C^0(\bar{D},\mathcal{N})\cap W^{1,2}(D,\mathcal{N}) \text{ is continuous,} \\ \sigma(\cdot,t)(\partial D)\subset\Gamma \text{ for all } t\in[0,1], \\ \sigma(\cdot,0) \text{ and } \sigma(\cdot,1) \text{ are constant maps.} \end{array} \right\}.$$

Each $\beta\in\Omega$ will be called a *sweepout*. Given a map $\beta\in\Omega$, we define Ω_β to be the homotopy class of β in Ω.

Here and in the following, we write $E(\cdot)$ and $\mathrm{Area}(\cdot)$ for the Dirichlet energy and area functionals on $C^0(\bar{D},\mathcal{N})\cap W^{1,2}(D,\mathcal{N})$. Associated to each homotopy class Ω_β, there is a min-max value, also called the *width* of Ω_β,

$$W = W(\Omega_\beta) := \inf_{\gamma\in\Omega_\beta}\max_{s\in[0,1]}\mathrm{Area}(\gamma(\cdot,s)). \tag{1}$$

As the first main result, we establish a direct variational construction for minimal surfaces associated with this critical value W.

Theorem 0.1 *Given $\beta\in\Omega$ with $W=W(\Omega_\beta)>0$, there is a sequence of sweepouts $\gamma^j\in\Omega_\beta$ with $\max_{s\in[0,1]}E(\gamma^j(\cdot,s))\to W$ such that for any given $\epsilon>0$, there exist $\bar{J}$ and $\delta>0$ such that if $j>\bar{J}$ and*

$$\mathrm{Area}(\gamma^j(\cdot,s))>W-\delta,$$

then there exist finitely many harmonic maps $u_k\colon D\to\mathcal{N}$ which have free boundary $u_k(\partial D)\subset\Gamma$, and finitely many (possibly empty) *harmonic maps $\hat{u}_l\colon S^2\to\mathcal{N}$, such that*

$$d_V\Big(\gamma^j(\cdot,s),\bigcup_k\{u_k\}\cup\bigcup_l\{\hat{u}_l\}\Big)<\epsilon.$$

Here we have identified each map u_i with the varifold associated to the map, and d_V denotes the varifold distance. Moreover we have the energy identity

$$\sum_k \mathrm{Area}(u_k) + \sum_l \mathrm{Area}(\hat{u}_l) = W.$$

Remark 0.2 (i) By the work of Sacks–Uhlenbeck [45, Corollary 1.7] and Fraser [15] (see also [14, Lemma 1.1]), we know that these harmonic maps are conformal and hence parametrize minimal disks with free boundary on Γ or minimal spheres.

(ii) Our result is an effective version of Fraser's result [15] in the sense that we obtain a strong convergence property. In particular, a sequence of maps $\{\gamma^{j_k}(\cdot, s_k)\}$ is usually called a min-max sequence if

$$\lim_{k\to\infty} \mathrm{Area}(\gamma^{j_k}(\cdot, s_k)) = W.$$

We prove that every min-max sequence will subconverge to a set of minimal disks with free boundary and possibly some minimal spheres in the varifold sense. The essential ingredient is to prove that all the min-max sequences converge to a bubble tree limit and the energy identity holds. Note that the energy identity has caught a lot of attention from mathematicians in conformally invariant variational problems; see eg Chen and Tian [4], Ding and Tian [12], Ding, Li and Liu [11], Jost [25], Parker and Wolfson [36; 37] and Qing and Tian [40], and for quantification results for harmonic maps with free boundary, see eg Jost, Liu and Zhu [26] and Laurain and Petrides [30]. A similar strong convergence property was first proven by Colding and Minicozzi [6] for the min-max construction of minimal spheres, and it played an essential role in their proof of the finite-time extinction for certain 3–dimensional Ricci flow. A similar property was also obtained for the min-max construction of closed minimal surfaces of higher genus by the last author [58; 60], and for min-max construction of closed minimal surfaces via viscosity method by Rivière [43]. To the authors' knowledge, our work is the first occasion to obtain such a strong property in the context of free boundary problems.

(iii) As a special case of our result, one can take $(\mathcal{N}, \Gamma)$ to be a compact manifold with convex boundary $(M, \partial M)$. Using the convex boundary as barriers, our theory applies in this case and the resulting minimal disks with free boundary on ∂M and minimal spheres will all lie inside M.

Our main result has an almost direct corollary for a min-max construction of minimal disks with fixed boundary. In particular, we now assume Γ to be a Jordan curve in $\mathcal{N}$. Suppose $\overline{v}_0\colon D\to\mathcal{N}$ and $\overline{v}_1\colon D\to\mathcal{N}$ are two area-minimizing minimal disks (conformal and harmonic maps), where $\overline{v}_i|_{\partial D}\colon \partial D\to\Gamma$ is a monotone parametrization for $i=1,2$. The total variational space for the fixed boundary problem will be

$$\Omega_f=\left\{\sigma\colon D\times[0,1]\to\mathcal{N}\;\middle|\;\begin{array}{l}\sigma\colon[0,1]\to C^0(\bar{D},\mathcal{N})\cap W^{1,2}(D,\mathcal{N})\ \text{is continuous},\\ \sigma(\cdot,t)(\partial D)\subset\Gamma,\\ \sigma(\cdot,0)=\overline{v}_0\ \text{and}\ \sigma(\cdot,1)=\overline{v}_1.\end{array}\right\}.$$

Given a map $\beta\in\Omega_f$, we define Ω_β to be the homotopy class of β in Ω_f. The width W associated with Ω_β can be defined in the same way, namely

$$W=W(\Omega_\beta):=\inf_{\gamma\in\Omega_\beta}\max_{s\in[0,1]}\mathrm{Area}(\gamma(\cdot,s)). \tag{2}$$

Then $W\geq\max(\mathrm{Area}(\overline{v}_0),\mathrm{Area}(\overline{v}_1))>0$. The next result is a slight variant of Theorem 0.1.

Theorem 0.3 *Given $\beta\in\Omega_f$ with $W=W(\Omega_\beta)>\max(\mathrm{Area}(\overline{v}_0),\mathrm{Area}(\overline{v}_1))$, there is a sequence of sweepouts $\gamma^j\in\Omega_\beta$ with $\max_{s\in[0,1]}E(\gamma^j(\cdot,s))\to W$ such that for any given $\epsilon>0$, there exist $\bar{J}$ and $\delta>0$ such that if $j>\bar{J}$ and*

$$\mathrm{Area}(\gamma^j(\cdot,s))>W-\delta,$$

then there exist finitely many harmonic disks $u_k\colon D\to\mathcal{N}$ with $u_k(\partial D)\subset\Gamma$ and finitely many (possibly empty) harmonic spheres $\hat{u}_l\colon S^2\to\mathcal{N}$ with

$$d_V\Big(\gamma^j(\cdot,s),\bigcup_k\{u_k\}\cup\bigcup_l\{\hat{u}_l\}\Big)<\epsilon.$$

Moreover, when considered as restricted maps from ∂D to Γ, only one map among the u_k has degree 1, whereas all others have degree 0. We also have the energy identity $\sum_k\mathrm{Area}(u_k)+\sum_l\mathrm{Area}(\hat{u}_l)=W$.

Remark We postpone the discussions of this result until Section 8.

We also want to mention the min-max theory for constructing minimal submanifolds with free boundary using geometric measure theory. In the 1960s, Almgren [2] initiated a program to develop a Morse theory for minimal submanifolds (with or without free boundary), and he obtained in [3] the existence of an integral varifold which is stationary with free boundary in the sense of first variation in any dimension and codimension; see more details in [31]. Later along this direction, higher regularity was established for hypersurfaces. In particular, Grüter and Jost [19] proved the existence of an unstable embedded minimal disk inside any compact *convex* domain in $\mathbb{R}^3$. Later, Jost [24, Theorem 4.1] generalized their work to any compact three-manifold which is diffeomorphic to a ball with mean convex boundary. Higher-dimensional results were developed very recently by De Lellis and Ramic [10] (for both free and fixed boundary problems in convex manifolds; see also Montezuma [33]), and by Li and Zhou [31] (for the free boundary problem in any compact manifolds with boundary). We refer to Marques and Neves [32] for other recent developments of the min-max theory using geometric measure theory.

Sketch of main ideas

Here we provide a brief summary of our main new ideas. Though the main scheme follows the approach laid out by Colding–Minicozzi [6] for minimal spheres (see also Zhou [58; 60] for minimal surfaces with higher genus), the presence of free boundary in our setting brings in several main new obstacles.

For the analytic aspect, there are two main ingredients that we have to establish for weakly harmonic maps with (partial) free boundary. The first ingredient is a version of energy convexity which says that the energy functional is strictly convex near a weakly harmonic map with mixed Dirichlet and free boundaries on the half 2–disk. Unlike the proof in [6] where Colding–Minicozzi used the moving frame method developed by Hélein [20] in order to get a Hardy-type estimate for weakly harmonic maps with Dirichlet boundary and small energy on the 2–disk D, we first use the Uhlenbeck–Rivière decomposition method developed by Rivière [41] to get a refined ϵ–regularity (gradient estimate) for weakly harmonic maps with mixed Dirichlet and free boundaries on the half 2–disk (Theorem 2.5), and then appeal to the first-order Hardy inequality (Lemma 2.4); cf Lamm and Lin [27], where the energy density $|\nabla u|^2$ is estimated in the local Hardy space $h^1(D)$. The key geometric observation in the proof of the energy convexity lies in two orthogonality conditions: one observed by Colding–Minicozzi [6], and the other by Zhou [59]. Very recently, this idea also permitted the first author and Lin [29] to obtain an energy convexity and uniqueness for *weakly intrinsic biharmonic maps* defined on the unit 4–ball with small bi-energy, which in particular yields a version of uniqueness for weakly harmonic maps in dimension 4. The other ingredient is the uniform continuity estimate up to the free boundary for weakly harmonic maps. Whereas the continuity up to the Dirichlet boundary was proven by Qing [39], we prove the uniform continuity for weakly harmonic maps with mixed Dirichlet and free boundaries by a careful covering argument using our previous gradient estimate.

We think that the results and techniques in both of the ingredients are of independent interest for the free boundary problem of harmonic maps.

The variational construction consists mainly of two parts. The first part is a two-stage tightening process, following closely the approach of Colding–Minicozzi [6] (this can be made a 2^k–stage process when the parameter space is k–dimensional). The main challenges here include the $W^{1,2} \cap C^0$–continuity for (partial) free boundary harmonic replacements, and an interpolation construction to prove energy improvement inequalities. In fact, the key idea in both places is to construct comparison maps and then use the energy minimality conditions. One natural candidate of comparison maps is the linear interpolation in $\mathbb{R}^N$ of two maps $u_i \colon D \to \mathcal{N}$, $i = 1, 2$. Though such a map will go outside of $\mathcal{N}$, Colding–Minicozzi in [6] used the nearest-point projection to pull it back to $\mathcal{N}$. However the projection does not necessarily map the image of ∂D into Γ. Therefore we have to develop new methods to overcome this issue. To prove the $W^{1,2}$–continuity, we find a way to reduce the problem to an interpolation between two curves; and in the second place, we choose to do interpolation in Fermi coordinates. Based on the two new results, the tightening process can then be carried through in the free boundary setting analogously to [6].

The second part is a bubbling convergence procedure for almost harmonic maps with free boundary. The bubbling convergence for almost harmonic maps on spheres was developed by Colding–Minicozzi [6, Appendix B], and for free boundary α–harmonic maps it has been investigated systematically by Fraser [15]. Our result can be viewed as a combination of the two results. Among several things, the most novel observation in this part is the asymptotic analysis for harmonic maps with free boundary defined on a long half-cylinder (which is conformally equivalent to a thin half-annulus). In particular, we prove that the angular energy is much smaller than the total energy of this map. Since boundary terms will appear in the integration by parts argument, we have to use our gradient estimate together with a delicate doubling argument to take care of these boundary terms, so as to carry out Colding–Minicozzi's method (which works on cylinders).

Layout of the paper

The paper is organized as follows. We fix some notation in Section 1. In Section 2, we prove the energy convexity for weakly harmonic maps defined on the half-disk with mixed Dirichlet and free boundaries (Theorem 2.2). In Section 3, we prove that a weakly harmonic map with mixed Dirichlet and free boundaries on the half-disk with continuous partial Dirichlet boundary is also continuous on the whole free boundary including the corner points; and we also present the proof on the existence of a weakly harmonic map with free boundary and with prescribed partial Dirichlet boundary, which we call the *partial free boundary harmonic replacement*. In Section 4, we outline the main ingredients needed to establish the min-max theory in the free boundary setting. In Section 5, we prove that the min-max values for the area and energy functionals are the same by using conformal reparametrizations. In Section 6, we carry out the construction of the tightening process; in particular, we show how to use a two-stage harmonic replacement procedure to make the sweepout as tight as possible. In Section 7, we prove that any min-max sequence of maps will converge to a bubble tree consisting of harmonic disks and harmonic spheres; here we also show that the bubbling convergence satisfies the energy identity, or equivalently, the total energy of the bubble tree is the same as the min-max value. In Section 8, we point out necessary changes needed to adapt our theory to the min-max Plateau problem (Theorem 0.3).

Acknowledgements The authors would like to thank Professor William Minicozzi for encouragement and helpful comments, and also Professors Rugang Ye and Tony Tromba for helpful discussions. Zhou is partially supported by NSF grant DMS-1811293.

Added in proof We were informed that Laurain and Petrides have recently proved a similar result, and Theorem 2.2 was announced in a seminar proceedings paper [28], but without a proof.

1 Notation

We first fix some notation.

- $\mathbb{R}^2$ denotes the Euclidean two-plane, where (x_1, x_2) (or (x, y)) and (r, θ) are the Cartesian and polar coordinates, respectively.
- $\mathbb{H}^2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\}$ denotes the upper half-plane.
- $D_r = D_r(0) = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 < r^2\}$ denotes the disk of radius r centered at the origin.
- $D_s^+ := D_s \cap \mathbb{H}^2 = \{(r, \theta) : 0 \le r < s,\ 0 < \theta < \pi\}$ denotes the upper half-disk with radius s.
- For simplicity, we sometimes write $D^+ = D_1^+$.

Then we write $\partial D_s^+ = \partial_s^C \cup \partial_s^A$, which is the union of the chord (diameter) and the arc of the upper semicircle, ie

$$\partial_s^C = \{(r,\theta) : 0 \le r \le s,\ \theta = 0 \text{ or } \pi\} \quad \text{and} \quad \partial_s^A = \{(r,\theta) : r \equiv s,\ 0 \le \theta \le \pi\}.$$

Similarly, we will write

$$\partial^C = \partial_1^C \quad \text{and} \quad \partial^A = \partial_1^A.$$

We write $u\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ if u is a map from D to $\mathcal{N}$ and $u(\partial D) \subset \Gamma$. Similarly, we write $u\colon (D_s^+, \partial_s^C) \to (\mathcal{N}, \Gamma)$ if u is a map from D_s^+ to $\mathcal{N}$ and $u(\partial_s^C) \subset \Gamma$.

Given $u\colon D \to \mathcal{N}$ or $u\colon D_s^+ \to \mathcal{N}$, the Dirichlet energy is defined as

$$E(u) := \frac{1}{2} \int_{D \text{ or } D_s^+} |\nabla u|^2 \, dx\, dy.$$

The Euler–Lagrange equation for the Dirichlet energy is the so-called *harmonic map equation*, which is a quasilinear system defined by

$$-\Delta u = A(u)(\nabla u, \nabla u), \tag{3}$$

where $A(u)$ is the second fundamental form of the embedding $\mathcal{N} \hookrightarrow \mathbb{R}^N$. See eg [20].

Definition 1.1 A $W^{1,2}$–map $u\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ is called a *weakly harmonic map with free boundary* if u satisfies the harmonic map equation weakly in D and

$$\frac{\partial u}{\partial r} \perp \Gamma \quad \text{along } \partial D.$$

Similarly a $W^{1,2}$–map $u\colon (D_s^+, \partial_s^C) \to (\mathcal{N}, \Gamma)$ is called a *weakly harmonic map with partial free boundary* if u satisfies the harmonic map equation weakly in D_s^+ and

$$\frac{\partial u}{\partial y} \perp \Gamma \quad \text{along } \partial_s^C.$$

2 Energy convexity for weakly harmonic maps with partial free boundary

In this section, we present the first main result, that is, the energy convexity and uniqueness for weakly harmonic maps with mixed Dirichlet and free boundaries. This is not only one of the key ingredients of our min-max existence theory for minimal disks with free boundary, but it also has its own interest from the point of view of PDEs and calculus of variations.

We first recall Colding–Minicozzi's energy convexity for weakly harmonic maps defined on the 2–disk D with Dirichlet boundary and small energy; see also [27].

Theorem 2.1 [6, Theorem 3.1] *There exists a constant $\varepsilon_0 > 0$, depending only on $\mathcal{N}$, such that if $u, v \in W^{1,2}(D,\mathcal{N})$, $u = v$ on ∂D, and u is weakly harmonic with $E(u) \le \varepsilon_0$, then*

$$\frac{1}{2}\int_D |\nabla v - \nabla u|^2\,dx \le \int_D |\nabla v|^2\,dx - \int_D |\nabla u|^2\,dx. \tag{4}$$

In this section, we prove a free boundary analogue of Theorem 2.1. We will abuse notation and still denote the energy threshold as ε_0. More precisely, we prove:

Theorem 2.2 (energy convexity for weakly harmonic maps with mixed Dirichlet and free boundaries) *There exists a constant $\varepsilon_0 > 0$, depending only on $\mathcal{N}$ and Γ, such that if $u, v \in W^{1,2}(D^+,\mathcal{N})$ with $u|_{\partial\mathcal{A}} = v|_{\partial\mathcal{A}}$, $u|_{\partial\mathcal{C}} \subset \Gamma$, $v|_{\partial\mathcal{C}} \subset \Gamma$, $E(u) \le \varepsilon_0$, and u is a weakly harmonic map with partial free boundary, then we have the energy convexity*

$$\frac{1}{2}\int_{D^+} |\nabla v - \nabla u|^2\,dx \le \int_{D^+} |\nabla v|^2\,dx - \int_{D^+} |\nabla u|^2\,dx. \tag{5}$$

An immediate corollary of Theorem 2.2 is the uniqueness of weakly harmonic maps with mixed Dirichlet and free boundaries on D^+.

Corollary 2.3 *There exists an $\varepsilon_0 > 0$, depending only on $\mathcal{N}$ and Γ, such that for any two weakly harmonic maps $u, v \in W^{1,2}(D^+,\mathcal{N})$ with Dirichlet boundary $u|_{\partial\mathcal{A}} = v|_{\partial\mathcal{A}}$ and free boundaries $u|_{\partial\mathcal{C}} \subset \Gamma$ and $v|_{\partial\mathcal{C}} \subset \Gamma$, if their energies satisfy $E(u) \le \varepsilon_0$ and $E(v) \le \varepsilon_0$, then we have $u \equiv v$ in D^+.*

In order to prove Theorem 2.2, we will use the following first-order Hardy inequality and appeal to the refined ϵ–regularity for weakly harmonic maps with mixed Dirichlet and free boundaries (Theorem 2.5).

Lemma 2.4 (Hardy inequality) *Let u, v be in $W^{1,2}(D^+,\mathbb{R}^N)$ with $u|_{\partial\mathcal{A}} = v|_{\partial\mathcal{A}}$. Then we have*

$$\int_{D^+} |v-u|^2 \cdot \frac{1}{(1-|x|)^2}\,dx \le 4\int_{D^+} |\nabla(v-u)|^2\,dx, \tag{6}$$

where $x = (x_1, x_2) \in D^+$.

Proof First we extend u and v across ∂^C by reflection, setting $\tilde{u}(x_1,x_2)=u(-x_1,x_2)$ and $\tilde{v}(x_1,x_2)=v(-x_1,x_2)$ for $x_1\le 0$, to get $\tilde{v}-\tilde{u}\in W_0^{1,2}(D)$. Approximate $\tilde{v}-\tilde{u}$ in $W^{1,2}$ by a sequence of smooth functions with compact support $\tilde{w}_i\in C_c^\infty(D)$, and let w_i be the restriction of $\tilde{w}_i$ to D^+. Now for each w_i we have

$$\begin{aligned}\text{(7)}\quad \int_{D^+}\frac{|w_i|^2}{(1-|x|)^2}\,dx &= \int_0^\pi\int_0^1 \frac{|w_i|^2 r}{(1-r)^2}\,dr\,d\theta \\ &= \int_0^\pi\left[\frac{|w_i|^2 r}{1-r}\bigg|_0^1 - \int_0^1\left(\frac{|w_i|^2}{1-r}+\frac{2w_i\cdot(w_i)_r r}{1-r}\right)dr\right]d\theta \\ &\le 2\left(\int_0^\pi\int_0^1\frac{|w_i|^2 r}{(1-r)^2}\,dr\,d\theta\right)^{\frac12}\left(\int_0^\pi\int_0^1|(w_i)_r|^2\,r\,dr\,d\theta\right)^{\frac12} \\ &\le 2\left(\int_{D^+}\frac{|w_i|^2}{(1-|x|)^2}\,dx\right)^{\frac12}\left(\int_{D^+}|\nabla w_i|^2\,dx\right)^{\frac12},\end{aligned}$$

which yields

$$\text{(8)}\qquad \int_{D^+}\frac{|w_i|^2}{(1-|x|)^2}\,dx \le 4\int_{D^+}|\nabla w_i|^2\,dx.$$

Now (6) follows from (8) and Fatou's lemma. □

The next result is a refined ϵ–regularity for weakly harmonic maps on D^+ with Dirichlet and free boundaries; cf [46; 30; 26] and Fraser [15]. This ϵ–regularity is crucial to the proof of Theorem 2.2. We shall remark that such ϵ–regularity is well-known for weakly harmonic maps defined on the 2–disk D with Dirichlet boundary, see eg Qing [39, Lemma 4].

Theorem 2.5 *There exists a constant $\varepsilon_0>0$, depending only on $\mathcal{N}$ and Γ, such that if $u\in W^{1,2}(D^+,\mathcal{N})$ is a weakly harmonic map with mixed Dirichlet boundary on ∂^A and free boundary $u|_{\partial^C}\subset\Gamma$ and $E(u)\le\varepsilon_0$, then for any $x\in D^+\cup(\partial^C)^\circ$ we have*

$$\text{(9)}\qquad |\nabla u|(x)\le\frac{C\sqrt{\varepsilon_0}}{1-|x|}$$

for some constant $C>0$ that only depends on $\mathcal{N}$ and Γ.

