

A. Carlotto and A. Mondino (2014) "Existence of Generalized Totally Umbilic 2-Spheres in Perturbed 3-Spheres,"  
International Mathematics Research Notices, Vol. 2014, No. 21, pp. 6020–6052  
Advance Access Publication July 19, 2013  
doi:10.1093/imrn/rnt155

## Existence of Generalized Totally Umbilic 2-Spheres in Perturbed 3-Spheres

**Alessandro Carlotto<sup>1</sup> and Andrea Mondino<sup>2</sup>**

<sup>1</sup>Department of Mathematics, Stanford University, Sloan Hall, Stanford, CA 94305, USA and <sup>2</sup>Department of Mathematics, ETH Zentrum, CH-8093 Zürich, Switzerland

Correspondence to be sent to: [andrea.mondino@math.ethz.ch](mailto:andrea.mondino@math.ethz.ch)

It was recently shown by Souam and Toubiana [34] that the (nonconstantly curved) Berger spheres do not contain totally umbilic surfaces. Nevertheless, in this article we show, by perturbative arguments, that all analytic metrics sufficiently close to the round metric  $g_0$  on  $\mathbb{S}^3$  possess *generalized* totally umbilic 2-spheres, namely critical points of the conformal Willmore functional  $\int_{\Sigma} |A^\circ|^2 d\mu_\gamma$ . The same is true in the smooth setting provided a suitable nondegeneracy condition on the traceless Ricci tensor holds. The proof involves a gluing process of two different finite-dimensional reduction schemes, a sharp asymptotic analysis of the functional on perturbed umbilic spheres of small radius and a quantitative Schur-type Lemma in order to treat the cases when the traceless Ricci tensor of the perturbation is degenerate but not identically zero. For left-invariant metrics on  $SU(2) \cong \mathbb{S}^3$ , our result implies the existence of uncountably many distinct Willmore spheres.

### 1 Introduction

One of the most general problems in *extrinsic* Riemannian geometry is to find the *best immersion*  $\Phi$  of a given smooth manifold  $M$  in a higher dimensional ambient space  $N$  endowed with a Riemannian metric  $g$ , by which we mean that some special

Received April 5, 2013; Revised June 18, 2013; Accepted June 27, 2013

curvature condition is required. For instance, one might prescribe the vanishing of the second fundamental form  $A$  of  $M$  in  $N$  (in which case  $\Phi$  is called *totally geodesic*), or the vanishing of its trace  $H$  (in this case,  $\Phi$  is *minimal*) or, instead, the vanishing of its traceless component  $A^\circ$ , which corresponds to  $\Phi$  being *totally umbilic*. Over the last century, a number of different obstructions to the existence of such *optimal* immersions have been found: for example, it is well known that if  $M$  is compact and has no boundary, then it cannot be minimally immersed in the Euclidean space  $\mathbb{R}^n$ , as a basic consequence of the classical monotonicity formula for minimal submanifolds (or, maybe more simply, as a direct consequence of the maximum principle). If  $\dim_{\mathbb{R}}(M) = 2$  and  $\dim_{\mathbb{R}}(N) = 3$ , a recent obstruction was found by Souam and Toubiana [34], who proved that if  $(N, g)$  belongs to the class of the *Berger spheres*, then there exist no totally umbilic immersions of  $M$  in  $N$ , unless  $(N, g)$  is in fact a space form. It should be remarked that this result was then extended to any compact, homogeneous 3-manifold by Manzano and Souam (see [23]). Given these facts, it is then natural to weaken our requirement and to ask whether there exist immersions having the property to minimize (or, more generally, to be critical points for) some integral functional of  $|A^\circ|$ . This is the object of study of the present article.

Before proceeding further, let us introduce some notation. We will always deal with a compact, isometrically immersed, surface  $(\Sigma, \gamma)$  in  $(N, g)$ : the corresponding principal curvatures will be denoted by  $\lambda_1, \lambda_2$  and the mean curvature  $H$  of  $\Sigma$  will be their sum, namely  $H = \lambda_1 + \lambda_2$ . Moreover, according to our sign convention the round unit sphere in  $\mathbb{R}^3$  has mean curvature equal to 2.

In this work, we consider the *conformal Willmore functional* given by

$$I(\Sigma, \gamma) = \int_{\Sigma} \left( \frac{H^2}{4} - D \right) d\mu_{\gamma} = \frac{1}{2} \int_{\Sigma} |A^\circ|^2 d\mu_{\gamma}, \quad (1)$$

where  $d\mu_{\gamma}$  is the Riemannian volume form associated to the metric  $\gamma$ ,  $A^\circ := A - \frac{1}{2}H\gamma$  is the traceless part of the second fundamental form  $A$  of  $\Sigma$  in  $M$  and  $D = \lambda_1\lambda_2$ . In the special case when the sphere  $\mathbb{S}^3$  is endowed with its standard round metric  $g_0$  (as embedded unit sphere in  $\mathbb{R}^4$ ) then the previous two functionals coincide, modulo a null Lagrangian, with the functional

$$W(\Sigma, \gamma) = \int_{\Sigma} \left( \frac{H^2}{4} + 1 \right) d\mu_{\gamma}, \quad (2)$$

as an immediate consequence of the Gauss equations. However, from the point of view of conformal geometry, for immersions in a general Riemannian manifold, the

functional  $I$  is more natural than  $W$ , indeed the former is *conformally invariant* (see, for instance, [37]) while the latter may not be.