Proof By the reflection across the free boundary Γ constructed by Scheven in [46], $u\in W^{1,2}(D^+)$ can be extended to $\tilde{u}\in W^{1,2}(D)$ so that $\tilde{u}$ weakly solves in D the system of equations

$$\text{(10)}\qquad \operatorname{div}(Q\nabla\tilde{u})=\omega\cdot Q\nabla\tilde{u}$$

for some $\omega = (\omega^i_j)_{1\le i,j\le n} \in L^2(D, \mathrm{so}(n)\otimes \bigwedge^1\mathbb{R}^2)$ (ie $\omega^j_i = -\omega^i_j$) and $Q \in W^{1,2}\cap L^\infty(D, \mathrm{GL}(n,\mathbb{R}))$ such that

$$|\omega| \le C|\nabla\tilde{u}| \quad \text{a.e. in } D \qquad \text{and} \qquad \|Q\|_{L^\infty(D)} + \|Q^{-1}\|_{L^\infty(D)} \le C, \tag{11}$$

where $C > 0$ is a constant depending only on $\mathcal{N}$ and Γ; see Jost, Liu and Zhu [26, Proposition 3.3]. Moreover, by the assumption $E(u) \le \varepsilon_0$ and the reflection construction, we have

$$\|\nabla\tilde{u}\|_{L^2(D)} \le C\sqrt{\varepsilon_0} \quad \text{and} \quad \|\omega\|_{L^2(D)} \le C\sqrt{\varepsilon_0}. \tag{12}$$

Then, using the Uhlenbeck–Rivière decomposition method developed by Rivière [41] and Rivière–Struwe [44] in the study of regularity of elliptic PDEs with antisymmetric structure, we can obtain a constant $\alpha > 0$ such that (cf [42, page 50])

$$\sup_{0<\rho<\frac{1}{4},\, p\in D_{1/2}} \rho^{-\alpha}\int_{D_\rho(p)} |\Delta\tilde{u}|\,dx \le C\sqrt{\varepsilon_0}. \tag{13}$$

Proof of (13) To see this, we first note that since $\tilde{u}-\bar{\tilde{u}}$ also satisfies (10) with the same Q and ω, where $\bar{\tilde{u}} = \frac{1}{\pi}\int_D \tilde{u}\,dx$ is the average of $\tilde{u}$ on D, without loss of generality we may assume that $\bar{\tilde{u}} = 0$. Then by the work of Sharp (see eg [47, Corollary 1.4, Proposition 3.1]) on the higher integrability for solutions to a system of PDEs similar to (10), we know that there exists a constant $C > 0$ such that

$$\|\nabla^2\tilde{u}\|_{L^1(D_{1/2})} \le C\|\tilde{u}\|_{L^1(D)} \le C\|\nabla\tilde{u}\|_{L^1(D)} \le C\sqrt{\varepsilon_0}. \tag{14}$$

See [48, Theorem 1.2] and [26, Theorem 2.4]. Here we have used the Poincaré–Wirtinger inequality in the second inequality. We note that this higher integrability (14) essentially follows from the stability of the local Hardy space $h^1(D)$ (see eg Lamm and Lin [27, Section A.2]) under multiplication by Hölder continuous functions, coupled with the Hölder continuity of $\tilde{u}$ in D proved by Scheven [46, Theorem 4.1].

Then by the continuous embedding of $W^{1,1}(D)$ into $L^{2,1}(D)$, where $L^{2,1}$ is the Lorentz space (see eg [22; 38; 55; 20]), we have

$$\|\nabla\tilde{u}\|_{L^{2,1}(D_{1/2})} \le C\|\nabla\tilde{u}\|_{W^{1,1}(D_{1/2})} \le C\|\nabla\tilde{u}\|_{L^1(D)} \le C\sqrt{\varepsilon_0}. \tag{15}$$

Now for any $p \in D_{\frac{1}{2}}$ and $0 < r < \frac{1}{2}$, we use the Hodge decomposition (see eg [23, Corollary 10.5.1]) to find $A \in W^{1,2}(D_r(p),\mathbb{R}^n)$ and $B \in W^{1,2}_0(D_r(p),\mathbb{R}^n)$ such that

$$Q\nabla\tilde{u} = \nabla A + \nabla^\perp B \quad \text{in } D_r(p), \tag{16}$$

where $\nabla^\perp := (-\partial_{x_2}, \partial_{x_1})$ and we have

$$\|\nabla A\|_{L^2(D_r(p))} + \|\nabla B\|_{L^2(D_r(p))} \le \widetilde{C}\|\nabla\widetilde{u}\|_{L^2(D_r(p))}.$$

Then we have, taking divergence on both sides of the first equation in (16),

$$\Delta A = \operatorname{div}(Q\nabla\widetilde{u}) = \omega\cdot Q\nabla\widetilde{u} \quad \text{in } D_r(p), \tag{17}$$

and, taking curl on both sides of the first equation in (16),

$$\begin{cases} \Delta B = \nabla^\perp Q\cdot\nabla\widetilde{u} & \text{in } D_r(p), \\ \quad B = 0 & \text{on } \partial D_r(p). \end{cases} \tag{18}$$

Now let $A = A_1 + A_2$ on $D_r(p)$ so that $\Delta A_2 = 0$ and

$$\begin{cases} \Delta A_1 = \operatorname{div}(Q\nabla\widetilde{u}) & \text{in } D_r(p), \\ \quad A_1 = 0 & \text{on } \partial D_r(p). \end{cases} \tag{19}$$

Then using [20, Theorem 3.3.3] (which implies that the standard L^p–theory extends to Lorentz spaces), we furthermore get

$$\|\nabla A_1\|_{L^{2,1}(D_r(p))} \le C\|Q\nabla\widetilde{u}\|_{L^{2,1}(D_r(p))} \le C\|\nabla\widetilde{u}\|_{L^{2,1}(D_r(p))}. \tag{20}$$

Hence by [20, Theorem 3.3.4] we conclude that

$$\begin{aligned} \|A_1\|_{L^\infty(D_r(p))} &\le C(\|A_1\|_{L^{2,1}(D_r(p))} + \|\nabla A_1\|_{L^{2,1}(D_r(p))}) \\ &\le C\|\nabla A_1\|_{L^{2,1}(D_r(p))} \\ &\le C\sqrt{\varepsilon_0}, \end{aligned} \tag{21}$$

where we again used [20, Theorem 3.3.3] (which ensures that the Poincaré inequality extends to Lorentz spaces) and the fact that $A_1 = 0$ on $\partial D_r(p)$ in the second estimate. Thus, an integration by parts yields

$$\begin{aligned} \|\nabla A_1\|^2_{L^2(D_r(p))} &= -\int_{D_r(p)} A_1\Delta A_1\,dx \\ &\le \|A_1\|_{L^\infty(D_r(p))}\int_{D_r(p)}|\omega\cdot Q\nabla\widetilde{u}|\,dx \\ &\le C\sqrt{\varepsilon_0}\|\nabla\widetilde{u}\|^2_{L^2(D_r(p))}, \end{aligned} \tag{22}$$

where we have used (12) and the fact that $Q \in W^{1,2}\cap L^\infty(D, \mathrm{GL}(n,\mathbb{R}))$. Note that again by integration by parts (using $A_1 = 0$ on $\partial D_r(p)$ and $\Delta A_2 = 0$) we have

$$\|\nabla A_2\|^2_{L^2(D_r(p))} = \|\nabla A\|^2_{L^2(D_r(p))} - \|\nabla A_1\|^2_{L^2(D_r(p))} \le \widetilde{C}\|\nabla\widetilde{u}\|^2_{L^2(D_r(p))}. \tag{23}$$

Now since A_2 is harmonic on $D_r(p)$ we know that for every $q \in D_r(p)$ the function

$$\rho \to \frac{1}{\rho^2} \int_{D_\rho(q)} |\nabla A_2|^2 \, dx$$

is increasing; see eg [42, Lemma IV.1]. Now let $\overline{C} > 0$ be such that $\|Q^{-1}\|_{L^\infty(D)} \le \overline{C}$ and

$$\delta = \min\left\{ \frac{1}{4\sqrt{\widetilde{C}\,\overline{C}}}, \frac{1}{2} \right\},$$

where $\widetilde{C}$ is from (23). Then we have

$$\int_{D_{\delta r}(p)} |\nabla A_2|^2 \, dx \le \frac{1}{16\widetilde{C}\,\overline{C}} \int_{D_r(p)} |\nabla A_2|^2 \, dx \le \frac{1}{16\overline{C}} \|\nabla \widetilde{u}\|^2_{L^2(D_r(p))}. \tag{24}$$

Now using (18) and by the results of Coifman–Lions–Meyer–Semmes [5] or Wente's lemma [56], we know

$$\|B\|_{L^\infty(D_r(p))} + \|\nabla B\|_{L^{2,1}(D_r(p))} \le C \|\nabla \widetilde{u}\|^2_{L^2(D_r(p))}. \tag{25}$$

Therefore, combining (16), (22), (24) and (25) we have

$$\begin{aligned} &\|\nabla \widetilde{u}\|^2_{L^2(D_{\delta r}(p))} \\ &\le \|Q^{-1}\|_{L^\infty(D)} \big(2\|\nabla A_1\|^2_{L^2(D_r(p))} + 2\|\nabla A_2\|^2_{L^2(D_{\delta r}(p))} + \|\nabla B\|^2_{L^2(D_r(p))} \big) \\ &\le C\sqrt{\varepsilon_0} \|\nabla \widetilde{u}\|^2_{L^2(D_r(p))} + \tfrac{1}{8} \|\nabla \widetilde{u}\|^2_{L^2(D_r(p))} + C \|\nabla \widetilde{u}\|^4_{L^2(D_r(p))}. \end{aligned} \tag{26}$$

Choosing ε_0 sufficiently small we arrive at

$$\|\nabla \widetilde{u}\|^2_{L^2(D_{\delta r}(p))} \le \tfrac{1}{2} \|\nabla \widetilde{u}\|^2_{L^2(D_r(p))} \tag{27}$$

for any $p \in D_{\frac{1}{2}}$ and $0 < r < \frac{1}{2}$. Iterating this inequality gives the existence of a constant $\alpha > 0$ such that for all $p \in D_{\frac{1}{2}}$ and all $\rho < \frac{1}{4}$, one has

$$\rho^{-2\alpha} \int_{D_\rho(p)} |\nabla \widetilde{u}|^2 \, dx \le C \int_D |\nabla \widetilde{u}|^2 \, dx \le C\varepsilon_0. \tag{28}$$

Now by (10) we have

$$\Delta \widetilde{u} = Q^{-1} \big(-\nabla Q \nabla \widetilde{u} + \omega \cdot Q \nabla \widetilde{u} \big).$$

Then using (11), for all $p \in D_{\frac{1}{2}}$ and all $\rho < \frac{1}{4}$ we have

$$\begin{aligned}(29)\quad \rho^{-\alpha}\int_{D_\rho(p)}|\Delta\tilde{u}|\,dx &\le C\rho^{-\alpha}\int_{D_\rho(p)}|\nabla Q||\nabla\tilde{u}|+|\nabla\tilde{u}|^2\,dx\\ &\le C\rho^{-\alpha}\bigg[\Big(\int_{D_\rho(p)}|\nabla Q|^2\,dx\Big)^{\frac{1}{2}}\Big(\int_{D_\rho(p)}|\nabla\tilde{u}|^2\,dx\Big)^{\frac{1}{2}}+\int_{D_\rho(p)}|\nabla\tilde{u}|^2\,dx\bigg]\\ &\le C\sqrt{\varepsilon_0}.\end{aligned}$$

Here we have used $\|\nabla Q\|_{L^2(D)}\le C$ for some $C>0$ depending only on $\mathcal{N}$ and Γ. This proves (13). □

Now using (13), a classical estimate on Riesz potentials then gives, for all $p\in D_{\frac{1}{4}}$,

$$(30)\qquad |\nabla\tilde{u}|(p)\le C\frac{1}{|x-p|}*\chi_{D_{1/2}}\,|\Delta\tilde{u}|+C\|\nabla\tilde{u}\|_{L^2(D)},$$

where $\chi_{D_{1/2}}$ is the characteristic function of the ball $D_{\frac{1}{2}}$. Together with injections proved by Adams in [1] (see also [18, Exercise 6.1.6]), the latter shows that

$$\|\nabla\tilde{u}\|_{L^r(D_{1/4})}\le C\sqrt{\varepsilon_0}$$

for some $r>2$. Reinjecting this into the equation (10) and bootstrapping the estimates, we get

$$\|\nabla\tilde{u}\|_{L^\infty(D_{1/8})}\le C\sqrt{\varepsilon_0}.$$

In particular, $|\nabla\tilde{u}|(0)\le C\sqrt{\varepsilon_0}$. Then by a scaling argument we get the desired estimate (9). □

Now we are ready to prove Theorem 2.2.

Proof of Theorem 2.2 In order to prove the energy convexity (5), it suffices to show

$$(31)\qquad \Psi\ge-\frac{1}{2}\int_{D^+}|\nabla(v-u)|^2\,dx,$$

where (using the boundary conditions and the harmonic map equation)

$$\begin{aligned}(32)\quad \Psi&:=\int_{D^+}|\nabla v|^2\,dx-\int_{D^+}|\nabla u|^2\,dx-\int_{D^+}|\nabla(v-u)|^2\,dx\\ &=2\int_{D^+}\langle\nabla(v-u),\nabla u\rangle\,dx\\ &=2\int_{D^+}\langle v-u,\ A(u)(\nabla u,\nabla u)\rangle\,dx+2\int_{\partial C}\Big\langle v-u,\ \frac{\partial u}{\partial\nu}\Big\rangle\,ds.\end{aligned}$$

Here $\nu = (0, -1)$ is the outward unit normal to ∂^C. Now we note that for any $p, q \in \mathcal{N}$ (resp. Γ), there exists a constant $C > 0$ depending only on $\mathcal{N}$ (resp. Γ) such that

$$|(p-q)^{\perp}| \le C|p-q|^2,$$

where the superscript $\perp$ denotes the normal component of the vector $p-q$ at q; see eg [7, Lemma A.1]. Therefore, using

$$-\Delta u = A(u)(\nabla u, \nabla u) \perp T_u\mathcal{N},$$

the Cauchy–Schwarz inequality together with (32) yields

$$\Psi \ge -C_1 \int_{D+} |v-u|^2 |\nabla u|^2 \, dx - C_2 \int_{\partial C} |v-u|^2 \left|\frac{\partial u}{\partial \nu}\right| ds, \tag{33}$$

where $C_1 > 0$ depends only on $\mathcal{N}$ and $C_2 > 0$ depends only on Γ. Now by Lemma 2.4, Theorem 2.5 and using the facts that $v = u$ on ∂^A and $x_2 = 0$ on $\partial^C = [-1, 1] \times \{0\}$, we have

$$\begin{aligned}
\Psi &\ge -C_3\varepsilon_0 \int_{D+} \frac{|v-u|^2}{(1-|x|)^2} \, dx - C_4\sqrt{\varepsilon_0} \int_{\partial C} \frac{|v-u|^2}{1-|x_1|} \, ds \\
&\ge -4C_3\varepsilon_0 \int_{D+} |\nabla(v-u)|^2 \, dx + C_4\sqrt{\varepsilon_0} \int_{-1}^{1} \int_0^{\sqrt{1-x_1^2}} \partial_{x_2}\left(\frac{|v-u|^2}{1-|x_1|}\right) dx_2 \, dx_1 \\
&= -4C_3\varepsilon_0 \int_{D+} |\nabla(v-u)|^2 \, dx + C_4\sqrt{\varepsilon_0} \int_{D+} \frac{\langle v-u, (v-u)_{x_2}\rangle}{1-|x_1|} \, dx \\
&\ge -4C_3\varepsilon_0 \int_{D+} |\nabla(v-u)|^2 \, dx - \frac{1}{4}\int_{D+} |\nabla(v-u)|^2 \, dx - C_5\sqrt{\varepsilon_0} \int_{D+} \frac{|v-u|^2}{(1-|x|)^2} \, dx \\
&\ge -C_6\sqrt{\varepsilon_0} \int_{D+} |\nabla(v-u)|^2 \, dx - \frac{1}{4}\int_{D+} |\nabla(v-u)|^2 \, dx.
\end{aligned}$$

In the second-to-last inequality we have used Young's inequality. Choosing ε_0 sufficiently small that $C_6\sqrt{\varepsilon_0} \le \frac{1}{4}$, we get the desired estimate (31). □

3 Existence and regularity of partial free boundary harmonic replacement

In this section, we discuss the existence and regularity of the (partial) free boundary version of the harmonic replacement. The fixed boundary harmonic replacement was discussed in Colding–Minicozzi [6].

3.1 Continuity of weakly harmonic maps with partial free boundary

In this section, we prove the C^0 regularity up to the boundary for weakly harmonic maps with mixed Dirichlet and free boundaries and small energy. For weakly harmonic maps with fixed continuous Dirichlet boundary, this C^0 boundary regularity was proved by Qing [39].

Let $u\colon (D^+, \partial^C) \to (\mathcal{N}, \Gamma)$ be a weakly harmonic map with mixed Dirichlet and free boundaries and small energy $E(u) \le \varepsilon_0$ (where ε_0 is from Theorem 2.5) such that the Dirichlet boundary $u\colon \partial^A \to \mathcal{N}$ is *continuous*. Then by Helein's interior regularity [20], u is smooth in D^+; by Qing's (Dirichlet) boundary regularity [39], u is continuous up to the interior of the Dirichlet boundary on ∂^A; moreover, by the (free) boundary regularity result of Scheven [46], u is also smooth up to the interior of the free boundary on ∂^C. So the only thing left to be verified is that u is continuous up to the two corner points $(1,0)$ and $(1,\pi)$ (in polar coordinates on $\overline{D^+}$), ie the endpoints of the Dirichlet boundary on ∂^A or the free boundary on ∂^C.

It suffices to prove the continuity around $p=(1,\pi)$.[1] We will first prove that there exists a sequence of points $\{x_i\}$ on ∂^C, converging to p, such that the $u(x_i)$ are all close to $u(p)$; next we show that all the intermediate points in $u([x_i, x_{i+1}])$ are also close to $u(p)$. For the convenience of later proof, we will parallell shift D^+ in $\mathbb{R}^2$ so that $p=(0,0)$. Now $D^+ = \{(x,y) \in \mathbb{R}^2 : y>0, (x-1)^2+y^2<1\}$ and $\partial^A = \{(x,y) \in \mathbb{R}^2 : y>0, (x-1)^2+y^2=1\}$, and we will use polar coordinates on $\mathbb{R}^2$.

We first need a variant of the Courant–Lebesgue lemma (cf [39, Lemma 2]) at the corner point p.

Lemma 3.1 *Given $0<l<\frac{1}{2}$, there exists $l' \in (l, 2l)$ such that*

$$\int_0^{\alpha(l')} \left|\frac{\partial u}{\partial \theta}\right|^2 (l', \theta)\, d\theta \le \frac{1}{\log 2} E(u|_{(D_{2l}(0)\setminus D_l(0))\cap D^+}), \tag{34}$$

where $\alpha(l')$ is the angle such that $(l', \alpha(l')) \in \partial^A$. Consequently, we have

$$|u(l',\theta_1) - u(l',\theta_2)| \le \sqrt{\frac{N}{\log 2}}\, E(u|_{(D_{2l}(0)\setminus D_l(0))\cap D^+})^{\frac{1}{2}} |\theta_1-\theta_2|^{\frac{1}{2}}, \tag{35}$$

where N is the dimension of the ambient space of the embedding of $\mathcal{N}$ into $\mathbb{R}^N$.

[1] Given Scheven's reflection construction and equation (10), one may also use the Uhlenbeck–Rivière decomposition method as in Müller–Schikorra [35] to prove boundary regularity.

Proof The proof is virtually the same as the proof of [39, Lemma 2]:

$$E(u|_{(D_{2l}(0)\setminus D_l(0))\cap D^+}) \geq \int_l^{2l}\int_0^{\alpha(r)} \left(\left|\frac{\partial u}{\partial r}\right|^2 + \frac{1}{r^2}\left|\frac{\partial u}{\partial \theta}\right|^2 \right) r\, d\theta\, dr,$$

where $\alpha(r)$ is the angle such that

$$(r, \alpha(r)) \in \partial^A = \{(x, y) \in \mathbb{R}^2 : y > 0,\ (x-1)^2 + y^2 = 1\}.$$

Therefore there exists $l' \in (l, 2l)$ such that

$$\int_0^{\alpha(l')} \left|\frac{\partial u}{\partial \theta}\right|^2 (l', \theta)\, d\theta \leq \frac{1}{\log 2} E(u|_{(D_{2l}(0)\setminus D_l(0))\cap D^+}).$$

Then we have

$$\begin{aligned} |u(l', \theta_1) - u(l', \theta_2)| &\leq \left| \int_{\theta_1}^{\theta_2} \frac{\partial u}{\partial \theta}(l', \theta)\, d\theta \right| \\ &\leq \sqrt{\frac{N}{\log 2}} E(u|_{(D_{2l}(0)\setminus D_l(0))\cap D^+})^{\frac{1}{2}} |\theta_1 - \theta_2|^{\frac{1}{2}}. \end{aligned}$$ □

Now we are ready to prove the continuity up to the corners on the boundary of D^+.

Theorem 3.2 *There exists an $\varepsilon_0 > 0$ such that if $u: (D^+, \partial^C) \to (\mathcal{N}, \Gamma)$ is a weakly harmonic map with mixed Dirichlet and free boundaries with $E(u) \leq \varepsilon_0$, and u is continuous on ∂^A, then u is continuous on $\overline{D^+}$.*

Proof Since u is continuous on ∂^A, by the discussion prior to Lemma 3.1 it suffices to show that given $\varepsilon > 0$, there exists $\delta > 0$ such that for $x \in \partial^C$ and $|x - p| \leq \delta$ we have $|u(x) - u(p)| \leq \varepsilon$. By the continuity of u on ∂^A, we can find $\delta_1 > 0$ such that if $p_l = (l, \alpha(l)) \in \partial^A$ and $|p_l - p| \leq \delta_1$, then $|u(p_l) - u(p)| \leq \frac{\varepsilon}{10}$.

Now we identify the unit interval $[0, 1]$ with half of $\partial^C = [-1, 1] \times \{0\} \subset \mathbb{R}^2$ and abuse notation so that $x \in [0, 1]$ represents a point on ∂^C. Consider the decomposition

$$[0, 1] = \bigcup_{k=0}^{\infty} [2^{-k-1}, 2^{-k}] = \bigcup_{k=0}^{\infty} I_k.$$

Note that we have

$$\sum_{k=0}^{\infty} E(u|_{(D_{2^{-k-1}}\setminus D_{2^{-k}})\cap D^+}) \leq \varepsilon_0.$$

Then there is some $K_1 > 0$ such that for all $k \geq K_1$, we have

$$E(u|_{(D_{2^{-k-1}}\setminus D_{2^{-k}})\cap D^+}) \leq \frac{\varepsilon^2}{1000N^2}.$$

Now using Lemma 3.1, for any $k \geq K_1$ we can pick $x_k \in I_k$ such that

$$|u(x_k) - u(p)| \leq |u(x_k) - u(p_{|x_k|})| + |u(p_{|x_k|}) - u(p)| \leq \tfrac{\varepsilon}{8}.$$

Now let us consider a family of half-disks $D_k^+ \subset D^+$, where D_k^+ has corner points $(2^{-k-2}, 0)$ and $(2^{-k} + 2^{-k-2}, 0)$. Then the center of D_k^+ is $y_k := (2^{-k-1} + 2^{-k-2}, 0)$ and the radius of D_k^+ is $r_k := 2^{-k-1}$. Each D_k^+ can only overlap with at most four of the other D_k^+, so we have

$$\sum_{k=0}^{\infty} E(u|_{D_k^+}) \leq 4\varepsilon_0.$$

Thus, for any $\varepsilon_0 > 0$, there exists a $K_2 > 0$ such that for any $k \geq K_2$ we have $E(u|_{D_k^+}) \leq \varepsilon_0$. Then by rescaling and Theorem 2.5, for any $y \in [2^{-k-1}, 2^{-k}]$ with $k \geq K_2$ we have

$$|\nabla u(y)| \leq C 2^{k+1} \sqrt{\varepsilon_0}. \tag{36}$$

Then integration along the interval between y and x_k gives

$$|u(y) - u(x_k)| \leq 2C\sqrt{\varepsilon_0}.$$

If we choose ε_0 small enough (depending only on ε), then for any $y \in [2^{-k-1}, 2^{-k}]$ with $k \geq K_2$ we have $|u(y) - u(x_k)| \leq \frac{\varepsilon}{4}$. So for all $y \in (0, 2^{-K})$, where $K = \max\{K_1, K_2\}$, we have

$$|u(y) - u(p)| \leq \varepsilon.$$

We conclude the continuity of u at p and hence finish the proof of the theorem. □

As a corollary of Theorem 3.2, we have:

Corollary 3.3 *A weakly harmonic map u with mixed Dirichlet and free boundaries on D^+ is continuous on $\overline{D^+}$ provided that u is continuous on ∂^A.*

3.2 Existence of (partial) free boundary harmonic replacement

Theorem 3.4 *There exists $\varepsilon_0 > 0$ such that for any $v \in C^0(\overline{D^+}, \mathcal{N}) \cap W^{1,2}(D^+, \mathcal{N})$ with $v(\partial^C) \subset \Gamma$ and $E(v) \leq \varepsilon_0$, there exists a unique harmonic map*

$$u \in C^0(\overline{D^+}, \mathcal{N}) \cap W^{1,2}(D^+, \mathcal{N})$$

such that $E(u) \leq \varepsilon_0$, $u = v$ on ∂^A, $u(\partial^C) \subset \Gamma$ and u meets Γ orthogonally along ∂^C.

Remark 3.5 The map u in Theorem 3.4 is usually called the (partial) *free boundary harmonic replacement* of v.

Proof of Theorem 3.4 Combining Hélein's interior regularity for weakly harmonic maps on two-dimensional domains, Qing's (Dirichlet) boundary regularity [39], the free boundary regularity of Scheven [46] and Theorem 3.2, we know that the (partial) free boundary harmonic replacement u of v, if it exists, is smooth in D^+ and continuous in $\overline{D^+}$.

The (partial) free boundary harmonic replacement can be constructed as follows: suppose $v \in C^0(\overline{D^+}, \mathcal{N}) \cap W^{1,2}(D^+, \mathcal{N})$ is such that $v(\partial^C) \subset \Gamma$ and $E(v) \le \varepsilon_0$. Let us define the space $\mathcal{F}$ to be the space of maps

$$\mathcal{F} := \{w \in W^{1,2}(D^+, \mathcal{N}) : w|_{\partial^A} = v|_{\partial^A},\ w(\partial^C) \subset \Gamma\}.$$

Now we choose an energy-minimizing sequence $u^i \in \mathcal{F}$, ie

$$\lim_{i \to \infty} E(u^i) = \inf\{E(w) : w \in \mathcal{F}\}.$$

By the Rellich compactness theorem, we can find a subsequence of $\{u^i\}$ that weakly converges to a $W^{1,2}$ map u. By first variation of the energy functional at u we know that u is a weakly harmonic map with mixed Dirichlet and free boundaries on D^+. Then indeed we know that $\{u^i\}$ converges strongly in $W^{1,2}$ to u by the energy convexity Theorem 2.2 and Poincaré inequality Lemma A.2. □

4 Overview of the variational approach

In this section, we provide an overview of the proof of Theorem 0.1.

As in the proof of the classical Plateau problem, the area functional is too weak to control the maps, so we have to change gear to the energy functional. The next result says that one can take an approximating sequence of sweepouts in a given homotopy class so that their maximal energy converges to the maximal area. In particular, we have:

Theorem 4.1 *Given $\beta \in \Omega$ with $W = W(\Omega_\beta) > 0$, there exists a sequence of sweepouts $\gamma^j \in \Omega_\beta$ with*

$$\max_{s \in [0,1]} E(\gamma^j(\cdot, s)) \searrow W.$$

Remark The proof is given in Section 5. Note that we have

$$W \le \max_{s \in [0,1]} \operatorname{Area}(\gamma^j(\cdot, s)) \le \max_{s \in [0,1]} E(\gamma^j(\cdot, s)) \searrow W.$$

The next ingredient is a tightening theorem. We first fix some notation concerning balls in D.

Definition 4.2 A *generalized ball* B in the unit disk D is either an interior ball of D or the intersection with D of a ball of $\mathbb{R}^2$ centered at some boundary point of ∂D. That is: $B \subset D$ or $B = B_r(x) \cap D$ where $x \in \partial D$ and $r < \frac{1}{2}$. We will call the first case a *classical ball* and the second case a *boundary ball*.

Given $\rho > 0$, when B is a classical ball we let $\rho B \subset D$ denote the ball with the same center as B and radius ρ times that of B; when B is a boundary ball, we can define $\rho B \subset D$ as follows: there exists a unique fractional linear transformation Π_B from D to the upper half-plane $\mathbb{H}^2$ such that $\Pi_B(B) = D^+$, and ρB is defined as $\Pi_B^{-1}(D_\rho^+)$.

Remark In the following, we will frequently identify a boundary ball B with its image $\Pi_B(B)$ in $\mathbb{H}^2$.

The following result plays the role of deformation lemma in nonlinear analysis; see eg [54].

Theorem 4.3 *Given $\beta \in \Omega$ with $W = W(\Omega_\beta) > 0$, there exist a constant $\epsilon_1 > 0$ and a continuous function $\Psi\colon [0,\infty) \to [0,\infty)$ with $\Psi(0) = 0$, both depending on $(\mathcal{N},\Gamma)$, such that given any $\widetilde{\gamma} \in \Omega$ with no nonconstant harmonic slices other than $\widetilde{\gamma}(\cdot\,,0)$ and $\widetilde{\gamma}(\cdot\,,1)$, there exists $\gamma \in \Omega_{\widetilde{\gamma}}$ such that $E(\gamma(\cdot\,,t)) \le E(\widetilde{\gamma}(\cdot\,,t))$ for each t, and for each t with $E(\widetilde{\gamma}(\cdot\,,t)) \ge \frac{1}{2}W$ we have the following property:*

(∗) *If $\mathcal{B}$ is any finite collection of disjoint generalized closed balls in D with*

$$\int_{\bigcup_{\mathcal{B}} B} |\nabla\gamma(\cdot\,,t)|^2 < \varepsilon_1,$$

and if $v\colon \bigcup_{\mathcal{B}} \frac{1}{8}B \to \mathcal{N}$ is the free boundary harmonic replacement of $\gamma(\cdot\,,t)$ on $\bigcup_{\mathcal{B}} \frac{1}{8}B$, then

$$\int_{\bigcup_{\mathcal{B}} \frac{1}{8}B} |\nabla\gamma(\cdot\,,t) - \nabla v|^2 \le \Phi\big(E(\widetilde{\gamma}(\cdot\,,t)) - E(\gamma(\cdot\,,t))\big).$$

Remark The proof is given in Section 6.

We also need the following compactness result. Let ε_{SU} and ε_F be the small thresholds (depending only on $\mathcal{N}$ and $(\mathcal{N},\Gamma)$, respectively) from [39, Lemma 4] (cf [45, Main Estimate 3.2]) and Theorem 2.5 (cf [15, Proposition 1.7]), respectively, giving us interior estimates for harmonic maps or free boundary harmonic maps with energy less than ε_{SU} and ε_F, respectively.

Theorem 4.4 *Let $\epsilon_1, E_0 > 0$ be such that $\epsilon_1 < \min\{\varepsilon_{SU}, \varepsilon_F\}$, and suppose that $u^j\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ is a sequence of maps in $C^0(\overline{D}, \mathcal{N}) \cap W^{1,2}(D, \mathcal{N})$ with $E_0 \geq E(u^j)$ which satisfies*

$$\text{Area}(u^j) > E(u^j) - \frac{1}{j} \tag{37}$$

as well as the following condition:

(†) *For any finite collection $\mathcal{B}$ of disjoint generalized closed balls in D with*

$$\int_{\bigcup_{\mathcal{B}} B} |\nabla u^j|^2 < \epsilon_1,$$

if $v\colon \bigcup_{\mathcal{B}} \frac{1}{8} B \to \mathcal{N}$ is the harmonic replacement of u^j on $\bigcup_{\mathcal{B}} \frac{1}{8} B$, then

$$\int_{\bigcup_{\mathcal{B}} \frac{1}{8} B} |\nabla u^j - \nabla v|^2 \leq \frac{1}{j}.$$

Then a subsequence of the varifolds for the maps u^j converges to a collection of harmonic maps $v_0, \dots, v_m\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ with free boundary, and harmonic spheres $\widetilde{v}_1, \dots, \widetilde{v}_l\colon S^2 \to \mathcal{N}$. Moreover, the energy identity holds:

$$\sum_{i=0}^{m} E(v_i) + \sum_{k=1}^{l} E(\widetilde{v}_k) = \lim_{j\to\infty} E(u^j).$$

Remark The proof is given in Section 7. Note that $E(v_i) = \text{Area}(v_i)$ and $E(\widetilde{v}_j) = \text{Area}(\widetilde{v}_j)$.

Now we prove the main Theorem 0.1 using Theorems 4.1, 4.3 and 4.4.

Proof of Theorem 0.1 Choose a sequence $\widetilde{\gamma}^j \in \Omega_\beta$ as in Theorem 4.1 and assume that

$$\max_{t\in[0,1]} E(\widetilde{\gamma}^j(\cdot, t)) < W + \frac{1}{j}.$$

We can slightly change the parametrization so that $\widetilde{\gamma}^j$ maps a small open subset of D to a point, so that each slice cannot be harmonic unless it is a constant map; we refer to [6, footnote 8] for details. Applying Theorem 4.3 to $\widetilde{\gamma}^j$ gives a sequence $\gamma^j \in \Omega_\beta$. We will show that $\{\gamma^j\}$ has the desired properties.

Let us argue by contradiction. Let $\mathcal{G}^W$ be the collection of harmonic maps from $S^2 \to \mathcal{N}$ and free boundary harmonic maps $(D, \partial D) \to (\mathcal{N}, \Gamma)$ such that the sum of the energies is exactly W. Suppose $\{\gamma^j\}$ does not have the desired property, which means that there exists some $\varepsilon > 0$ such that there exist $j_k \to \infty$ and $s_k \in [0,1]$ with $d_V(\gamma^{j_k}(\cdot, s_k), \mathcal{G}^W) \geq \varepsilon$ and $\mathrm{Area}(\gamma^{j_k}(\cdot, s_k)) > W - \frac{1}{k}$. Then by $E(u) \geq \mathrm{Area}(u)$ we get $\lim_{k\to\infty} E(\gamma^{j_k}(\cdot, s_k)) = W$, and

$$E(\tilde{\gamma}^{j_k}(\cdot, s_k)) - E(\gamma^{j_k}(\cdot, s_k)) \leq E(\tilde{\gamma}^{j_k}(\cdot, s_k)) - \mathrm{Area}(\gamma^{j_k}(\cdot, s_k)) \leq \frac{1}{k} + \frac{1}{j_k} \to 0.$$

Since the tightening process decreases the energy, we get

$$E(\gamma^{j_k}(\cdot, s_k)) - \mathrm{Area}(\gamma^{j_k}(\cdot, s_k)) \to 0 \quad \text{as } k \to \infty.$$

By Theorem 4.3 we have that if $\mathcal{B}$ is any finite collection of disjoint generalized closed balls in D with

$$\int_{\bigcup_{\mathcal{B}} \boldsymbol{B}} |\nabla \gamma^{j_k}(\cdot, s_k)|^2 < \varepsilon_1,$$

and if $v\colon \bigcup_{\mathcal{B}} \frac{1}{8} B \to \mathcal{N}$ is the harmonic replacement of $\gamma^{j_k}(\cdot, s_k)$ on $\bigcup_{\mathcal{B}} \frac{1}{8} B$, then

$$\int_{\bigcup_{\mathcal{B}} \frac{1}{8} \boldsymbol{B}} |\nabla \gamma^{j_k}(\cdot, s_k) - \nabla v|^2 \leq \Phi\Big(\frac{1}{k} + \frac{1}{j_k}\Big) \to 0.$$

Therefore applying Theorem 4.4 gives a subsequence of the $\gamma^{j_k}(\cdot, s_k)$ that varifold-converges to a collection of desired harmonic disks with free boundary and harmonic spheres. The energy identity implies that the sum of the energies of the limit is exactly W. However this contradicts $d_V(\gamma^{j_k}(\cdot, s_k), \mathcal{G}^W) \geq \varepsilon$. This finishes the proof of Theorem 0.1. □

5 Conformal parametrization

We now prove Theorem 4.1. The main idea of the proof follows [6, Appendix D]; see also [58, Section 3] and [60, Section 3]. For a given sweep-out, we will find conformal reparametrizations of the regularization of this sweep-out, so that the energy of each slice of this family can not be much larger than its area.

Proof of Theorem 4.1 First we claim that for a given sweep-out $\tilde{\gamma}(\cdot, t) \in \Omega_{\tilde{\gamma}}$, we can find a regularized sweep-out $\gamma(\cdot, t) \in \Omega_{\tilde{\gamma}}$ which lies in $C^0([0,1], C^2(D, \mathcal{N}))$ as a map of t and is such that $\gamma(\cdot, t)$ is close to $\tilde{\gamma}(\cdot, t)$ uniformly in the $W^{1,2} \cap C^0$–norm for all $t \in [0,1]$. This follows from a standard argument using mollification just like [6, Lemma D.1]. Here we only point out the necessary modifications of [6, Lemma D.1] to make sure that the images of ∂D under each slice $\gamma(\cdot, t)$ lie in the constraint submanifold Γ. In particular, near the boundary ∂D we can first enlarge the domain D

to $D_{1+\alpha}$ (for small $\alpha > 0$) by reflecting $\tilde{\gamma}(\cdot, t)$ across Γ in the Fermi coordinates around Γ (cf [15]) and then do classical mollification under the Fermi coordinates. The mollified maps when restricted to D will map ∂D to Γ by our construction. By the classical mollification result [13, page 250, Theorem 1] and Lemma A.1, we can make sure that the obtained maps are close to $\tilde{\gamma}(\cdot, t)$ uniformly in $W^{1,2} \cap C^0$ near ∂D. In the interior of D, we can just mollify $\tilde{\gamma}(\cdot, t)$ in $\mathbb{R}^N$. To combine them, we can choose a partition of unity to glue these two mollifications together, and by the same argument as [13, page 252, Theorem 3] the obtained maps are also close to $\tilde{\gamma}(\cdot, t)$ uniformly in $W^{1,2} \cap C^0(D)$. Up to this step, under the obtained maps the boundary ∂D goes into Γ, but the interior of D may get out of $\mathcal{N}$. Finally, one can follow [6, Lemma D.1] to project these maps to $\mathcal{N}$ using the nearest-point projection to get the desired $\gamma(\cdot, t)$. By choosing the mollification parameter small enough, we can make sure $\max_{t\in[0,1]}\|\gamma(\cdot, t) - \tilde{\gamma}(\cdot, t)\|$ is as small as we want. Note that an explicit homotopy between γ and $\tilde{\gamma}$ is given by letting the mollification parameter go to 0. So we finish the sketch of the proof of the claim.

Then $\gamma(\cdot, t)$ induces a continuous one-parameter family of C^1 metrics

$$g(t) = \gamma(t)^*(\text{metric on } \mathcal{N})$$

on D. This family of metrics may be degenerate, so we define the perturbed metrics as $\tilde{g}(t) = g(t) + \varepsilon g_0$, where g_0 is the standard flat metric on D. Then by the same methods as in [6, Lemma D.6], [58, Proposition 3.1] and [60, Proposition 3.4], we can construct a family of conformal reparametrizations $h_t\colon D_{g_0} \to D_{\tilde{g}(t)}$ (which fix three given points on ∂D) that varies continuously in $C^0 \cap W^{1,2}(D, D)$. Then with the conformality we can control the energy:

$$\begin{aligned} E(\gamma(\cdot, t) \circ h_t) &= E(h_t\colon D_{g_0} \to D_{g(t)}) \\ &\le E(h_t\colon D_{g_0} \to D_{\tilde{g}(t)}) = \text{Area}(D_{\tilde{g}(t)}) \\ &= \int_D \big(\det(g_0^{-1} g(t)) + \varepsilon \operatorname{Tr}(g_0^{-1} g(t)) + \varepsilon^2\big)^{\frac{1}{2}}\, d\text{Vol}_{g_0} \\ &\le \text{Area}(D_{g(t)}) + \pi\big(\varepsilon^2 + 2\varepsilon \sup_t |g_0^{-1} g(t)|\big)^{\frac{1}{2}}. \end{aligned} \tag{38}$$

Choose $\varepsilon > 0$ so that $\pi(\varepsilon^2 + 2\varepsilon \sup_t |g_0^{-1} g(t)|)^{\frac{1}{2}} < \frac{\delta}{2}$, and we get

$$E(\gamma(\cdot, t) \circ h_t) - \text{Area}(D_{g_t}) \le \tfrac{\delta}{2}.$$

If $\tilde{\gamma}^j$ is a sequence of sweepouts in Ω_β, then $\gamma^j \circ h_t$ constructed as above is a sequence as desired. □

6 Construction of the tightening process

This section is devoted to the proof of Theorem 4.3.

6.1 Continuity of harmonic replacement

In this part we want to prove that the free boundary harmonic replacement process is actually continuous as a map from $C^0(\overline{D^+}) \cap W^{1,2}(D^+)$ to itself, if we restrict to maps with small energy. This generalizes the continuity of the harmonic replacement process on classical interior balls by Colding–Minicozzi [6].

Theorem 6.1 *Let ε_0 be as in Theorem 2.2 and set*

$$\mathcal{M} = \{u \in C^0(\overline{D^+}, \mathcal{N}) \cap W^{1,2}(D^+, \mathcal{N}) : E(u) \leq \varepsilon_0,\ u(\partial^C) \subset \Gamma\}.$$

Given $u \in \mathcal{M}$, let $w \in \mathcal{M}$ be the unique free boundary harmonic replacement of u (produced by Theorem 3.4). *Then $u \to w$ is continuous as a map from $C^0(\overline{D^+}) \cap W^{1,2}(D^+)$ to itself.*

6.1.1 $W^{1,2}$–continuity First we prove that this map is $W^{1,2}$ continuous. In [6], Colding–Minicozzi's idea was to construct a comparison map by interpolating in $\mathbb{R}^N$ between two maps of the same boundary value in $W^{1,2}(D, \mathcal{N})$, and then projecting the interpolation back to $\mathcal{N}$. In our free boundary setting, however, such an interpolation–projection trick may not leave the image of ∂^C lying inside Γ. Here we prove the $W^{1,2}$–continuity by contradiction. The main idea is that if the $W^{1,2}$–continuity fails, then we can find a sequence u_i converging to u_∞, but the replacements w_i of the u_i have energy strictly greater than the energy of w_∞. In this scenario, we are able to construct some comparison maps v_i sharing the same fixed boundary value with w_i, but having energy strictly smaller than that of w_i for i large. This contradicts the uniqueness of free boundary harmonic maps with small energy; see Corollary 2.3.

The key ingredient in our proof is the construction of certain new comparison maps. We first collect a few preliminary results for the construction.

There are five components in the comparison map. One of the components consists of "small" free boundary harmonic maps. The following lemma shows that the energy of these maps is actually small.

Lemma 6.2 *There exists a $\delta_0 > 0$ such that if a Lipschitz map $f\colon [0,\pi]\to\mathcal{N}$ satisfies $f(0)\in\Gamma$, $f(\pi)\in\Gamma$ and $\int_0^\pi |f'|^2 \le \delta \le \delta_0$, then there exists a map $v\colon D^+\to\mathcal{N}$ such that $v(1,\theta)=f(\theta)$, $v(\partial^C)\subset\Gamma$ and $E(v)\le C\delta$ for some universal constant $C>0$ depending only on $\mathcal{N}$, Γ and δ_0.*

Proof By the gradient bound of f and the fact that $f(0), f(\pi)\in\Gamma$, we can assume that the image of f lies in the Fermi coordinates $\{y^1,\dots,y^n\}$ around Γ, where Lemma A.1 in the appendix applies. We can further assume that locally Γ is identified as a subset of $\{y^{k+1}=\dots=y^n=0\}$ and $f(0)=0$. Define

$$v(r,\theta)=r\cdot f(\theta),$$

where $r\cdot f(\theta)$ is the scalar multiplication under Fermi coordinates; hence v is a map from $D^+\to\mathcal{N}$, and $v(r,0)$ and $v(r,\pi)$ both lie in Γ for any $r\in[0,1]$, thus $v(\partial^C)\subset\Gamma$.

It is only left to check that the energy of v is small. By Lemma A.1, we only need to check that the energy of v is small as a map to the Euclidean space $\mathbb{R}^N$. In particular, Lemma A.1 implies that $\int_0^\pi |f'|^2\le(1+\alpha)\delta$ under the Euclidean metric for some small $\alpha>0$. Also, as $f(0)=0$, we get $|f(\theta)|\le\sqrt{\pi(1+\alpha)\delta}$ for $\theta\in[0,\pi]$. Now we have the energy estimates for v as a map into standard Euclidean space $\mathbb{R}^N$

$$\begin{aligned}E(v)&=\int_0^1\int_0^\pi\left(\left|\frac{\partial v}{\partial r}\right|^2+\frac{1}{r^2}\left|\frac{\partial v}{\partial\theta}\right|^2\right)r\,d\theta\,dr\\&\le\int_0^1\int_0^\pi(|f|^2+|f'|^2)\,r\,d\theta\,dr\le C\delta\end{aligned}\tag{39}$$

for some universal constant $C>0$. This completes the proof. □

Another component in our comparison map is a modified interpolation band. First we recall a lemma in [6] to construct the interpolation of band between two circles.

Lemma 6.3 *There exists $\delta_0>0$ such that for $\delta<\delta_0$ the following statement holds. Let $f,g\colon[0,\pi]\to\mathcal{N}$ be two C^0 maps such that*

$$|f-g|\le\delta,\quad \int|f'|^2\le\delta'\quad\text{and}\quad\int|g'|^2\le\delta'$$

for some $\delta'>0$. Then there exists a $\rho\in\left(0,\frac12\right]$ and a map $\tilde v\colon D^+\setminus D^+_{1-\rho}\to\mathcal{N}$ such that

$$\tilde v(1-\rho,\theta)=f(\theta),\quad \tilde v(1,\theta)=g(\theta)\quad\text{and}\quad E(\tilde v)\le C\delta^{\frac12}\delta'^{\frac12}$$

for some constant $C>0$ depending only on $\mathcal{N}$.

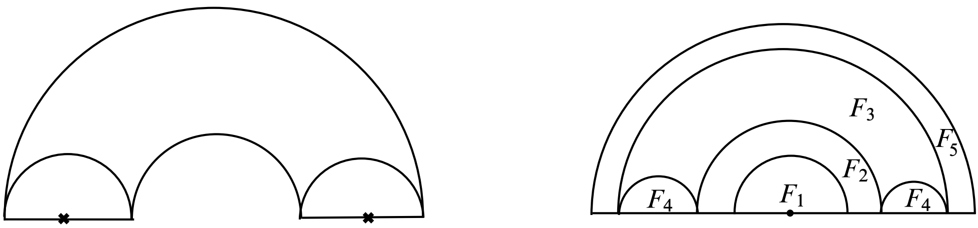

Figure 1

Proof We refer the proof to [6, Lemma 3.11], where the construction works for the whole circles, but it can be generalized to half-circles without any modification. Note that our assumption here is even stronger. □

With this construction on hand, we can construct an interpolation between two arcs on a modified band. Let us first define a modified band. A *modified band* $MB_{a,b}$ is defined as

(40) $$MB_{a,b} = \{(r,\theta) : r \in [a,b], \theta \in [0,\pi]\} \setminus \bigcup_{\xi=0,\pi} \{(r,\theta) : \big|(r,\theta) - \big(\tfrac{1}{2}(b+a), \xi\big)\big| < \tfrac{1}{2}(b-a)\}.$$

See Figure 1. In the following context we may use ϕ to denote the angle parameter of the arcs of the removed disks.

Lemma 6.4 *There exists $\delta_0 > 0$ such that for $\delta < \delta_0$ the following statement holds. Let $f, g\colon [0,\pi] \to \mathcal{N}$ be two C^0 maps such that $\|f-g\|_{L^\infty} \le \delta$, $\int |f'|^2 \le \delta'$ and $\int |g'|^2 \le \delta'$ for some $\delta' > 0$. Then there exists $\rho \in \big(0, \frac{1}{2}\big]$ and a map $v\colon MB_{1-\rho,1} \to \mathcal{N}$ such that $v(1-\rho,\theta) = f(\theta)$, $v(1,\theta) = g(\theta)$ and $E(v) \le C\delta^{\frac{1}{2}}\delta'^{\frac{1}{2}}$ for some constant C depending only on $\mathcal{N}$ and ρ. Moreover, for the arc of the removed half-disks of the modified band* (reparametrized by angle parameter ϕ)*, we have*

$$\int_0^\pi |v'(\phi)|^2\, ds(\phi) \le C\delta^2,$$

where $ds(\phi)$ is the intrinsic arc-length integral.

Proof We construct a change of variables. Construct a map $h\colon D^+ \setminus D^+_{1-\rho} \to MB_{1-\rho,1}$ as follows: given $s \in [1-\rho, 1]$, we consider the arc $MB_{1-\rho,1} \cap \{r \equiv s\}$ and denote the length of this arc by l_s; for any fixed $(r,\theta) \in D^+ \setminus D^+_{1-\rho}$, let h map it to (r,θ'), where the ratios satisfy

$$\left[\left|\theta-\tfrac{\pi}{2}\right|:\tfrac{\pi}{2}\right]=\left[\left|\theta'-\tfrac{\pi}{2}\right|:l_r\right].$$

Then the Jacobian and the gradient of this change of variables is bounded by some universal constant. Let $\tilde{v}$ be given by Lemma 6.3; then $v=\tilde{v}\circ(h^{-1})$ is the desired interpolation map. The only thing we need to check is the last estimate. In particular, the energy of v on the boundary of the removed half-disks is bounded by that of $\tilde{v}$ on $\theta=0$ times a universal constant, and the energy of $\tilde{v}$ therein is bounded. □

Now we start proving that the harmonic replacement $u\to w$ is continuous as a map from $C^0\cap W^{1,2}$ to $W^{1,2}$ for u with small energy.

Proof of $W^{1,2}$–continuity We prove by contradiction. Suppose the $W^{1,2}$–continuity of the harmonic replacement process fails. Then we can find a sequence of maps $\{u_i\colon D^+\to\mathcal{N}\}$ converging to u_∞ in $C^0(\overline{D^+},\mathcal{N})\cap W^{1,2}(D^+,\mathcal{N})$, but whose free boundary harmonic replacements $\{w_i\}$ do not converge to the corresponding replacement w_∞ (of u_∞) in $W^{1,2}$. Note that a subsequence of w_i (still denoted by w_i) must converge weakly to w_∞, so by lower-semicontinuity of energy, $E(w_\infty)\le\liminf E(w_i)$.

The fact that w_i does not converge to w_∞ in $W^{1,2}$ implies that there exists $\varepsilon>0$ and a subsequence (still denoted by w_i) such that $\|w_i-w_\infty\|^2_{W^{1,2}}\ge 2\varepsilon$ holds, ie

$$\int\langle w_i-w_\infty,w_i-w_\infty\rangle+\langle\nabla(w_i-w_\infty),\nabla(w_i-w_\infty)\rangle\ge 2\varepsilon.$$

The weak convergence implies that $\|w_i-w_\infty\|_{L^2}\to 0$ and $\int\langle\nabla w_i,\nabla w_\infty\rangle\to E(w_\infty)$. Thus we get $E(w_i)-E(w_\infty)\ge\varepsilon$ for i large.

We divide D^+ into five different pieces, $D^+=\bigcup_{j=1}^5 F_j$, where

(i) $F_1=\{(r,\theta):r\in[0,\eta\lambda(1-\rho)]\}$,

(ii) $F_2=\{(r,\theta):r\in[\eta\lambda(1-\rho),\lambda(1-\rho)]\}$,

(iii) $F_3=MB_{\lambda(1-\rho),\lambda}$,

(iv) $F_4=\{y\in D^+:\left|y-\left(\lambda-\tfrac12\lambda\rho,0\right)\right|\le\tfrac12\lambda\rho\}\cup\{y\in D^+:\left|y-\left(\lambda-\tfrac12\lambda\rho,\pi\right)\right|\le\tfrac12\lambda\rho\}$,

(v) $F_5=\{(r,\theta):r\in[\lambda,1]\}$.

See Figure 1. Here $\eta, \rho, \lambda \in (0,1)$ will be determined later. Note that the F_j have some common boundaries, and we will see in the construction below that the maps on the common boundaries share the same value. Let us choose a radius $r_0 > \frac{1}{2}$ such that $\int_{D^+ \setminus D^+_{1-r_0}} |\nabla u_\infty|^2 \le \frac{\varepsilon}{16}$. Then given i large enough that $\|u_i - u_\infty\|^2_{W^{1,2}} \le \frac{\varepsilon}{16}$, we have $\int_{D^+ \setminus D^+_{1-r_0}} |\nabla u_i|^2 \le \frac{\varepsilon}{8}$. Using the coarea formula, we can pick some radius $r_1 \in (r_0, 1)$ such that $\int_{r=r_1} |\partial_\theta u_\infty|^2 \le C(r_0)\varepsilon$ and $\int_{r=r_1} |\partial_\theta u_i|^2 \le C(r_0)\varepsilon$.

Construction of v_i

- On F_1, we define v_i to be w_∞ after rescaling to a suitable scale to fit F_1.
- On F_2, we define v_i to be the part of u_∞ defined on $D^+ \setminus D^+_{1-r_1}$ after inversion (in polar coordinates, we change r to $2 - r_1 - r$) and rescaling. Note that the energy of v_i on F_2 is bounded by the energy of u_∞ on $D^+ \setminus D^+_{1-r_1}$ times a universal constant 4.
- On F_3, we define v_i to be the interpolation between $u_\infty|_{\{r=r_1\}}$ and $u_i|_{\{r=r_1\}}$ from Lemma 6.4 on the modified band, then rescale to fit F_3.
- On F_4, we define v_i to be the map constructed in Lemma 6.2 (up to rescaling), which provides two maps with small energy.
- On F_5, we define v_i to be the part of u_i defined on $D^+ \setminus D^+_{1-r_1}$ (up to rescaling).

Here ρ comes from Lemma 6.4, $\eta = 1 - r_1$ and $\lambda = 1 - r_1$. These parameters are chosen in order to guarantee that in the definition of the v_i, the rescalings are possible; ie the ratios between inner radius and outer radius of the bands or modified bands do not change.

Now we claim some properties of the v_i that we just constructed. First, on ∂^A the fixed boundary, $v_i = u_i$; on ∂^C the free boundary, v_i always has image in Γ. Second, v_i is continuous. So it is an eligible comparison map (with w_∞).

Finally we estimate the energy of v_i. We will repeatedly use (without mention) the fact that the energy is invariant under conformal reparametrization on the domain. Note that the energy of v_i on F_1 equals $E(w_\infty)$; on F_2 and F_5 the total energy of v_i is bounded by $\varepsilon/3$; on F_3 by Lemma 6.4, the energy of v_i is bounded by a constant times $\|u_\infty - u_i\|^{1/2}_{C^0}$, and on F_4 by Lemma 6.2 the energy of v_i is bounded by a constant times $\|u_\infty - u_i\|^{1/2}_{C^0}$. Combining all the pieces gives

$$E(v_i) \le E(w_\infty) + \tfrac{2\varepsilon}{3} + C\|u_\infty - u_i\|^{1/2}_{C^0}.$$

Since u_i converges to u_∞ in C^0, when i is large enough $C\|u_\infty - u_i\|_{C^0}^{1/2} \le \frac{\varepsilon}{6}$. In conclusion, when i is large enough, $E(v_i) \le E(w_\infty) + \frac{5\varepsilon}{6} \le E(w_i) - \frac{\varepsilon}{6} < E(w_i)$.

Note that v_i shares the same fixed boundary data with w_i, and thus we have constructed a map v_i with same fixed boundary data as w_i along ∂^A and $v_i(\partial^C) \subset \Gamma$, but v_i has energy strictly less than that of w_i on D^+; this is a contradiction to the energy-minimizing property of the free boundary harmonic replacement (Theorem 2.2). Therefore the harmonic replacement is continuous in $W^{1,2}$. □

6.1.2 C^0–continuity

Proof of C^0–continuity Now we are ready to prove the C^0–continuity of the harmonic replacement process. We argue by contradiction. If the harmonic replacement process is not C^0, we can find a sequence of maps $u_i\colon D^+ \to \mathcal{N}$ converging to u_∞ in $C^0(\overline{D^+}) \cap W^{1,2}(D^+)$ while for the corresponding free boundary harmonic replacements w_i (of u_i) we have $\|w_i - w_\infty\|_{C^0(\overline{D^+})} \ge \varepsilon > 0$.

From the proof of Theorem 3.2, we know that the w_i are equicontinuous on $\overline{D^+}$. Thus by the Arzelà–Ascoli theorem, up to a subsequence w_i must converge to some w'_∞ in $C^0(\overline{D^+})$. Now using the $W^{1,2}$–continuity of the harmonic replacement process that we have just proved, we have $\|w_i - w_\infty\|_{W^{1,2}(D^+)} \to 0$ as $i \to \infty$, which implies $w_\infty = w'_\infty$. This contradicts $\|w_i - w_\infty\|_{C^0(\overline{D^+})} \ge \varepsilon > 0$. □

6.2 Uniform continuity of energy improvement

In this part, we prove two inequalities regarding energy improvements for two sets of free boundary harmonic replacements on generalized balls; see Definition 4.2. These inequalities will play a key role in the proof of Theorem 4.3. Similar results were proved by Colding–Minicozzi [6, Section 3] for fixed boundary harmonic replacements. We suggest to first-time readers to come back to this section after reading Section 6.3.

Let ε_0 be the minimum of the ε_0 in Theorems 2.1 and 2.2. We adopt the following notation: given a map $u \in C^0(\bar{D}, \mathcal{N}) \cap W^{1,2}(D, \mathcal{N})$ with $u(\partial D) \subset \Gamma$ and a finite collection $\mathcal{B}$ of disjoint *generalized* closed balls in D such that the energy of u on $\bigcup_{\mathcal{B}} B := \bigcup_{B \in \mathcal{B}} B$ is at most $\frac{1}{3}\epsilon_0$, let $H(u, \mathcal{B})\colon D \to \mathcal{N}$ denote the map that coincides with u on $D \setminus \bigcup_{\mathcal{B}} B$ and is equal to the harmonic replacements of u on $\bigcup_{\mathcal{B}} B$; cf Theorem 2.1 for the harmonic replacement with fixed boundary used by Colding–Minicozzi, and see Remark 3.5 for the free boundary harmonic replacement, where the existence is guaranteed by Theorem 3.4. Note we also have $H(u,\mathcal{B})(\partial D) \subset \Gamma$, and without loss of generality we will call $H(u,\mathcal{B})$ the *free boundary harmonic replacement of u on $\mathcal{B}$*. Given two such disjoint collections $\mathcal{B}_1$ and $\mathcal{B}_2$, we use $H(u,\mathcal{B}_1,\mathcal{B}_2)$ to denote $H(H(u,\mathcal{B}_1),\mathcal{B}_2)$. Recall that for $\alpha \in (0,1]$, we denote by $\alpha\mathcal{B}$ the collection of concentric balls with radii that are shrunk by the factor α in the sense of Definition 4.2.

First we have the following interpolation formula for free boundary replacements; cf [6, Lemma 3.11]. Let κ denote the radius for which the Fermi coordinates system exists in a tubular neighborhood of radius κ surrounding Γ where Lemma A.1 in the appendix applies.

Lemma 6.5 *There exists $\tau > 0$ such that given two $C^0 \cap W^{1,2}$ maps $f, g\colon \partial^A_R \to \mathcal{N}$ with $f(0), f(\pi), g(0), g(\pi) \in \Gamma$, if f and g agree at one point on ∂^A_R and satisfy*

$$R\int_{\partial^A_R} |f'-g'|^2 \le \tau^2 \qquad \text{and} \qquad |f(\theta)-f(0)| \le \tfrac{\kappa}{3} \quad \text{for all } 0 \le \theta \le \pi,$$

then there exist $\rho \in (0, R/2]$ and a $C^0 \cap W^{1,2}$–map $w\colon D^+_R \setminus D^+_{R-\rho} \to \mathcal{N}$ such that

$$w(R-\rho,\theta) = f(\theta), \quad w(R,\theta) = g(\theta),$$

the image of $w|_{\partial^C_R \setminus \partial^C_{R-\rho}}$ lies in Γ, and the following estimate holds:

$$\int_{D^+_R \setminus D^+_{R-\rho}} |\nabla w|^2 \le C\left(R\int_{\partial^A_R} |f'|^2 + |g'|^2\right)^{\frac12}\left(R\int_{\partial^A_R} |f'-g'|^2\right)^{\frac12},$$

where $C > 0$ is a universal constant depending only on κ, τ, Γ and $\mathcal{N}$.

Remark 6.6 This lemma generalizes an interpolation formula by Colding–Minicozzi for maps defined on circles [6, Lemma 3.11] to maps defined on half-circles, and the constructed interpolating map has the chord boundary $\partial^C_R \setminus \partial^C_{R-\rho}$ lying on Γ. In [6, Lemma 3.11], Colding–Minicozzi first took the linear interpolation of f, g in $\mathbb{R}^N$, and then projected it back to the ambient manifold $\mathcal{N}$. However this method doesn't work in the free boundary setting because the projection $\mathbb{R}^N \to \mathcal{N}$ may not map the boundary ∂^C_R to the constraint submanifold Γ.

Here we use the Fermi coordinate system to construct the desired interpolation between f and g. In the proof we have to work in two different coordinate systems, and Lemma A.1 will be used to show the equivalence.

Proof of Lemma 6.5 There is one point $0 \le \theta \le \pi$ such that $f(\theta) = g(\theta)$. Choose $\tau = \kappa/3\pi$. Then by the assumptions and integrating, we get $|f(\theta) - g(\theta)| \le \frac{\kappa}{2}$ for all $\theta \in [0, \pi]$, so that $|g(\theta) - g(0)| \le \kappa$ for all $\theta \in [0, \pi]$.

As a result, we can assume the images of f and g both lie in a convex neighborhood U_κ of $f(0)$, and in U we can pick the Fermi coordinate systems $\{y^1, \dots, y^n\}$ as in Lemma A.1, where Γ is a subset of $\{y^{k+1} = \cdots = y^n = 0\}$. In Lemma A.1, g^1, g^2 and g^3 denote, respectively, the metric of $\mathcal{N}$, the flat metric in $\{y^1, \dots, y^n\}$ and the flat metric of $\mathbb{R}^N$. We will use $|\cdot|$ to denote the norm under the metric g^3.

Since the statement is scaling-invariant, it suffices to prove the case $R = 1$. For $\rho \in \left(0, \frac{1}{2}\right]$ to be determined, define $w\colon D^+ \setminus D^+_{1-\rho} \to \mathbb{R}^N$ by

$$w(r, \theta) = f(\theta) + \left(\frac{r + \rho - 1}{\rho}\right)(g(\theta) - f(\theta)).$$

On $\partial^C \setminus \partial^C_{1-\rho}$, since $f(0), f(\pi), g(0), g(\pi) \in \Gamma$, we know that in the Fermi coordinate chart, $f(0)$, $f(\pi)$, $g(0)$ and $g(\pi)$ lie in the plane $\{y^{k+1} = \cdots = y^n = 0\}$. As a result, the interpolation function w also has the image $w|_{\partial^C \setminus \partial^C_{1-\rho}}$ lying in the same plane, and thus lying in Γ.

The energy density in this coordinate system is

$$e(w) = \sum_{i,j=1}^{n} g^1_{ij}\left(\frac{\partial w^i}{\partial x}\frac{\partial w^j}{\partial x} + \frac{\partial w^i}{\partial y}\frac{\partial w^j}{\partial y}\right) \le (1+\alpha)\|\nabla w\|^2_{g^2}.$$

Now we proceed to prove the estimate of $\|\nabla w\|^2_{g^2}$. First we need a Wirtinger-type inequality. Suppose $h(\cdot)$ is a function on ∂^A and $h(s) = 0$ for some $s \in [0, \pi]$. Then for any $t \in [0, \pi]$, we have

$$|h(t)| = |h(t) - h(s)| \le \int_s^t |h'(x)|\, dx \le \int_{\partial^A} |h'|.$$

Integrating the square of both sides for $t \in [0, \pi]$, we get

$$\int_{\partial^A} |h(t)|^2 \le \pi \left(\int_{\partial^A} |h'|\right)^2 \le \pi^2 \int_{\partial^A} |h'|^2.$$

Then we get, setting $h = f - g$,

$$\int_{\partial^A} |f - g|^2 \le \pi^2 \int_{\partial^A} |f' - g'|^2.$$

Thus

$$
\begin{aligned}
E(D^+\backslash D^+_{1-\rho})
&\le \int_{D^+\backslash D^+_{1-\rho}} (1+\alpha)\|\nabla w\|^2_{g^2} \\
&\le (1+\alpha)\int_{1-\rho}^{1}\left(\frac{1}{\rho^2}\int_0^\pi \|f-g\|^2_{g^2}(\theta)\,d\theta+\frac{1}{r^2}\int_0^\pi (\|f'\|^2_{g^2}+\|g'\|^2_{g^2})(\theta)\,d\theta\right) r\,dr \\
&\le (1+\alpha)^2\int_{1-\rho}^{1}\left(\frac{1}{\rho^2}\int_0^\pi |f-g|^2(\theta)\,d\theta+\frac{1}{r^2}\int_0^\pi (|f'|^2+|g'|^2)(\theta)\,d\theta\right) r\,dr \\
&\le (1+\alpha)^2\left(\frac{16}{\rho}\int_0^\pi |f'-g'|^2(\theta)\,d\theta+2\rho\int_0^\pi (|f'|^2+|g'|^2)(\theta)\,d\theta\right) \\
&\le 18(1+\alpha)^2\left(\int_{\partial A}|f'|^2+|g'|^2\right)^{\frac{1}{2}}\left(\int_{\partial A}|f'-g'|^2\right)^{\frac{1}{2}},
\end{aligned}
$$

once we pick $\rho=\bigl(\int_{\partial A}|f'-g'|^2\bigr)^{\frac{1}{2}}/\bigl(\int_{\partial A}|f'|^2+|g'|^2\bigr)^{\frac{1}{2}}$. □

Now we are ready to prove the energy improvement inequalities for free boundary harmonic replacements. Similar results were first obtained by Colding–Minicozzi for fixed boundary harmonic replacements [6, Lemma 3.8]. Again we let ε_0 be the minimum of the ε_0 in Theorems 2.1 and 2.2.

Lemma 6.7 *There is a constant $k>0$ such that if $u\colon D\to\mathcal{N}$ is in $C^0\cap W^{1,2}$ and $\mathcal{B}_1$ and $\mathcal{B}_2$ are two finite collections of disjoint closed generalized balls in D such that the energy of u on each $\bigcup_{\mathcal{B}_i} B$ is at most $\frac{1}{3}\epsilon_0$, then*

$$E(u)-E(H(u,\mathcal{B}_1,\mathcal{B}_2))\ge k\bigl(E(u)-E\bigl(H\bigl(u,\tfrac{1}{2}\mathcal{B}_2\bigr)\bigr)\bigr)^2. \tag{41}$$

Furthermore, for any $\mu\in\left[\frac{1}{8},\frac{1}{2}\right]$ we have

$$\frac{\bigl(E(u)-E(H(u,\mathcal{B}_1))\bigr)^{\frac{1}{2}}}{k}+E(u)-E(H(u,2\mu\mathcal{B}_2))\ge E(H(u,\mathcal{B}_1))-E(H(u,\mathcal{B}_1,\mu\mathcal{B}_2)). \tag{42}$$

Proof The proof is analogous to the proof of [6, Lemma 3.8], and the main difficulty arises from those boundary balls. We will include the details for completeness, and focus on how we use the new interpolation result, ie Lemma 6.5.

Let $\mathcal{B}_1=\{B^1_\alpha\}$ and $\mathcal{B}_2=\{B^2_j\}$. We need to clarify the second replacement. Observe that the total energy of u on the union of the balls $\mathcal{B}_1\cup\mathcal{B}_2$ is at most $\frac{2}{3}\varepsilon_0$, and the free boundary harmonic replacement on $\mathcal{B}_1$ does not change the map outside these balls and is energy-nonincreasing, then it follows that the total energy of $H(u,\mathcal{B}_1)$ on $\mathcal{B}_2$ is at most $\frac{2}{3}\varepsilon_0$.

We will divide $\mathcal{B}_2$ into two disjoint subsets $\mathcal{B}_{2,+}$ and $\mathcal{B}_{2,-}$, set

$$\mathcal{B}_{2,+} = \left\{B_j^2 \in \mathcal{B}_2 : \tfrac{1}{2}B_j^2 \subset B_\alpha^1 \text{ for some } B_\alpha^1 \in \mathcal{B}_1\right\} \quad \text{and} \quad \mathcal{B}_{2,-} = \mathcal{B}_2 \setminus \mathcal{B}_{2,+}.$$

Since the balls in $\mathcal{B}_2$ are disjoint, we have

$$E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_2\bigr)\bigr) = \bigl(E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,+}\bigr)\bigr)\bigr) + \bigl(E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,-}\bigr)\bigr)\bigr).$$

Now we have two cases.

Case 1 Suppose

$$E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,+}\bigr)\bigr) \geq E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,-}\bigr)\bigr);$$

then

$$E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,+}\bigr)\bigr) \geq \tfrac{1}{2}\bigl(E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_2\bigr)\bigr)\bigr).$$

Since the balls in $\frac{1}{2}\mathcal{B}_{2,+}$ are contained in balls in $\mathcal{B}_1$ and harmonic replacements minimize energy, we get

$$E(H(u,\mathcal{B}_1,\mathcal{B}_2)) \leq E(H(u,\mathcal{B}_1)) \leq E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,+}\bigr)\bigr),$$

so that

$$\begin{aligned} E(u) - E(H(u,\mathcal{B}_1,\mathcal{B}_2)) &\geq E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,+}\bigr)\bigr) \\ &\geq \tfrac{1}{2}\bigl(E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_2\bigr)\bigr)\bigr) \\ &\geq k\bigl(E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_2\bigr)\bigr)\bigr)^2 \end{aligned} \tag{43}$$

if $k \leq 1/(2\varepsilon_0)$.

Case 2 Suppose

$$E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,+}\bigr)\bigr) \leq E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,-}\bigr)\bigr);$$

then

$$E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_{2,-}\bigr)\bigr) \geq \tfrac{1}{2}\bigl(E(u) - E\bigl(H\bigl(u, \tfrac{1}{2}\mathcal{B}_2\bigr)\bigr)\bigr).$$

Let $\tau > 0$ be given by Lemma 6.5. We can assume that

$$9\int_D |\nabla u - \nabla H(u,\mathcal{B}_1)|^2 \leq \tau^2. \tag{44}$$

Otherwise Theorems 2.1 and 2.2 give (41) with $k=\tau^2/(2\varepsilon_0^2)$. In fact, if

$$9\int_D |\nabla u-\nabla H(u,\mathcal{B}_1)|^2>\tau^2,$$

then applying Theorems 2.1 and 2.2 to each classical ball or boundary ball in $\mathcal{B}_1$, we get

$$E(u)-E(H(u,\mathcal{B}_1,\mathcal{B}_2))\ge E(u)-E(H(u,\mathcal{B}_1))\ge \frac{1}{2}\int_D |\nabla u-\nabla H(u,\mathcal{B}_1)|^2>\frac{\tau^2}{18}.$$

Noting that $E(u)-E\big(H\big(u,\tfrac{1}{2}\mathcal{B}_2\big)\big)\le \tfrac{1}{3}\varepsilon_0$, we get the desired estimate.

Assuming (44), we want to show the following estimate for each $B_j^2\in\mathcal{B}_{2,-}$:

$$\begin{aligned}(45)\quad \int_{B_j^2}|\nabla H(u,\mathcal{B}_1)|^2-\int_{B_j^2}|\nabla H(u,\mathcal{B}_1,B_j^2)|^2 &\\ \ge \int_{\frac12 B_j^2}|\nabla u|^2-\int_{\frac12 B_j^2}\big|\nabla H\big(u,\tfrac12 B_j^2\big)\big|^2 &\\ -C\bigg(\int_{B_j^2}|\nabla u|^2+|\nabla H(u,\mathcal{B}_1)|^2\bigg)^{\frac12}\bigg(\int_{B_j^2}|\nabla u-\nabla H(u,\mathcal{B}_1)|^2\bigg)^{\frac12}&.\end{aligned}$$

If (45) is true, summing it over balls in $\mathcal{B}_{2,-}$ and using the Cauchy–Schwartz inequality for discrete sums $\big|\sum a_jb_j\big|\le\big(\sum a_j^2\big)^{\frac12}\big(\sum b_j^2\big)^{\frac12}$ and Theorems 2.1 and 2.2, we get

$$\begin{aligned}(46)\quad E(H(u,\mathcal{B}_1))-E(H(u,\mathcal{B}_1,\mathcal{B}_{2,-}))&\\ \ge E(u)-E\big(H\big(u,\tfrac12\mathcal{B}_{2,-}\big)\big)-C\varepsilon_0^{\frac12}\big(E(u)-E(H(u,\mathcal{B}_1))\big)^{\frac12}&\\ \ge E(u)-E\big(H\big(u,\tfrac12\mathcal{B}_{2,-}\big)\big)-C\varepsilon_0^{\frac12}\big(E(u)-E(H(u,\mathcal{B}_1,\mathcal{B}_{2,-}))\big)^{\frac12}&.\end{aligned}$$

Then noting that

$$\begin{aligned}E(H(u,\mathcal{B}_1))-E(H(u,\mathcal{B}_1,\mathcal{B}_{2,-}))&\le E(u)-E(H(u,\mathcal{B}_1,\mathcal{B}_{2,-}))\\ &\le E(u)-E(H(u,\mathcal{B}_1,\mathcal{B}_2))\le\tfrac23\varepsilon_0<\varepsilon_0,\end{aligned}$$

we get the desired estimate (41).

So we only need to prove (45) to conclude Case 2. If B_j^2 is a classical ball, this is just the result in [6]. Now let us consider the case when B_j^2 is the boundary ball D_R^+ of radius R centered at 0 in the upper half-plane $\mathbb{H}^2$. Set $u_1=H(u,\mathcal{B}_1)$. By the coarea formula, there exists some $r \in \left[\frac{3}{4}R, R\right]$ with

$$\int_{\partial_r^A} |\nabla u_1 - \nabla u|^2 \le \frac{9}{R} \int_{\frac{3}{4}R}^{R} \left(\int_{\partial_s^A} |\nabla u_1 - \nabla u|^2 \right) ds \le \frac{9}{r} \int_{D_R^+} |\nabla u_1 - \nabla u|^2,$$

$$\begin{aligned} \int_{\partial_r^A} (|\nabla u_1|^2 + |\nabla u|^2) &\le \frac{9}{R} \int_{\frac{3}{4}R}^{R} \left(\int_{\partial_s^A} |\nabla u_1|^2 + |\nabla u|^2 \right) ds \\ &\le \frac{9}{r} \int_{D_R^+} (|\nabla u_1|^2 + |\nabla u|^2) \le \frac{24\varepsilon_0}{R}. \end{aligned}$$

The second estimate indicates that the length of the image of $u|_{\partial_r^A}$ is bounded by a universal constant times ε_0. So if ε_0 is smaller than a constant multiple of κ (where κ is the radius for Fermi coordinates of Γ; see Lemma 6.5), we will get $|u(x)-u(0)| \le \frac{1}{2}\kappa$ for all $x \in \partial_r^A$. Also note that $B_R \in \mathcal{B}_{2,-}$, by definition $\frac{1}{2}B_R$ does not completely lie in $\mathcal{B}_1$, so ∂_r^A does not lie in $\mathcal{B}_1$ for all $r \in \left[\frac{1}{2}R, R\right]$. Hence u_1 and u must agree at one point on ∂_r^A. Then we can apply Lemma 6.5 to get some $\rho \in \left(0, \frac{1}{2}r\right]$ and a map $w\colon D_r^+ \setminus D_{r-\rho}^+ \to \mathcal{N}$, with $w(r,\theta) = u_1(r,\theta)$ and $w(r-\rho,\theta) = u(r,\theta)$, such that

$$\int_{D_r^+ \setminus D_{r-\rho}^+} |\nabla w|^2 \le C \left(\int_{D_r^+} |\nabla u|^2 + |\nabla u_1|^2 \right)^{\frac{1}{2}} \left(\int_{D_r^+} |\nabla u - \nabla u_1|^2 \right)^{\frac{1}{2}}.$$

The map $x \to H(u, D_r^+)(rx/(r-\rho))$ maps $D_{r-\rho}^+$ to $\mathcal{N}$ and agrees with w on $\partial_{r-\rho}^A$. So we get a map from D_R^+ to $\mathcal{N}$ which is equal to $H(u,\mathcal{B}_1)$ on $D_R^+ \setminus D_r^+$, equal to w on $D_r^+ \setminus D_{r-\rho}^+$ and equal to $H(u, D_r^+)(r\cdot/(r-\rho))$ on $D_{r-\rho}^+$. This new map gives an upper bound for the energy of $H(u_1, D_R^+)$:

$$\int_{D_R^+} |\nabla H(u_1, D_R^+)|^2 \le \int_{D_R^+ \setminus D_r^+} |\nabla u_1|^2 + \int_{D_r^+ \setminus D_{r-\rho}^+} |\nabla w|^2 + \int_{D_r^+} |\nabla H(u, D_r^+)|^2.$$

Using the previous estimate and the fact $\left||\nabla u_1|^2 - |\nabla u|^2\right| \le (|\nabla u_1| + |\nabla u|)|\nabla(u-u_1)|$, we get

$$\begin{aligned} &\int_{D_R^+} |\nabla u_1|^2 - \int_{D_R^+} |\nabla H(u_1, D_R^+)|^2 \\ &\ge \int_{D_r^+} |\nabla u_1|^2 - \int_{D_r^+} |\nabla H(u, D_r^+)|^2 - \int_{D_r^+ \setminus D_{r-\rho}^+} |\nabla w|^2 \\ &\ge \int_{D_r^+} |\nabla u|^2 - \int_{D_r^+} |\nabla H(u, D_r^+)|^2 - C \left(\int_{D_R^+} |\nabla u|^2 + |\nabla u_1|^2 \right)^{\frac{1}{2}} \left(\int_{D_R^+} |\nabla(u-u_1)|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Since $r > \frac{1}{2}R$, we have

$$\begin{aligned}\int_{D^+_{R/2}} |\nabla u|^2 &= \int_{D^+_r} |\nabla u|^2 - \int_{D^+_r \setminus D^+_{R/2}} |\nabla u|^2 \\ &= \int_{D^+_r} |\nabla u|^2 - \int_{D^+_r} |\nabla H(u, D^+_{R/2})|^2 + \int_{D^+_{R/2}} |\nabla H(u, D^+_{R/2})|^2 \\ &\leq \int_{D^+_r} |\nabla u|^2 - \int_{D^+_r} |\nabla H(u, D^+_r)|^2 + \int_{D^+_{R/2}} |\nabla H(u, D^+_{R/2})|^2.\end{aligned}$$

Combining this with the previous estimate we get (45). Hence we can conclude the proof of (41).

Now one can prove (42) by the same argument as in [6], based on the proof of (41) above. This completes the proof. □

6.3 Tightening process

Now we have developed the necessary tools to prove Theorem 4.3. In fact, the proof is virtually the same as the proof of [6, Theorem 2.1]. All the main ingredients in the free boundary setting needed in the proof have been obtained in previous sections. We give the proof here for completeness and the convenience of readers.

Again, let ε_0 be the minimum of the ε_0 in Theorems 2.1 and 2.2. Given a sweep-out $\sigma \in \Omega$ and $0 < \varepsilon \leq \varepsilon_0$, we define the maximal improvement for free boundary harmonic replacement on families of generalized closed balls with energy at most ε by

$$e_{\sigma,\varepsilon}(t) = \sup_{\mathcal{B}} \bigl\{E(\sigma(\cdot,t)) - E\bigl(H\bigl(\sigma(\cdot,t), \tfrac{1}{2}\mathcal{B}\bigr)\bigr)\bigr\}.$$

Here the supremum is taken over all collections $\mathcal{B}$ of disjoint closed generalized balls in D where the total energy of $\sigma(\cdot,t)$ on $\mathcal{B}$ is at most ε. Note that $e_{\sigma,\varepsilon}(t)$ is positive if $\sigma(\cdot,t)$ is not harmonic.

We first show that the maximal improvement of a given slice (which is not harmonic) can control the maximal improvement of any nearby slices.

Lemma 6.8 *Given $t \in (0,1)$, if $\sigma(\cdot,t)$ is not harmonic and $0 < \varepsilon \leq \varepsilon_0$, then there is an open interval I^t containing t such that $e_{\sigma,\varepsilon/2}(s) \leq 2e_{\sigma,\varepsilon}(t)$ for all s in the double interval $2I^t$.*

Proof By Theorem 6.1 and [6, Corollary 3.4], there exists $\delta_1 > 0$ depending on t such that if $\|\sigma(\cdot,t) - \sigma(\cdot,s)\|_{C^0 \cap W^{1,2}} < \delta_1$ and $\mathcal{B}$ is a finite collection of disjoint generalized closed balls in D where both $\sigma(\cdot,t)$ and $\sigma(\cdot,s)$ have energy at most ε_0, then

$$\big|E\big(H\big(\sigma(\cdot,s), \tfrac{1}{2}\mathcal{B}\big)\big) - E\big(H\big(\sigma(\cdot,t), \tfrac{1}{2}\mathcal{B}\big)\big)\big| \le \tfrac{1}{2} e_{\sigma,\varepsilon}(t). \tag{47}$$

Since $t \to \sigma(\cdot,t)$ is continuous as a map to $C^0 \cap W^{1,2}$, we can choose I^t so that for all $s \in 2I^t$ we have $\|\sigma(\cdot,t) - \sigma(\cdot,s)\|_{C^0 \cap W^{1,2}} < \delta_1$ and

$$\int_D \big||\nabla\sigma(\cdot,t)|^2 - |\nabla\sigma(\cdot,s)|^2\big| \le \min\big\{\tfrac{\varepsilon}{2}, \tfrac{1}{2} e_{\sigma,\varepsilon}(t)\big\}. \tag{48}$$

Now suppose that $s \in 2I^t$ and the energy of $\sigma(\cdot,s)$ is at most $\frac{\varepsilon}{2}$ on a collection $\mathcal{B}$. It follows from (48) that the energy of $\sigma(\cdot,t)$ on this $\mathcal{B}$ is at most ε. Combining (47) with (48) we get

$$\big|E(\sigma(\cdot,s)) - E\big(H\big(\sigma(\cdot,s), \tfrac{1}{2}\mathcal{B}\big)\big) - E(\sigma(\cdot,t)) + E\big(H\big(\sigma(\cdot,t), \tfrac{1}{2}\mathcal{B}\big)\big)\big| \le e_{\sigma,\varepsilon}(t).$$

Since this estimate applies to any such $\mathcal{B}$, and noting that $e_{\sigma,\varepsilon}$ is monotone nondecreasing in ε, we get that $e_{\sigma,\varepsilon/2}(s) \le 2e_{\sigma,\varepsilon}(t)$. □

Theorem 4.3 indicates that our tightening process should effectively decrease the energy of those nonharmonic slices with large energy, ie when $E \ge \frac{1}{2}W$. The next lemma shows that we can find a harmonic replacement to decrease the energy of those slices by a certain amount.

Lemma 6.9 *If $W > 0$ and $\widetilde{\gamma} \in \Omega$ has no harmonic slices other than $\widetilde{\gamma}(\cdot,0)$ and $\widetilde{\gamma}(\cdot,1)$, then we get an integer m depending on $\widetilde{\gamma}$, along with m collections $\mathcal{B}_1,\dots,\mathcal{B}_m$ of generalized closed balls in D where the balls in each collection $\mathcal{B}_j$ are pairwise disjoint, and m continuous functions $r_1,\dots,r_m\colon [0,1] \to [0,1]$ such that, for each t:*

(i) *At most two $r_j(t)$ are positive and $\sum_{B\in\mathcal{B}_j} \int_{r_j(t)B} |\nabla\widetilde{\gamma}(\cdot,t)|^2 < \frac{1}{3}\varepsilon_0$ for each j.*

(ii) *If $E(\widetilde{\gamma}(\cdot,t)) \ge \frac{1}{2}W$, then there exists $j(t)$ such that the harmonic replacement on $\big(\frac{1}{2}r_{j(t)}\big)\mathcal{B}_{j(t)}$ decreases energy by at least $\frac{1}{8}e_{\widetilde{\gamma},\varepsilon_0/8}(t)$.*

Proof The energy of the slices is continuous in t, so the set $I = \big\{t : E(\widetilde{\gamma}(\cdot,t)) \ge \frac{1}{2}W\big\}$ is compact. For each $t \in I$, choose a finite collection $\mathcal{B}^t$ of disjoint closed balls in D with $\int_{\bigcup_{\mathcal{B}^t}} |\nabla\widetilde{\gamma}(\cdot,t)|^2 \le \frac{1}{4}\varepsilon_0$ so that

$$E(\widetilde{\gamma}(\cdot,t)) - E\big(H\big(\widetilde{\gamma}(\cdot,t), \tfrac{1}{2}\mathcal{B}^t\big)\big) \ge \tfrac{1}{2} e_{\widetilde{\gamma},\varepsilon_0/4}(t) > 0. \tag{49}$$

Note that $0, 1 \notin I$. Lemma 6.8 gives an open interval I^t containing t such that for all $s \in 2I^t$,

$$e_{\tilde{\gamma},\varepsilon_0/8}(s) \leq 2e_{\tilde{\gamma},\varepsilon_0/4}(t). \tag{50}$$

Using the continuity of $\tilde{\gamma}(\cdot, s)$ in $C^0 \cap W^{1,2}$, we can shrink I^t so that $\tilde{\gamma}(\cdot, s)$ has energy at most $\frac{1}{3}\varepsilon_0$ in $\mathcal{B}^t$ for $s \in 2I^t$ and in addition, using Theorem 6.1 and [6, Corollary 3.4], we have

$$\begin{aligned} \big|E(\tilde{\gamma}(\cdot, s)) - E\big(H\big(\tilde{\gamma}(\cdot, s), \tfrac{1}{2}\mathcal{B}^t\big)\big) - E(\tilde{\gamma}(\cdot, t)) + E\big(H\big(\tilde{\gamma}(\cdot, t), \tfrac{1}{2}\mathcal{B}^t\big)\big)\big| \\ \leq \tfrac{1}{4} e_{\tilde{\gamma},\varepsilon_0/4}(t). \end{aligned} \tag{51}$$

Since I is compact, we can cover I by finitely many of the I^t, say $I^{t_1}, \dots, I^{t_m}$. By discarding some of them, we can arrange that each t lies in at least one $\overline{I^{t_j}}$ and at most two consecutive $\overline{I^{t_j}}$. That is, we get a family of intervals I^{t_j} such that I^{t_j} only intersects $I^{t_{j-1}}$ and $I^{t_{j+1}}$, and $I^{t_{j-1}}$ and $I^{t_{j+1}}$ do not intersect each other. Now for each $j = 1, \dots, m$, we choose a continuous function $r_j(t)\colon [0, 1] \to [0, 1]$ with the following properties: $r_j(t) = 1$ on $\overline{I^{t_j}}$ and $r_j(t) = 0$ for $t \notin 2I^{t_j} \cap (I^{t_{j-1}} \cup I^{t_j} \cup I^{t_{j+1}})$; $r_j(t) = 0$ on the intervals that do not intersect $\overline{I^{t_j}}$.

Property (1) follows directly from the construction of the $r_j(t)$, and property (2) follows from combining (49), (50) and (51). □

Now using this lemma we can prove Theorem 4.3.

Proof of Theorem 4.3 Let $\mathcal{B}_1, \dots, \mathcal{B}_m$ and $r_1, \dots, r_m$ be given by Lemma 6.9. We will use an m–step replacement process to define γ.

We first set $\gamma^0 = \tilde{\gamma}$. Then for $j = 1, 2, \dots, m$ define $\gamma^j(\cdot, t) = H(\gamma^{j-1}(\cdot, t), r_j(t)\mathcal{B}_j)$. Finally we set $\gamma = \gamma^m$.

We first claim that this is a well-defined process and γ is again in $\Omega_{\tilde{\gamma}}$. In fact, property (1) in Lemma 6.9 implies that each energy-minimizing map replaces a map with energy at most $\frac{2}{3}\varepsilon_0 < \varepsilon_0$. Moreover, Theorem 6.1 and [6, Corollary 3.4] imply that the replacement depends continuously on the boundary values, which are themselves continuous in t. Finally it is clear that γ is homotopic to $\tilde{\gamma}$ since continuously shrinking the disjoint closed balls on which we make harmonic replacement gives an explicit homotopy.

Now we show that this γ satisfies the requirements of Theorem 4.3. Suppose $t \in [0, 1]$ is chosen with $E(\tilde{\gamma}(\cdot, t)) \geq \frac{1}{2}W$, then property (2) of Lemma 6.9 implies that the harmonic replacement for $\tilde{\gamma}(\cdot, t)$ on $\left(\frac{1}{2}r_j(t)\right)\mathcal{B}_{j(t)}$ decreases the energy by at least $\frac{1}{8}e_{\tilde{\gamma},\varepsilon_0}(t)$. Thus using (41) of Lemma 6.7 we get

$$E(\tilde{\gamma}(\cdot, t)) - E(\gamma(\cdot, t)) \geq k\left(\tfrac{1}{8}e_{\tilde{\gamma},\varepsilon_0/8}(t)\right)^2 \quad \text{for some } k > 0. \tag{52}$$

Suppose that $\mathcal{B}$ is a finite collection of disjoint generalized closed balls in D such that the energy of $\gamma(\cdot, t)$ on $\mathcal{B}$ is at most $\frac{1}{12}\varepsilon_0$. We can assume $\gamma^k(\cdot, t)$ has energy at most $\frac{1}{8}\varepsilon_0$ on $\mathcal{B}$ for every k, since otherwise using Theorems 2.1 and 2.2 we could chose Ψ to be a linear function $\Psi(x) = Cx$ for some $C > 0$ sufficiently large. Now we apply (42) in Lemma 6.7 twice with $\mu = \frac{1}{8}$ and then $\mu = \frac{1}{4}$ to get

$$\begin{aligned} E(\gamma(\cdot, t)) &- E\left(H\left(\gamma(\cdot, t), \tfrac{1}{8}\mathcal{B}\right)\right) \\ &\leq E(\tilde{\gamma}(\cdot, t)) - E\left(H\left(\tilde{\gamma}(\cdot, t), \tfrac{1}{2}\mathcal{B}\right)\right) + \tfrac{2}{k}\left(E(\tilde{\gamma}(\cdot, t)) - E(\gamma(\cdot, t))\right)^{\frac{1}{2}} \\ &\leq e_{\tilde{\gamma},\varepsilon_0/8}(t) + \tfrac{2}{k}\left(E(\tilde{\gamma}(\cdot, t)) - E(\gamma(\cdot, t))\right)^{\frac{1}{2}}. \end{aligned} \tag{53}$$

Combining (52) and (53) with Theorems 2.1 and 2.2, we complete the proof by choosing $\Psi(x) = Cx + Cx^{\frac{1}{2}}$ for some sufficiently large $C > 0$ and $\varepsilon_1 = \frac{1}{12}\varepsilon_0$. □

7 Compactness of maximal slices

This section is devoted to the proof of Theorem 4.4. In particular, we will prove that any sequence of maps (as slices of approximating sweepouts) whose energy converges to the width will converge to a bubble tree of free boundary harmonic disks and harmonic spheres. Similar bubble tree convergence was first studied by Fraser [15] for α–harmonic disks with free boundary (where the notion of α–maps was introduced by Sacks–Uhlenbeck [45]). In our proof, we adopt the schemes in both [6, Appendix B] and [15]. We generalize the notion of almost harmonic maps with their asymptotic analysis and compactness in [6, Appendix B] to our free boundary setting.

7.1 Compactness of free boundary almost harmonic maps

We first introduce our notion of almost harmonic maps with free boundary. Note that generalized balls were defined in Definition 4.2.

Definition 7.1 We say a sequence of maps $u^j\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ is ε_1*–almost harmonic in the free boundary sense* if:

(B_0) For any generalized ball $B \subset D$ with $\int_B |\nabla u^j|^2 < \varepsilon_1$, there is a harmonic replacement $v\colon \frac{1}{8}B \to \mathcal{N}$ of u^j (with possible free boundary along $v\big(\partial \frac{1}{8}B \cap \partial D\big) \subset \Gamma$ if it is not empty) which satisfies the bound

$$\int_{\frac{1}{8}B} |\nabla u^j - \nabla v|^2 \le \frac{1}{j}.$$

Next we have the following preliminary compactness result for a sequence of almost harmonic maps in the free boundary sense with finite energy. In particular, any such sequence converges to a harmonic map with free boundary in $W^{1,2}$ locally away from only finitely many points in D. Let ε_{SU} and ε_F be as in Section 4.

Theorem 7.2 *Let $\epsilon_1 > 0$ be such that $\epsilon_1 < \min\{\varepsilon_{SU}, \varepsilon_F\}$. Let $u^j\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ be a sequence of $W^{1,2}$–maps satisfying property (B_0) in Definition 7.1 and having $E(u^j) \le E_0 < \infty$. Then there exist a finite collection of points $\{x_1, \dots, x_k\} \subset \bar{D}$, a subsequence of maps* (still denoted by u^j), *and a harmonic map $u\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ with free boundary, such that $u^j \to u$ weakly in $W^{1,2}(D)$ and for any compact subset $K \subset \bar{D} \setminus \{x_1, \dots, x_k\}$, $u^j \to u$ strongly in $W^{1,2}(K)$. Furthermore, the measures $|\nabla u^j|^2\,dx$ converge to a measure ν on $\bar{D}$ with $\nu(\{x_i\}) \ge \varepsilon_1$ for all $1 \le i \le k$ and $\nu(\bar{D}) \le E_0$.*

Proof After passing to a subsequence we can assume that the u^j converge weakly in $W^{1,2}(D)$ to a $W^{1,2}$–map $u\colon D \to \mathcal{N}$, and the measures $|\nabla u^j|^2\,dx$ converge weakly to a limiting measure ν on $\bar{D}$ with $\nu(\bar{D}) \le E_0$.

So there are at most E_0/ε_1 points $x_1, \dots, x_k \in \bar{D}$, with $\lim_{r\to 0} \nu(B_r(x_j)) \ge \varepsilon_1$.

Next we show that away from these points the convergence is strong in $W^{1,2}$ and u is a harmonic map with free boundary in Γ. Given any $x \in \bar{D} \setminus \{x_1, \dots, x_k\}$, by definition there exist a generalized ball B_x and an integer J_x such that $\int_{B_x} |\nabla u^j|^2 < \varepsilon_1$ for $j > J_x$. If x is an interior point we can choose B_x to be a classical ball which is contained in D, otherwise B_x will be some boundary ball. By condition (B_0) we get a harmonic replacement $v_x^j\colon \frac{1}{8}B_x \to \mathcal{N}$ such that

$$\int_{\frac{1}{8}B_x} |\nabla u^j - \nabla v_x^j| \le \frac{1}{j}.$$

Note that the energy of v_x^j on the ball $\frac{1}{8}B_x(x)$ is less than ε_1, so by the ε–regularity of Qing [39, Lemma 4] and Theorem 2.5, we get a uniform C^k–bound for v_x^j in $\frac{1}{9}B_x$ for $k \ge 1$. Hence a subsequence v_x^j converges strongly in $W^{1,2}\big(\frac{1}{9}B_x\big)$ to a harmonic map $v_x\colon \frac{1}{9}B_x \to \mathcal{N}$ with possible free boundary along $v_x\big(\partial\frac{1}{9}B_x \cap \partial D\big) \subset \Gamma$. By the triangle inequality we get

$$\int_{\frac{1}{9}B_x} |\nabla u^j - \nabla v_x|^2 \le 2\int_{\frac{1}{9}B_x} |\nabla u^j - \nabla v_x^j|^2 + 2\int_{\frac{1}{9}B_x} |\nabla v_x^j - \nabla v_x|^2 \to 0$$

as $j \to \infty$.

We can also derive the L^2–convergence of u^j to v_x on $\frac{1}{9}B_x$ by the inequality

$$\int_{\frac{1}{9}B_x} |u^j - v_x|^2 \le 2\int_{\frac{1}{9}B_x} |u^j - v_x^j|^2 + 2\int_{\frac{1}{9}B_x} |v_x^j - v_x|^2$$

and the Poincaré inequality (when B_x is a classical ball of D) or its variant Lemma A.2 (when B_x is a boundary ball of D).

We have proved that the sequence u^j converges to v_x strongly in $W^{1,2}\big(\frac{1}{9}B_x\big)$, and hence $u = v_x$ in $\frac{1}{9}B_x$. Therefore we conclude that u is a free boundary harmonic map on $D \setminus \{x_1, \ldots, x_k\}$. Furthermore, for K relatively compact in $\bar{D} \setminus \{x_1, \ldots, x_k\}$, the $W^{1,2}(K)$–convergence of u^j to u follows from a standard covering argument.

Finally, since u has finite energy, we can apply the removable singularity theorem by Sacks–Uhlenbeck [45, Theorem 3.6] or Fraser [15, Theorem 1.10] at each x_i for interior points or free boundary points respectively, so that u extends to a global harmonic map on the whole of D with free boundary along $u(\partial D) \subset \Gamma$. □

7.2 Harmonic maps on half-cylinders

In this and the following subsections, we will generalize the analysis of harmonic maps defined on cylinders in [6, Appendix B] to harmonic maps with free boundary defined on half-cylinders. The analysis of harmonic maps and almost harmonic maps on cylinders in [6] is essential to the proof of the so-called *energy identity*. More precisely, in the blow-up process (see Section 7.4), the energy could be lost (so that the energy identity fails) only when some energy escapes from the "necks" (modeled by cylinders or half-cylinders), and this is the case we want to rule out in our scenario. In the free boundary setting, not only spherical bubbles but also disk bubbles may appear during the blow-up process. So we need to carefully analyze the maps defined on the "necks" between the spheres and disks, ie cylinders and half-cylinders.

Let us set up some notation. Let $\mathcal{C}_{a,b}$ denote the flat half-cylinder $[a,b] \times [0,\pi]$, where $[0,\pi]$ can be viewed as half of a circle S^1. We will use (t,θ) as parameters on $\mathcal{C}_{a,b}$. Note that $\mathcal{C}_{a,b}$ is conformally equivalent to the half-annulus in plane: $[e^a, e^b] \times [0,\pi]$

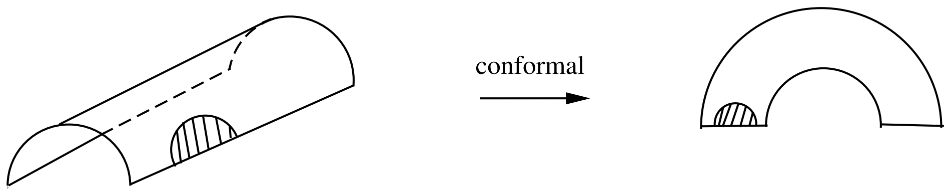

Figure 2

under polar coordinates, see Figure 2. We will also abuse the notation $\partial^C \mathcal{C}$ to denote the part of the boundary $[a,b]\times\{0,\pi\}$, and later this will be the free boundary part of our maps and we will only focus on the boundary behavior of the maps along this part. Moreover, when we say a map $u\colon \mathcal{C}\to\mathcal{N}$ is a harmonic map with free boundary on Γ, we will always assume $u|_{\partial^C\mathcal{C}}$ is the free boundary with $u(\partial^C\mathcal{C})\subset\Gamma$.

Theorem 7.3 *Given $\delta>0$, there exist $\varepsilon_2>0$ and $l>1$ depending on δ such that if u is a nonconstant C^3–harmonic map from the flat half-cylinder $\mathcal{C}_{-3l,3l}$ to $\mathcal{N}$ with free boundary along Γ, and the energy $E(u)\le\varepsilon_2$, then*

$$\int_{\mathcal{C}_{-l,l}}|u_\theta|^2<\delta\int_{\mathcal{C}_{-2l,2l}}|\nabla u|^2. \tag{54}$$

Roughly speaking, this theorem implies that for nonconstant free boundary harmonic maps with small energy, the θ–energy (on a sub-half-cylinder) is far smaller than the total energy. Then by the Cauchy–Schwartz inequality this implies that the area of the image of u is strictly less than the energy of u.

In order to prove Theorem 7.3, we follow the idea of Colding–Minicozzi to prove a differential inequality for various energies of a free boundary harmonic map. In particular we need some free boundary versions of the lemmas in [6, Appendix B].

Lemma 7.4 *Given a C^3–free boundary harmonic map u from $\mathcal{C}_{-3l,3l}$ to $\mathcal{N}\subset\mathbb{R}^N$ with $E(u)\le\varepsilon_2$ for some $\varepsilon_2>0$ sufficiently small, we have*

$$\partial_t^2\int_t|u_\theta|^2\ge\frac{1}{C}\int_t|u_\theta|^2-C\int_t|\nabla u|^4-C\sqrt{\varepsilon_2}\,\frac{1}{\lambda}\int_{t-\lambda}^{t+\lambda}\int_s|\nabla u|^2\,ds \tag{55}$$

for all $t\in[-2l,2l]$ and $\lambda\in(0,1)$. Here $\int_t(\cdot)$ denotes the integral over the t–slice $\{(t,\theta)\in\mathcal{C}_{-3l,3l}:\theta\in[0,\pi]\}$, and $C>0$ is some universal constant depending only on $\mathcal{N}$ and Γ.

Proof We can assume $\varepsilon_2 \le \varepsilon_0$, as in Theorem 2.5. Differentiating $\int_t |u_\theta|^2$ twice gives

$$
\begin{aligned}
(56)\qquad \frac{1}{2}\partial_t^2 \int_t |u_\theta|^2 &= \int_t |u_{t\theta}|^2 + \int_t \langle u_\theta, u_{tt\theta}\rangle \\
&= \int_t |u_{t\theta}|^2 - \int_t \langle u_{\theta\theta}, u_{tt}\rangle + \langle u_\theta, u_{tt}\rangle|_0^\pi \\
&= \int_t |u_{t\theta}|^2 - \int_t \langle u_{\theta\theta}, (\Delta u - u_{\theta\theta})\rangle + \langle u_\theta, A^\Gamma(u_t,u_t)\rangle|_0^\pi \\
&\ge \int_t |u_{t\theta}|^2 + \int_t |u_{\theta\theta}|^2 - \sup_{\mathcal{N}}|A^{\mathcal{N}}| \int_t |u_{\theta\theta}|\,|\nabla u|^2 \\
&\qquad\qquad - 2\sup_\Gamma |A^\Gamma|\, C\sqrt{\varepsilon_2}\,\frac{1}{\lambda}\int_{t-\lambda}^{t+\lambda}\int_s |\nabla u|^2\, ds.
\end{aligned}
$$

Here A^Γ and $A^{\mathcal{N}}$ are the second fundamental forms of $\Gamma \hookrightarrow \mathcal{N}$ and $\mathcal{N} \hookrightarrow \mathbb{R}^N$. In the second equality we used integration by parts; in the third inequality, since u is a free boundary map, on the boundary (where $\theta = 0$ or $\theta = \pi$) u_θ is a tangent vector of $\mathcal{N}$ and also perpendicular to Γ, and $u_{tt} = \nabla^{\mathcal{N}}_{u_t} u_t$ when projected to the tangent space $T_u\mathcal{N}$, where u_t is tangent to Γ; in the last inequality we used $|\Delta u| \le \sup_{\mathcal{N}}|A^{\mathcal{N}}||\nabla u|^2$ by the harmonic map (3), and also the gradient estimate (Theorem 2.5) applied to half-disks of radius λ centered at $(t,0)$ and (t,π). By the Cauchy–Schwartz inequality we get

$$
\frac{1}{2}\partial_t^2 \int_t |u_\theta|^2 \ge \int_t |u_{t\theta}|^2 + \frac{3}{4}\int_t |u_{\theta\theta}|^2 - C\int_t |\nabla u|^4 - C\sqrt{\varepsilon_2}\,\frac{1}{\lambda}\int_{t-\lambda}^{t+\lambda}\int_s |\nabla u|^2\, ds
$$

for some universal constant $C > 0$ depending only on $\mathcal{N}$ and Γ.

Next we claim: for u satisfying the assumptions of the theorem,

$$
(57)\qquad \int_t |u_\theta|^2 \le C\int_t |u_{\theta\theta}|^2 \quad \text{for } t \in [-2l, 2l],
$$

where $C > 0$ is a constant depending only on $\mathcal{N}$ and Γ. Note that once we have this inequality we get the desired inequality (55) in the lemma.

To prove (57), we reflect the map $u(t,\cdot)\colon [0,\pi] \to \mathbb{R}^N$ across Γ to obtain a map defined on the circle $\hat{u}(t,\cdot)\colon S^1 \to \mathbb{R}^N$. In particular, let P_Γ denote the nearest-point projection map from a tubular neighborhood of Γ in $\mathbb{R}^N$ to Γ. When the tubular neighborhood is chosen small enough, we can assume that

$$
|DP_\Gamma| \le 1 \quad\text{and}\quad |D^2P_\Gamma| \le C
$$

for some universal constant $C > 0$. By the gradient estimate Theorem 2.5, we can assume that the image of $u(\mathcal{C}_{-2l,2l})$ lies in this tubular neighborhood when $\varepsilon_2 > 0$ is chosen small enough. Now define the map $\widehat{u}(t,\cdot)\colon S^1\to\mathbb{R}^N$ by

$$\widehat{u}(t,\theta)=\begin{cases}u(t,\theta) & \text{when } \theta\in[0,\pi],\\ 2P_\Gamma(u(t,-\theta))-u(t,-\theta) & \text{when } \theta\in[-\pi,0].\end{cases}$$

By the free boundary assumption, we know that $\widehat{u}$ is $C^{1,1}$ on S^1. As $\int_{S^1}\widehat{u}_\theta=0$, by the Wirtinger inequality we get $\int_{S^1}|\widehat{u}_\theta|^2\le\int_{S^1}|\widehat{u}_{\theta\theta}|^2$, and hence we can deduce that

$$\begin{aligned}\int_0^\pi|u_\theta|^2&\le\int_0^\pi|u_{\theta\theta}|^2+\int_0^\pi|(2P_\Gamma(u(t,\theta))-u(t,\theta))_{\theta\theta}|^2\\&\le\int_0^\pi|u_{\theta\theta}|^2+\int_0^\pi|2DP_\Gamma(u_{\theta\theta})-u_{\theta\theta}+2D^2P_\Gamma(u_\theta,u_\theta)|^2\\&\le C\int_0^\pi|u_{\theta\theta}|^2+C\int_0^\pi|u_\theta|^4.\end{aligned}\tag{58}$$

Note that by the gradient estimates Theorem 2.5, $C\int_0^\pi|u_\theta|^4\le C^2\varepsilon_2\int_0^\pi|u_\theta|^2$. So the desired estimates (57) follow by taking $\varepsilon_2>0$ small enough, so that $C^2\varepsilon_2<\frac{1}{2}$. □

The next lemma is an ODE comparison lemma.

Lemma 7.5 *Suppose f is a nonnegative C^2 function on $[-2l,2l]\subset\mathbb{R}$ satisfying*

$$f''\ge\frac{1}{4C}f-a\tag{59}$$

for some constants $C,a>0$. If $\max_{[-l,l]}f\ge 8Ca$, then

$$\int_{-2l}^{2l}f\ge 4\sqrt{2C}a\sinh\Big(\frac{l}{2\sqrt{2C}}\Big).\tag{60}$$

Proof Let $\tilde{f}(t)=f(\sqrt{4C}t)$. Then we get a differential inequality for $\tilde{f}$,

$$\tilde{f}''\ge\tilde{f}-4Ca,$$

where $\tilde{f}$ is defined on $[-l/\sqrt{C},l/\sqrt{C}]$, and

$$\max_{[-l/2\sqrt{C},l/2\sqrt{C}]}\tilde{f}\ge 8Ca.$$

Then applying [6, Lemma B.4] gives

$$\int_{-l/\sqrt{C}}^{l/\sqrt{C}}\tilde{f}\ge 8\sqrt{2}Ca\sinh\Big(\frac{l}{2\sqrt{2C}}\Big),$$

which implies

$$\int_{-2l}^{2l}f\ge 4\sqrt{2C}a\sinh\Big(\frac{l}{2\sqrt{2C}}\Big).$$ □

Now we can prove the main theorem of this subsection.

Proof of Theorem 7.3 First we analyze $\int_t(|u_t|^2-|u_\theta|^2)$. Differentiating it and applying integration by parts gives

$$(61)\quad \frac{1}{2}\partial_t\int_t(|u_t|^2-|u_\theta|^2)=\int_t(\langle u_t,u_{tt}\rangle-\langle u_\theta,u_{\theta t}\rangle)=\int_t\langle u_t,u_{tt}+u_{\theta\theta}\rangle=0.$$

Here again we use the fact that u_θ and u_t are perpendicular to each other on the free boundary, and $u_{tt}+u_{\theta\theta}=\Delta u$, which is normal to $\mathcal{N}$ and hence perpendicular to u_t. Thus $\int_t(|u_t|^2-|u_\theta|^2)$ is a constant, and

$$(62)\quad \int_t(|u_t|^2-|u_\theta|^2)=\frac{1}{4l}\int_{\mathcal{C}_{-2l,2l}}(|u_t|^2-|u_\theta|^2)\le\frac{1}{4l}\int_{\mathcal{C}_{-2l,2l}}|\nabla u|^2.$$

Moreover, we get

$$(63)\quad \int_t|\nabla u|^2=2\int_t|u_\theta|^2+\int_t(|u_t|^2-|u_\theta|^2)\le 2\int_t|u_\theta|^2+\frac{1}{4l}\int_{\mathcal{C}_{-2l,2l}}|\nabla u|^2.$$

Let us choose ε_2 smaller than the values of ε_0 in [45, Lemma 3.4] and Theorem 2.5. Then the interior gradient estimates for harmonic maps (see [45, Lemma 3.4]) and for free boundary harmonic maps (Theorem 2.5) imply that

$$(64)\quad \sup_{\mathcal{C}_{-2l,2l}}|\nabla u|^2\le C\varepsilon_2.$$

Let $f(t)=\int_t|u_\theta|^2$. Then by Lemma 7.4 we get

$$\begin{aligned}(65)\quad f''(t)&\ge\frac{1}{C}f(t)-C\varepsilon_2\int_t|\nabla u|^2-C\sqrt{\varepsilon_2}\frac{1}{\lambda}\int_{t-\lambda}^{t+\lambda}\int_s|\nabla u|^2\,ds\\ &\ge\frac{1}{C}f(t)-2C\varepsilon_2f(t)-2C\sqrt{\varepsilon_2}\frac{1}{\lambda}\int_{t-\lambda}^{t+\lambda}f(t)\\ &\qquad-\frac{C\varepsilon_2}{4l}\int_{\mathcal{C}_{-2l,2l}}|\nabla u|^2-\frac{2C\sqrt{\varepsilon_2}}{4l}\int_{\mathcal{C}_{-2l,2l}}|\nabla u|^2\\ &\ge\frac{1}{2C}\Big(f(t)-\frac{1}{4\lambda}\int_{t-\lambda}^{t+\lambda}f\Big)-\frac{C\sqrt{\varepsilon_2}}{l}\int_{\mathcal{C}_{-2l,2l}}|\nabla u|^2,\end{aligned}$$

where $C>0$ depends only on $\mathcal{N}$ and Γ, and we can further assume that ε_2 is small enough that $C^2\varepsilon_2\le\frac{1}{4}$ and $2C^2\varepsilon_2^{1/4}\le 1$.

Letting $\lambda\to 0$, by continuity of f we get the differential inequality

$$(66)\quad f''(t)\ge\frac{1}{4C}f(t)-a,$$

where $a = (C\sqrt{\varepsilon_2}/l)\int_{\mathcal{C}_{-2l,2l}} |\nabla u|^2$. Then we apply Lemma 7.5 to get either

$$\max_{[-l,l]} f < 8\frac{C^2\sqrt{\varepsilon_2}}{l}\int_{\mathcal{C}_{-2l,2l}} |\nabla u|^2 \tag{67}$$

or

$$\int_{\mathcal{C}_{-2l,2l}} |u_\theta|^2 = \int_{-2l}^{2l} f(t)\,dt \geq 4\sqrt{2C}C\sqrt{\varepsilon_2}\frac{\sinh(l/2\sqrt{2C})}{l}\int_{\mathcal{C}_{-2l,2l}} |\nabla u|^2. \tag{68}$$

If we choose l large enough then the second inequality can not hold. Then we get

$$\int_{\mathcal{C}_{-l,l}} |u_\theta|^2 \leq 2l \max_{[-l,l]} f < 8C^2\sqrt{\varepsilon_2}\int_{\mathcal{C}_{-2l,2l}} |\nabla u|^2. \tag{69}$$

Then the inequality (54) holds if $8C^2\sqrt{\varepsilon_2} < \delta$. So we can choose ε_2 small and then choose l large to get the desired inequality. □

7.3 Almost harmonic maps on half-cylinders

The main results in this part generalize the results in the previous subsection to almost harmonic maps on half-cylinders.

Let us first fix some notation. Given a half-cylinder $\mathcal{C}_{r_1,r_2}$, we will view it as its conformally equivalent half-annulus $D^+_{e^{r_2}} \setminus D^+_{e^{r_1}} \subset D^+$. A generalized ball $B \subset D^+_{e^{r_2}} \setminus D^+_{e^{r_1}}$ is either a ball in the interior of the annulus $D^+_{e^{r_2}} \setminus D^+_{e^{r_1}}$, or is a half-ball centered along the chord boundary $\partial^C_{e^{r_2}} \setminus \partial^C_{e^{r_1}}$. When B is a half-ball, we write $\partial^C B = \partial B \cap \partial^C$. Note that this definition is the same as that in Definition 4.2.

Definition 7.6 Given $\mu > 0$ and a half-cylinder $\mathcal{C}_{r_1,r_2}$, we will say that a $W^{1,2}$–map $u\colon \mathcal{C}_{r_1,r_2} \to \mathcal{N}$ with $u(\partial^C\mathcal{C}) \subset \Gamma$ is *μ–almost harmonic with free boundary* if for any finite collection of disjoint generalized balls $\mathcal{B}$ in the conformally equivalent half-annulus $D^+_{e^{r_2}} \setminus D^+_{e^{r_1}}$, there is a free boundary harmonic replacement $v\colon \bigcup_{\mathcal{B}} \frac{1}{8}B \to \mathcal{N}$ with free boundary along $u\big(\bigcup_{\mathcal{B}} \frac{1}{8}\partial^C B\big) \subset \Gamma$ such that

$$\int_{\bigcup_{\mathcal{B}} \frac{1}{8}B} |\nabla u - \nabla v|^2 \leq \frac{\mu}{2}\int_{\mathcal{C}_{r_1,r_2}} |\nabla u|^2.$$

The first lemma of this subsection shows that for almost harmonic maps with free boundary the estimate in the previous subsection still holds.

Lemma 7.7 *Given $\delta > 0$ there exists $\mu > 0$ depending on δ, $\mathcal{N}$ and Γ such that if $u\colon \mathcal{C}_{-3l,3l} \to \mathcal{N}$ is a μ–almost harmonic map with free boundary and $E(u) \le \varepsilon_2$, where l is from Theorem 7.3, then*

(70)
$$\int_{\mathcal{C}_{-l,l}} |u_\theta|^2 \le \delta \int_{\mathcal{C}_{-3l,3l}} |\nabla u|^2.$$

Proof We will argue by contradiction. Suppose the lemma does not hold. Then there exists a sequence u^j of $\frac{1}{j}$–almost harmonic maps from $\mathcal{C}_{-3l,3l}$ to $\mathcal{N}$ with free boundary $u_j(\partial^C \mathcal{C}_{-3l,3l}) \subset \Gamma$, energy $E(u^j) \le \varepsilon_2$, and

(71)
$$\int_{\mathcal{C}_{-l,l}} |u^j_\theta|^2 > \delta \int_{\mathcal{C}_{-3l,3l}} |\nabla u^j|^2.$$

Now we have two cases depending on whether the energy of the limit is zero.

Case 1 Suppose $\limsup_{j\to\infty} E(u^j) > 0$. Then up to a subsequence, $\int_{\mathcal{C}_{-l,l}} |u^j_\theta|^2$ is uniformly bounded from below by (71). We will apply the compactness result (Theorem 7.2) to this sequence. In particular, we can use the same argument as Theorem 7.2 to find a subsequence that converges weakly to a free boundary harmonic map $u\colon (\mathcal{C}_{-3l,3l}, \partial^C \mathcal{C}_{-3l,3l}) \to (\mathcal{N}, \Gamma)$, and strongly in $W^{1,2}$ on any compact subset of $\mathcal{C}_{-3l,3l}$. Note that since $E(u^j) \le \varepsilon_2$, there will be no energy concentration points. The uniform lower bound of $\int_{\mathcal{C}_{-l,l}} |u^j_\theta|^2$ and the $W^{1,2}$–convergence on $\mathcal{C}_{-l,l}$ imply that u can not be a constant map. Finally by the lower semicontinuity of energy along $W^{1,2}$–weak convergence,

$$\int_{\mathcal{C}_{-l,l}} |u_\theta|^2 \ge \delta \int_{\mathcal{C}_{-3l,3l}} |\nabla u|^2,$$

which contradicts Theorem 7.3.

Case 2 Suppose $\lim_{j\to\infty} E(u^j) = 0$. We will use a blow-up argument. Let

$$v^j = \frac{u^j - u^j(0)}{E(u^j)^{1/2}}.$$

This is a sequence of maps from $\mathcal{C}_{-3l,3l}$ to $\mathcal{N}_j = (\mathcal{N} - u^j(0))/(E(u^j)^{1/2})$. Here $0 = (0,0)$ is a boundary point on $\mathcal{C}_{-3l,3l}$, so $u^j(0) \in \Gamma$; hence we can always see the free boundary $\Gamma_j = (\Gamma - u^j(0))/(E(u^j)^{1/2})$ in the blow-up process.

Note that $E(v^j) = 1$ and by (71), we have

$$\int_{\mathcal{C}_{-l,l}} |v^j_\theta|^2 > \delta > 0.$$

Furthermore, the sequence of the v^j are still $\frac{1}{j}$–almost harmonic because this property is invariant under dilation. So we can argue as before to get a subsequence that converges in $W^{1,2}$ on compact subsets of $\mathcal{C}_{-3l,3l}$ to a free boundary harmonic map $v\colon \mathcal{C}_{-3l,3l} \to \mathbb{R}^n \subset \mathbb{R}^N$ with free boundary $\Gamma = \mathbb{R}^k \subset \mathbb{R}^N$. As before, we get

$$\int_{\mathcal{C}_{-l,l}} |v_\theta|^2 \geq \delta,$$

which again contradicts Theorem 7.3. (Note that for free boundary harmonic maps into $(\mathbb{R}^n, \mathbb{R}^k)$ we do not need the assumption $E(v) \leq \varepsilon_2$; see also [6, Remark B.3].) □

With this lemma we can prove that the θ–energy of a free boundary almost harmonic map on a long half-cylinder would be far less than the total energy.

Theorem 7.8 *Given $\delta > 0$ there exists $\nu > 0$ depending on δ, $\mathcal{N}$ and Γ such that if m is any positive integer and u is ν–almost harmonic from $\mathcal{C}_{-(m+3)l,3l}$ to $\mathcal{N}$ with free boundary along $u(\partial^C \mathcal{C}_{-(m+3)l,3l}) \subset \Gamma$, and $E(u) \leq \varepsilon_2$, then*

$$\int_{\mathcal{C}_{-ml,0}} |u_\theta|^2 \leq 7\delta \int_{\mathcal{C}_{-(m+3)l,3l}} |\nabla u|^2. \tag{72}$$

Proof The proof follows by covering $\mathcal{C}_{-(m+3)l,3l}$ by sub-half-cylinders of length $6l$, together with Lemma 7.7. We refer to the proof of [6, Proposition B.19] for details. □

The following simple lemma will be useful in the next subsection.

Lemma 7.9 *Suppose $u\colon \mathcal{C}_{-(m+3)l,3l} \to \mathcal{N}$ is a map satisfying*

$$\int_{\mathcal{C}_{-ml,0}} |u_\theta|^2 \leq \tfrac{1}{9} E(u) \quad \text{and} \quad \int_{\mathcal{C}_{-(m+3)l,-ml} \cup \mathcal{C}_{0,3l}} |\nabla u|^2 \leq \tfrac{1}{9} E(u).$$

Then

$$\operatorname{Area}(u) \leq \tfrac{8}{9} E(u).$$

Proof First note that

$$\operatorname{Area}(u) = \int_{\mathcal{C}_{-(m+3)l,3l}} (|u_\theta|^2 |u_t|^2 - \langle u_\theta, u_t\rangle^2)^{1/2} \leq \int_{\mathcal{C}_{-(m+3)l,3l}} |u_\theta|\,|u_t|.$$

By $ab \le \frac{4}{3}a^2 + \frac{3}{16}b^2$, we get

$$\begin{aligned}\text{(73)}\qquad \int_{C_{-ml,0}} |u_\theta||u_t| &\le \frac{4}{3}\int_{C_{-ml,0}} |u_\theta|^2 + \frac{3}{16}\int_{C_{-ml,0}} |u_t|^2 \\ &\le \frac{1}{3}\int_{C_{-ml,0}} |u_\theta|^2 + \frac{3}{16}\int_{C_{-ml,0}} |u_t|^2 + \frac{1}{9}E(u) \\ &\le \tfrac{7}{9}E(u).\end{aligned}$$

Combining with the second assumption, we have

$$\text{(74)}\qquad \operatorname{Area}(u) \le \tfrac{7}{9}E(u) + \tfrac{1}{9}E(u) = E(u) - \tfrac{1}{9}E(u) \le \tfrac{8}{9}E(u). \qquad \square$$

7.4 Proof of Theorem 4.4

Now we are ready to prove Theorem 4.4. We will follow the same scheme as in [6, Section B.6]. One key point in our setting is that we may get two different kinds of bubbles. We may get spherical bubbles as in [6, Section B.6], as well as free boundary disk bubbles as [15]. The new techniques developed in previous sections will be essentially used to study these bubbles.

A boundary ball B of D is always the intersection of a classical ball $B_r(x)$ of $\mathbb{R}^2$ with D, ie $B = B_r(x) \cap D$ for some $x \in \partial D$ and $r > 0$. In the following proof we say r is the radius of B and sometime abuse notation to write $B = B_r(x)$.

We use $\tilde{\Pi}$ to denote a fixed conformal map that maps the upper half-plane $\overline{\mathbb{H}^2}$ to the unit disk $\bar{D}$ which maps $(0, 1, \infty)$ to three given distinct points on ∂D. We denote by p^+ the image of ∞ and by D^- the image of $D_1^+ \subset \overline{\mathbb{H}^2}$ under the map Π. For a given boundary ball $B_r(x)$, we define the *conformal dilation of* $B_r(x)$ to be the map $\Psi_{r,x}\colon \bar{D} \to \bar{D}$ such that $\Psi_{r,x} = \tilde{\Pi} \circ \Phi_{r,x} \circ \tilde{\Pi}^{-1}$, where $\Phi_{r,x}$ is the composition of a dilation of $\overline{\mathbb{H}^2}$ by the factor $\frac{1}{r}$ and a translation of $\overline{\mathbb{H}^2}$ by $-\tilde{\Pi}^{-1}(x)$. (Note $\tilde{\Pi}^{-1}(x)$ is a boundary point of $\overline{\mathbb{H}^2}$.)

Proof of Theorem 4.4 We divide the whole proof into two parts. The first part is about the bubbling compactness, and the second part is about the energy identity. Note that bubbling convergence with energy identity implies varifold convergence by [6, Proposition A.3]. (Even though in [6, Proposition A.3] the domain of the maps is the sphere, the proof works in our case for a disk domain with no change.)

Bubbling convergence Let u^j be a sequence as in the theorem. Then property (†) implies property (B_0) in Definition 7.1. By the compactness result Theorem 7.2, we can find a free boundary harmonic map $v_0\colon (D,\partial D)\to(\mathcal{N},\Gamma)$ (which may be trivial), and a finite collection of singular points $\mathcal{S}_0\subset\bar{D}$, such that a subsequence (still denoted by u^j) converges to v_0 weakly in $W^{1,2}(D)$ and strongly in $W^{1,2}(K)$ for any compact subset $K\subset\bar{D}\setminus\mathcal{S}_0$. The measures $|\nabla u^j|^2dx$ converge to a measure ν_0 with $\nu_0(\bar{D})\le E_0$, and at each singular point $x\in\mathcal{S}_0$, we have $\nu_0(\{x\})\ge\varepsilon_1$.

Next we want to renormalize the maps near the singular points. Let us start with boundary points. Suppose $x\in\mathcal{S}_0$ lies on ∂D. Let $\varepsilon_3>0$ be smaller than $\frac{1}{2}\varepsilon_1$ and ε_2. Fix a radius $\rho>0$ so that x is the only singular point in the boundary ball $B_{2\rho}(x)$ and $\int_{B_\rho(x)}|\nabla v_0|^2\le\varepsilon_3$. For each j, we choose $r_j>0$ to be the smallest radius for which

$$\inf_{y\in B_{\rho-r_j}(x)\cap\partial D}\int_{B_\rho(x)\setminus B_{r_j}(y)}|\nabla u^j|^2=\varepsilon_3,$$

and choose a point $y_j\in\partial D$ so that $B_{r_j}(y_j)\subset B_\rho(x)$ with $\int_{B_\rho(x)\setminus B_{r_j}(y_j)}|\nabla u^j|^2=\varepsilon_3$. Since u_j converges strongly to v_0 on any compact subset of $B_\rho(x)\setminus\{x\}$, by the energy bound we get $y_j\to x$ and $r_j\to 0$.

For each j, since the energy functional is invariant under conformal changes, the dilated sequence of maps $\tilde{u}_1^j=u^j\circ\Psi_{r_j,y_j}$ still satisfies the almost harmonic property (B_0) in Definition 7.1, and they all have the same energy as the u^j. Using the compactness result Theorem 7.2 again, we get a subsequence (still denoted by $\tilde{u}_1^j$), and a finite collection of singular points $\mathcal{S}_1\subset\bar{D}$, and a free boundary harmonic map $v_1\colon D\to\mathcal{N}$, such that $\tilde{u}_1^j$ converges to v_1 weakly in $W^{1,2}(D)$ and strongly in $W^{1,2}(K)$ for any compact $K\subset\bar{D}\setminus\mathcal{S}_1$. Moreover, the measures $|\nabla\tilde{u}_1^j|^2dx$ converge to a measure ν_1 on $\bar{D}$.

The choice of $B_{r_j}(y_j)$ guarantees that

$$\nu_1(\bar{D}\setminus\{p^+\})\le\nu_0(\{x\})\quad\text{and}\quad\nu_1(D^-)\le\nu_0(\{x\})-\varepsilon_3.$$

Next we want to show the following claim:

Claim *The maximal energy concentration at any $y\in\mathcal{S}_1\setminus\{p^+\}$ is at most $\nu_0(\{x\})-\varepsilon_3$.*

Proof of the claim Note that any such point y satisfies $\nu_1(\{y\})\ge\varepsilon_1>\varepsilon_3$, hence it can only lie in D^-. Then the fact $\nu_1(D^-)\le\nu_0(\{x\})-\varepsilon_3$ implies $\nu_1(\{y\})\le\nu_0(\{x\})-\varepsilon_3$.

Now we iterate this blowing-up construction at every boundary singular point in $\mathcal{S}_0$ and $\mathcal{S}_1$, and we will get subsequent singular sets $\mathcal{S}_0,\mathcal{S}_1,\mathcal{S}_2,\dots$, and dilated sequences of maps $\{u^j\},\{\tilde{u}_1^j\},\{\tilde{u}_2^j\},\dots$, one more singular set after each blowing-up process.

From the claim we know that this process must terminate after at most E_0/ε_3 steps, and we have in total m singular sets $\mathcal{S}_0, \mathcal{S}_1, \dots, \mathcal{S}_m$. Then for each sequence of dilated maps $\widetilde{u}^j_\alpha$, $\alpha \in \{0, \dots, m\}$, there are no boundary singular points away from $\mathcal{S}_\alpha$. Lastly we can apply the blowing-up process in [6, Appendix B.6] to each $\widetilde{u}^j_\alpha$ at those interior singular points, and finally get that the sequence u^j converges to a collection of free boundary harmonic disks $v_0, v_1, \dots, v_m\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ and harmonic spheres $\widetilde{v}_1, \dots, \widetilde{v}_k\colon S^2 \to \mathcal{N}$. Note that these harmonic spheres arise as blow-up limits near interior singular points.

Energy identity In this part we will show the summation of the energy of all the v_i is equal to the limit of the energy of the u^j, ie no energy was lost in the bubbling convergence process.

In order to prove this, we need to reexamine what happens to the energy during the blowing-up process. The "no loss of energy" for blowing up of interior singular points has already been proved by Colding–Minicozzi in [6, Appendix B.6, Step 4], so we only need to analyze the case for blowing up at boundary singular points.

At each blowing-up step, the energy is taken away from a singular point x and then goes to one of two places:

- it can show up in the new limiting free boundary harmonic disk of a singular point in $\partial D \setminus \{p^+\}$, or
- it can disappear at p^+.

In the first case, the energy is accounted for in the final summation and no energy is lost. So we only need to rule out energy loss in the second case. With out loss of generality, we can only prove the "no loss of energy" for the first blow-up process, ie for the convergence of $\{\widetilde{u}^j_1\}$. Note that if there is energy loss, then $\nu_1(\bar{D} \setminus \{p^+\}) < \nu_0(\{x\})$.

We argue by contradiction. Suppose $\nu_1(\bar{D} \setminus \{p^+\}) \le \nu_0(\{x\}) - \widehat{\delta}$ for some $\widehat{\delta} > 0$. Note that we must have $\widehat{\delta} \le \varepsilon_3$. Thus we can choose two sequences of radii $s_j > t_j$ so that each $A_j = B_{s_j}(y_j) \setminus B_{t_j}(y_j)$ is a half-cylinder with

$$s_j \to 0, \quad \frac{t_j}{r_j} \to \infty, \quad \int_{A_j} |\nabla u^j|^2 \ge \widehat{\delta} > 0. \tag{75}$$

Actually we may choose s_j close to ρ, and t_j close to s_j/λ_j, and $s_j/t_j > \lambda_j$ for a sequence $\lambda_j \to \infty$. After a conformal change, the A_j are a sequence of half-cylinders whose length goes to ∞. Moreover, there is quite a bit of freedom in choosing s_j

and t_j, ie we can change s_j and t_j a little bit and the above conditions are still satisfied. So we may also assume u^j has small energy near the two ends of the half-cylinder A_j as the second condition in Lemma 7.9.

Theorem 7.3 (with $\delta = \frac{1}{63}$) together with $\hat{\delta} \leq \varepsilon_3 \leq \varepsilon_2$ implies the theta energy of u^j on A_j is small, so the first condition of Lemma 7.9 is satisfied for j sufficiently large. Then by Lemma 7.9, we get that the area of the image of u^j on A_j must be strictly less than the energy of u^j on A_j for j large, which is a contradiction to the area assumption (37). Thus we complete the proof. □

8 Modifications for the proof of Theorem 0.3 and discussions

In this part, we record necessary modifications to adapt the proof of Theorem 0.1 to Theorem 0.3. There are only two places where we have to do some modifications.

Modification for Theorem 4.1 Again, the first step is to show that we can approximate a given $\gamma(\cdot, t) \in \Omega$ by some $\tilde{\gamma}(\cdot, t) \in \Omega$ which lies in $C^0([0,1], C^2(\bar{D}, \mathcal{N}))$. So we need to do mollifications on γ. However, direct mollifications as we did in Section 5 cannot work here, because the endpoint maps $\gamma(\cdot, 0) = \bar{v}_0$ and $\gamma(\cdot, 1) = \bar{v}_1$ may change after mollifications. In order to handle this issue, we first mollify the whole family $\gamma(\cdot, t)$ as in Section 5 to get a continuous family of C^2 maps $\bar{\gamma}(\cdot, t)\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$; next, by reparametrizing $t \to s(t) = (1-2\mu)t + \mu$, we get a new family $\tilde{\gamma}(\cdot, s)$ which is defined for $s \in [\mu, 1-\mu]$. Moreover, since $s(t) \to t$ as $\mu \to 0$ and $\bar{\gamma}(\cdot, t)$ is C^0 as a function of t to $C^2(\bar{D}, \mathcal{N})$, for any given $\varepsilon > 0$, we can choose μ small so that $\max_{s\in[\mu,1-\mu]} \|\tilde{\gamma}(\cdot, s) - \bar{\gamma}(\cdot, s)\|_{C^2} \leq \varepsilon$.

Note that $\gamma(\cdot, 0)$ and $\gamma(\cdot, 1)$ are both smooth, so by varying the mollification parameter we can connect $\gamma(\cdot, 0)$ and $\gamma(\cdot, 1)$ to $\tilde{\gamma}(\cdot, \mu)$ and $\tilde{\gamma}(\cdot, 1-\mu)$ respectively, and hence get a continuous family of C^2–maps $\tilde{\gamma}(\cdot, s)\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ for $s \in [0,1]$. The slices of $\tilde{\gamma}(\cdot, s)$ for $s \in [0, \mu]$ are mollifications of $\gamma(\cdot, 0)$ and the slices of $\tilde{\gamma}(\cdot, s)$ for $s \in [1-\mu, 1]$ are mollifications of $\gamma(\cdot, 1)$. Thus we get a regularization of $\gamma(t)$ which stays close to $\gamma(t)$ in $C^0(\bar{D}, \mathcal{N}) \cap W^{1,2}(D, \mathcal{N})$. The conformal reparametrization procedure works in the same way as Section 5. Note that since $\tilde{\gamma}(\cdot, 0) = \bar{v}_0$ and $\tilde{\gamma}(\cdot, 1) = \bar{v}_1$ are both conformal harmonic maps, the pull-back metrics of g on $\mathcal{N}$ are already conformal to the standard metric g_0 on D, so the conformal reparametrization maps $h(t)$ satisfy $h(0) = h(1) = \mathrm{id}$. Therefore after the conformal reparametrization procedure, the endpoint maps are still $\bar{v}_0$ and $\bar{v}_1$, so it is a legitimate sweepout homotopic to γ. We know that for this family $\tilde{\gamma}(\cdot, s)$, the area and the energy are close.

Modification for Theorem 4.3 In the tightening process, we only want to pull tight the slice with energy larger than the energy of $\overline{v}_0$ and $\overline{v}_1$. Therefore, we need to modify the theorem so that in the statement, $(*)$ holds for those t with $E(\tilde{\gamma}(\cdot,t)) \geq W/\lambda$, where $\lambda > 1$ is chosen with $W/\lambda > \max(\mathrm{Area}(\overline{v}_0), \mathrm{Area}(\overline{v}_1))$. All the proofs in Section 6 work in exactly the same way by changing $\frac{1}{2}W$ to W/λ.

Modification for Theorem 4.4 In the bubbling convergence procedure, we have one additional assumption on the sequence u^j: that is, when restricted to ∂D, $u^j\colon \partial D \to \Gamma$ is monotone. This is because the monotonicity is preserved when taking the limit.

Now we show that there is only one possible disk bubble. In order to show this, we only need to argue that each blow-up on the boundary can only generate a punctured sphere rather than a harmonic disk. Let's say u_k is generated from blowing up a boundary point on D. Then by the classical Courant–Lebesgue lemma, the length of $u_k|_{\partial D}$ is shorter than $\frac{1}{3}$ of the total length of Γ. Since the limiting series after conformal dilation are all monotone on the boundary, u_k should also be monotone on the boundary, which implies that u_k maps the whole boundary to a single point, ie u_k is a punctured sphere.

Some further discussions One main goal of the min-max construction for the fixed boundary problem is to produce a third nonminimizing minimal disk spanning Γ, which is a direct generalization of the work of Morse–Tompkins [34] and Shiffman [49] (see also Struwe [53]) to the Riemannian setting. There is one issue left open in our current result. In fact, it will be good if one can restrict the sweepouts to all those $\sigma \in \Omega_f$ where $\sigma(\cdot,t)\colon \partial D \to \Gamma$ is a monotone parametrization for each $t \in [0,1]$; if this could be done, our proof will show that one of the disk bubbles has monotone boundary parametrization, and all other disk bubbles must map the boundary ∂D to a point on Γ, so that they are punctured harmonic spheres. Thus if we assume additionally that the ambient manifold $\mathcal{N}$ has nonpositive curvature, then by the uniqueness of harmonic maps all of these (punctured) harmonic spheres must be constant, so the min-max solution we obtain in Theorem 0.3 gives an effective construction of a third nonminimizing (unstable) minimal disk; cf Hohrein's PhD thesis [21] for the general existence result of unstable minimal surfaces of higher genus in manifolds with nonpositive curvature. Unfortunately our current mollification process could possibly destroy the monotonicity property of the boundary parametrization in general. It will also be good to reduce the regularity of the boundary curve Γ to be just rectifiable, but our theory needs a nice Fermi neighborhood of Γ, which requires smoothness.

Appendix

For the definition of Fermi coordinate system near $\Gamma \subset \mathcal{N}$ we refer to [31, Appendix A]. We will say a quantity A is *α–close* to quantity B if we have $1-\alpha \le A/B \le 1+\alpha$.

Lemma A.1 *There is a constant $\kappa > 0$ depending only on $\mathcal{N}$, Γ and the isometric embedding $\mathcal{N} \hookrightarrow \mathbb{R}^N$, such that for any $x \in \Gamma$, the κ–neighborhood U_κ of x has a local Fermi coordinate system, satisfying the following condition: let g^1 be the metric of $\mathcal{N}$ under Fermi coordinates, g^2 be the standard Euclidean metric under the Fermi coordinate system, and g^3 be the standard Euclidean metric of $\mathbb{R}^N$. Then there is a constant $\alpha > 0$ depending on κ such that:*

(i) *For any vector $V \in T_p\mathcal{N} \subset T_p\mathbb{R}^N$ where $p \in U_\kappa$, $\|V\|_{g^k}$ is α–close to $\|V\|_{g^l}$ for $k, l \in \{1,2,3\}$.*

(ii) *For any pair of points $p_1, p_2 \in U_\kappa$, $\|p_1 - p_2\|_{g^2}$ is α–close to $\|p_1 - p_2\|_{g^3}$.*

Proof We only need to show that κ exists for any $x \in \Gamma$, then by a covering argument we can prove κ exists globally. Fix $x \in \Gamma$.

(i) By the definition of Fermi coordinates, we can choose the Fermi coordinates so that $g^1_{ij}(x) = g^2_{ij}(x) = \delta_{ij}$. Then if κ is small enough, locally $(1-\alpha)\,\mathrm{id} \le (g^1)^{-1} g^2 \le (1+\alpha)\,\mathrm{id}$, so we get (i) for $k, l \in \{1,2\}$. Since $\mathcal{N} \hookrightarrow \mathbb{R}^N$ is an isometric embedding, the lengths of V measured by g^1 and g^3 are close when κ is small enough, hence (i) is true.

(ii) First we show $\|p_1 - p_2\|_{g^2}$ is α–close to the distance between p_1, p_2 in $\mathcal{N}$. Note that for any curve connecting p_1 with p_2 in U_κ, the length of the curve evaluated under g^1 and g^2 should be α–close, since the lengths of the tangent vectors of the curve are α–close. Since the length between p_1 and p_2 is the shortest distance among all curves, we see that $\|p_1 - p_2\|_{g^2}$ is α–close to the distance between p_1 and p_2 on $\mathcal{N}$ (note that we can choose κ small so that U_κ is convex).

Since at each point x of $\mathcal{N}$ scaling up makes $\mathcal{N}$ converge to the tangent space at x, the distance for points close to x on $\mathcal{N}$ is equivalent to the distance in $\mathbb{R}^N$. Then apply this argument to all Fermi neighborhoods with small κ we get that $\|p_1 - p_2\|_{g^3}$ is α–close to the distance between p_1, p_2 in $\mathcal{N}$. It is easy to see that $\|p_1 - p_2\|_{g^3} \le \mathrm{dist}_{\mathcal{N}}(p_1, p_2)$, thus $\mathrm{dist}_{\mathcal{N}}(p_1, p_2)$ is α–close to $\|p_1 - p_2\|_{g^3}$ for some α. This proves the lemma. □

We also need the following Poincaré-type inequality with partial zero boundary values:

Lemma A.2 *Given a function* $f\colon D_r^+ \to \mathbb{R}$ *with* $f|_{\partial_r^A} = 0$*, there exists a constant* $C = C(r)$ *such that*

$$\int_{D_r^+} f^2 \le C \int_{D_r^+} |\nabla f|^2.$$

Proof We just extend f to the whole of D_r by letting $f(x,-y) = f(x,y)$ in Cartesian coordinates. Then f is a function on D_r which vanishes on the boundary, hence we can apply the classical Poincaré inequality. Note that this reflection doubles $\int f^2$ and $\int |\nabla f|^2$ at the same time, so we get the desired Poincaré inequality. □

A.1 Bubble convergence implies varifold convergence

The main goal of this section is to prove that bubble convergence implies varifold convergence; cf [6, Proposition A.3]. Let us recall some notions from geometric measure theory; for more details see [6, Section 1.3], and for further details see [50]. Let $\pi\colon G_2\mathcal{N} \to \mathcal{N}$ be the Grassmannian bundle of 2–planes over $\mathcal{N}$, and let us consider the pairs (X, F) where X is a compact surface (not necessarily connected) and $F\colon X \to G_2\mathcal{N}$ is a measurable map such that $f := \pi \circ F$ is in $W^{1,2}(X, \mathcal{N})$. We also use J_f to denote the Jacobian of f. We say that a sequence $X_i = (X_i, F_i)$ with uniformly bounded areas *varifold converges* to (X, F) if for all $h \in C^0(G_2\mathcal{N})$ we have

$$\int_{X_i} h \circ F_i J_{f_i} \to \int_X h \circ F J_f.$$

This is a kind of weak notion of convergence of measures. There exists a distance function d_V, the varifold distance, which induces this topology.

Here is one important example: a varifold induced by a map. Let $u\colon D \to \mathcal{N}$ be a $W^{1,2}$–map, then a pair (X, F) induced by u is constructed as follows: X is just D, and $F\colon X \to G_2\mathcal{N}$ is given by sending x to $du(T_x X)$. This is only defined on the measurable space where J_u is nonzero, but we can extend it arbitrarily to all of X since the corresponding Radon measure on $G_2\mathcal{N}$ (ie the *varifold induced by* u) given by $h \to \int_X h \circ F J_u$ is independent of the extension.

Proposition A.3 *If a sequence* $v^j\colon D \to \mathcal{N}$ *of* $W^{1,2}$*–maps bubble converges to a finite collection of smooth maps* $\{u_0, \dots, u_m\}$ *such that either* $u_i\colon (D, \partial D) \to (\mathcal{N}, \Gamma)$ *is a harmonic disk with free boundary or* $u_i\colon S^2 \to \mathcal{N}$ *is a harmonic sphere, and the energy identity holds, then this sequence also varifold converges to the varifold induced by* $\{u_0, \dots, u_m\}$.

We want to emphasize that the energy equality, which appears in the last part of Theorem 4.4, plays a key role in the proof.

Proof (see also [6, proof of Proposition A.3]) For each v^j we let V^j denote the corresponding map to $G_2\mathcal{N}$. Similarly for each u_i let U_i denote the corresponding map to $G_2\mathcal{N}$. By the construction of bubble convergence (in the proof of Theorem 4.4), we can choose $m+1$ sequences of domains $\Omega_0^j, \dots, \Omega_m^j \subset D$ that are pairwise disjoint for each j, so that for each $i=0,\dots,m$, applying $\Psi_{i,j}^{-1}$ (the inverse of the corresponding conformal dilation) to Ω_i^j gives a sequence of domains converging to either $D\setminus\mathcal{S}_i$ (if it is a disk bubble) or $S^2\setminus\mathcal{S}_i$ (if it is a sphere bubble), and they account for all the energy by the energy identity, ie

$$\lim_{j\to\infty}\int_{D\setminus(\bigcup_i \Omega_i^j)} |\nabla v^j|^2 = 0.$$

In order to show varifold convergence, we only need to show for any $h\in C^0(G_2\mathcal{N})$ that

$$\int_D h\circ U_i J_{u_i} = \lim_{j\to\infty}\int_{\Omega_i^j} h\circ V^j J_{v^j} = \lim_{j\to\infty}\int_{\Psi_{i,j}^{-1}(\Omega_i^j)} h\circ V^j\circ\Psi_{i,j} J_{(v^j\circ\Psi_{i,j})},$$

where the last equality is the change of variables formula.

Given $\varepsilon>0$ and i, let Ω_ε^i be the set where $J_{u_i}\geq\varepsilon$. Then we only need to show

$$\int_{\Omega_\varepsilon^i} h\circ U_i J_{u_i} = \lim_{j\to\infty}\int_{\Psi_{i,j}^{-1}(\Omega_i^j)} h\circ V^j\circ\Psi_{i,j} J_{(v^j\circ\Psi_{i,j})}.$$

Note that $J_{v^j\circ\Psi_{i,j}}\to J_{u_i}$ in L^1 since $v_i^j\to u_i$ in $W^{1,2}$, so the measure of

$$\{x\in\Omega_\varepsilon^i : J_{v^j\circ\Psi_{i,j}} < \tfrac{\varepsilon}{2}\}$$

goes to zero; the $W^{1,2}$–convergence implies that for given $\delta>0$, the measure of

$$\{x\in\Omega_\varepsilon^i : J_{v^j\circ\Psi_{i,j}} \geq \tfrac{\varepsilon}{2},\ |V^j\circ\Psi_{i,j}-U_i|\geq\delta\}$$

goes to zero. Then by the dominated convergence theorem we get the desired identity. Thus we conclude the varifold convergence. □

References

[1] **D R Adams**, *A note on Riesz potentials*, Duke Math. J. 42 (1975) 765–778 MR

[2] **F J Almgren, Jr**, *The homotopy groups of the integral cycle groups*, Topology 1 (1962) 257–299 MR

[3] **F J Almgren, Jr**, *The theory of varifolds*, mimeographed notes, Princeton University (1965)

[4] **J Chen**, **G Tian**, *Compactification of moduli space of harmonic mappings*, Comment. Math. Helv. 74 (1999) 201–237 MR

[5] **R Coifman**, **P-L Lions**, **Y Meyer**, **S Semmes**, *Compensated compactness and Hardy spaces*, J. Math. Pures Appl. 72 (1993) 247–286 MR

[6] **T H Colding**, **W P Minicozzi, II**, *Width and finite extinction time of Ricci flow*, Geom. Topol. 12 (2008) 2537–2586 MR

[7] **T H Colding**, **W P Minicozzi, II**, *Width and mean curvature flow*, Geom. Topol. 12 (2008) 2517–2535 MR

[8] **R Courant**, *Dirichlet's principle, conformal mapping, and minimal surfaces*, Springer (1977) MR

[9] **R Courant**, **N Davids**, *Minimal surfaces spanning closed manifolds*, Proc. Nat. Acad. Sci. U.S.A. 26 (1940) 194–199 MR

[10] **C De Lellis**, **J Ramic**, *Min-max theory for minimal hypersurfaces with boundary*, Ann. Inst. Fourier (Grenoble) 68 (2018) 1909–1986 MR

[11] **W Ding**, **J Li**, **Q Liu**, *Evolution of minimal torus in Riemannian manifolds*, Invent. Math. 165 (2006) 225–242 MR

[12] **W Ding**, **G Tian**, *Energy identity for a class of approximate harmonic maps from surfaces*, Comm. Anal. Geom. 3 (1995) 543–554 MR

[13] **L C Evans**, *Partial differential equations*, 2nd edition, Graduate Studies in Mathematics 19, Amer. Math. Soc., Providence, RI (2010) MR

[14] **A M Fraser**, *On the free boundary variational problem for minimal disks*, PhD thesis, Stanford University (1998) MR Available at `https://search.proquest.com/docview/304448599`

[15] **A M Fraser**, *On the free boundary variational problem for minimal disks*, Comm. Pure Appl. Math. 53 (2000) 931–971 MR

[16] **A Fraser**, **R Schoen**, *The first Steklov eigenvalue, conformal geometry, and minimal surfaces*, Adv. Math. 226 (2011) 4011–4030 MR

[17] **A Fraser**, **R Schoen**, *Sharp eigenvalue bounds and minimal surfaces in the ball*, Invent. Math. 203 (2016) 823–890 MR

[18] **L Grafakos**, *Modern Fourier analysis*, 2nd edition, Graduate Texts in Mathematics 250, Springer (2009) MR

[19] **M Grüter**, **J Jost**, *On embedded minimal disks in convex bodies*, Ann. Inst. H. Poincaré Anal. Non Linéaire 3 (1986) 345–390 MR

[20] **F Hélein**, *Harmonic maps, conservation laws and moving frames*, 2nd edition, Cambridge Tracts in Mathematics 150, Cambridge Univ. Press (2002) MR

[21] **J Hohrein**, *Existence of unstable minimal surfaces of higher genus in manifolds of nonpositive curvature*, PhD thesis, Universität Heidelberg (1994)

[22] **R A Hunt**, *On $L(p, q)$ spaces*, Enseign. Math. 12 (1966) 249–276 MR

[23] **T Iwaniec**, **G Martin**, *Geometric function theory and non-linear analysis*, Oxford Univ. Press (2001) MR

[24] **J Jost**, *Existence results for embedded minimal surfaces of controlled topological type, II*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. 13 (1986) 401–426 MR

[25] **J Jost**, *Two-dimensional geometric variational problems*, Wiley, Chichester (1991) MR

[26] **J Jost**, **L Liu**, **M Zhu**, *The qualitative behavior at the free boundary for approximate harmonic maps from surfaces*, Math. Ann. 374 (2019) 133–177 MR

[27] **T Lamm**, **L Lin**, *Estimates for the energy density of critical points of a class of conformally invariant variational problems*, Adv. Calc. Var. 6 (2013) 391–413 MR

[28] **P Laurain**, *Analyse des problèmes conformément invariants*, from "Séminaire Laurent Schwartz–Équations aux dérivées partielles et applications, 2016–2017" (F Golse, F Merle, editors), Ed. Éc. Polytech., Palaiseau (2017) Exposé XII MR

[29] **P Laurain**, **L Lin**, *Energy convexity of intrinsic bi-harmonic maps and applications, I: Spherical target*, preprint (2018) arXiv

[30] **P Laurain**, **R Petrides**, *Regularity and quantification for harmonic maps with free boundary*, Adv. Calc. Var. 10 (2017) 69–82 MR

[31] **M Li**, **X Zhou**, *Min-max theory for free boundary minimal hypersurfaces, I: Regularity theory*, preprint (2016) arXiv

[32] **F C Marques**, **A Neves**, *Min-max theory of minimal surfaces and applications*, from "Mathematical Congress of the Americas" (J A de la Peña, J A López-Mimbela, M Nakamura, J Petean, editors), Contemp. Math. 656, Amer. Math. Soc., Providence, RI (2016) 13–25 MR

[33] **R Montezuma**, *A mountain pass theorem for minimal hypersurfaces with fixed boundary*, preprint (2018) arXiv

[34] **M Morse**, **C Tompkins**, *The existence of minimal surfaces of general critical types*, Ann. of Math. 40 (1939) 443–472 MR

[35] **F Müller**, **A Schikorra**, *Boundary regularity via Uhlenbeck–Rivière decomposition*, Analysis (Munich) 29 (2009) 199–220 MR

[36] **T H Parker**, *Bubble tree convergence for harmonic maps*, J. Differential Geom. 44 (1996) 595–633 MR

[37] **T H Parker**, **J G Wolfson**, *Pseudo-holomorphic maps and bubble trees*, J. Geom. Anal. 3 (1993) 63–98 MR

[38] **S Poornima**, *An embedding theorem for the Sobolev space* $W^{1,1}$, Bull. Sci. Math. 107 (1983) 253–259 MR

[39] **J Qing**, *Boundary regularity of weakly harmonic maps from surfaces*, J. Funct. Anal. 114 (1993) 458–466 MR

[40] **J Qing**, **G Tian**, *Bubbling of the heat flows for harmonic maps from surfaces*, Comm. Pure Appl. Math. 50 (1997) 295–310 MR

[41] **T Rivière**, *Conservation laws for conformally invariant variational problems*, Invent. Math. 168 (2007) 1–22 MR

[42] **T Rivière**, *Conformally invariant variational problems*, preprint (2012) arXiv

[43] **T Rivière**, *A viscosity method in the min-max theory of minimal surfaces*, Publ. Math. Inst. Hautes Études Sci. 126 (2017) 177–246 MR

[44] **T Rivière**, **M Struwe**, *Partial regularity for harmonic maps and related problems*, Comm. Pure Appl. Math. 61 (2008) 451–463 MR

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

[46] **C Scheven**, *Partial regularity for stationary harmonic maps at a free boundary*, Math. Z. 253 (2006) 135–157 MR

[47] **B Sharp**, *Higher integrability for solutions to a system of critical elliptic PDE*, Methods Appl. Anal. 21 (2014) 221–240 MR

[48] **B Sharp**, **M Zhu**, *Regularity at the free boundary for Dirac-harmonic maps from surfaces*, Calc. Var. Partial Differential Equations 55 (2016) art. id. 27 MR

[49] **M Shiffman**, *The Plateau problem for non-relative minima*, Ann. of Math. 40 (1939) 834–854 MR

[50] **L Simon**, *Lectures on geometric measure theory*, Proceedings of the Centre for Mathematical Analysis, Australian National University 3, Aust. National Univ. Centre for Mathematical Analysis, Canberra (1983) MR

[51] **B Smyth**, *Stationary minimal surfaces with boundary on a simplex*, Invent. Math. 76 (1984) 411–420 MR

[52] **M Struwe**, *On a free boundary problem for minimal surfaces*, Invent. Math. 75 (1984) 547–560 MR

[53] **M Struwe**, *Plateau's problem and the calculus of variations*, Mathematical Notes 35, Princeton Univ. Press (1988) MR

[54] **M Struwe**, *Variational methods: applications to nonlinear partial differential equations and Hamiltonian systems*, 4th edition, Ergeb. Math. Grenzgeb. 34, Springer (2008) MR

[55] **L Tartar**, *Imbedding theorems of Sobolev spaces into Lorentz spaces*, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. 1 (1998) 479–500 MR

[56] **H C Wente**, *An existence theorem for surfaces of constant mean curvature*, J. Math. Anal. Appl. 26 (1969) 318–344 MR

[57] **R Ye**, *On the existence of area-minimizing surfaces with free boundary*, Math. Z. 206 (1991) 321–331 MR

[58] **X Zhou**, *On the existence of min-max minimal torus*, J. Geom. Anal. 20 (2010) 1026–1055 MR

[59] **X Zhou**, *On the free boundary min-max geodesics*, Int. Math. Res. Not. 2016 (2016) 1447–1466 MR

[60] **X Zhou**, *On the existence of min-max minimal surface of genus $g \geq 2$*, Commun. Contemp. Math. 19 (2017) art. id. 1750041 MR

Mathematics Department, University of California, Santa Cruz
Santa Cruz, CA, United States

Department of Mathematics, Massachusetts Institute of Technology
Cambridge, MA, United States

Department of Mathematics, University of California, Santa Barbara
Santa Barbara, CA, United States

lzlin@ucsc.edu, aosun@mit.edu, zhou@math.ucsb.edu

Proposed: Tobias H Colding Received: 28 November 2018
Seconded: Gang Tian, David Gabai Accepted: 18 July 2019