Since we deal with isometric immersions,  $\gamma$  is nothing but the pullback of the metric  $g$ , hence we will omit the  $\gamma$ -dependence of  $I$  and  $W$  (therefore we will only write  $I(\Sigma)$  and  $W(\Sigma)$ ).

In the last five decades, the study of the existence of critical points for these functionals and their geometric characterization has been the object of a number of works also due to the connections with other key questions in Geometric Analysis, like the classification of positive genus minimal surfaces in the round 3-sphere  $\mathbb{S}^3$  (this link, inspired by the works of Ros [30] and Urbano [35], was smartly exploited by Marques and Neves [23] in their proof of the Willmore conjecture), the regularity of complete properly embedded minimal surfaces in the Hyperbolic 3-space (see the recent paper of Alexakis and Mazzeo [1]), the study of sharp eigenvalue estimates in relation to the conformal volume theory by Li and Yau [21], etc.

Starting with the nowadays classical paper of Simon [33], the variational study of Willmore-type functionals for immersions *in the flat Euclidean space*  $\mathbb{R}^n$  has been extensively carried through, both in terms of existence and of regularity results (we recall the remarkable results of Kuwert and Schätzle, for instance, [15]). In parallel with Simon's ambient approach (involving geometric measure theory), recently Rivière developed a parametric approach to successfully attach the existence and regularity issues regarding the Willmore functional (see [31, 32]).

We stress that all the aforementioned existence results concern immersions *in the flat euclidean space*  $\mathbb{R}^n$  (or equivalently, thanks to the stereographic projections and the conformal invariance of the Willmore functional, for immersions in the round sphere  $\mathbb{S}^n$ ). Explicit examples of Willmore surfaces, or explicit bounds on the energy of some special submanifolds, in very symmetric ambient manifolds have been constructed by several authors (see, for instance, [6, 12, 20, 36], etc.). The existence of Willmore surfaces in nonconstantly curved ambient manifolds is a very recent topic started by the second author in a perturbative setting in [24, 25] (for existence of Willmore surfaces under area constraint, still in a perturbative setting, see the papers of Lamm and Metzger [16, 17] and Lamm et al. [18]). The minimization (i.e., the existence of a minimizer and the corresponding regularity theory), among smooth immersions of 2-spheres in a 3-manifold, of quadratic curvature functionals of the type  $\int |A|^2$  and  $\int |H|^2 + 1$  has been achieved by the second author in collaboration with Kuwert and Schygulla in [14] (see also [28] for the noncompact case). Finally, in collaboration with Rivière (see [26, 27]), the second author developed a parametric approach for studying the regularity of possibly

branched immersions which are critical points (possibly with constraints) of such curvature functionals in Riemannian manifolds (also of higher codimension), and applied this theory to the minimization of these functionals among possibly branched immersions of 2-spheres in homotopy groups (see also the paper of Chen and Li [9] for related results).

Let us remark that all these results in general Riemannian manifolds follow a *minimization* scheme; on the other hand, such a method cannot be applied for finding interesting spherical type critical points of the conformal Willmore functional  $I$  defined in (1). Indeed, given a point  $p$  in the 3-manifold  $M$  and denoted with  $S_{p,\rho}$  the geodesic sphere of center  $p$  (i.e., the sphere in geodesic coordinates centered at  $p$ ) and radius  $\rho > 0$ , it is easy to see that  $I(S_{p,\rho}) \rightarrow 0$  as  $\rho \rightarrow 0$ . Therefore, the infimum, among smooth immersions of  $\mathbb{S}^2$  into  $M$ , of the functional  $I$  is zero and every minimizing sequence either collapses to a point or converges, in a suitable sense, to a totally umbilic surface. But both the situations are not interesting for our purposes: the former is a degeneration which does not give a geometric object in the limit, while the latter simply may not happen (and indeed it *does not* happen in our case, by the result of Souam and Toubiana).

Therefore, in order to study the critical points of the functional  $I$ , one could either perform a min – max scheme or use a perturbative method. The present paper is related to the second technique, the first one will be studied in a forthcoming work.

Now we can state the main theorem of this paper which answers, in a perturbative setting, the question of the existence of an *umbilically best immersion* in relation to the aforementioned obstruction given in [34].

**Theorem 1.1.** Let  $g_\varepsilon = g_0 + \varepsilon h$  be a Riemannian metric on  $\mathbb{S}^3$  for some analytic, symmetric  $(0, 2)$ -tensor  $h$ . There exists  $\bar{\varepsilon} \in \mathbb{R}_{>0}$  such that if  $\varepsilon \in (-\bar{\varepsilon}, \bar{\varepsilon})$ , then there exist embedded critical points for the conformal Willmore functional in metric  $g_\varepsilon$ .

$$I_\varepsilon(\Sigma) = \frac{1}{2} \int_\Sigma |A^\circ|^2 d\mu_{g_\varepsilon}.$$

More precisely, every Willmore surface we construct is a normal graph over a totally umbilic sphere of a smooth function  $w_\varepsilon$  converging to 0 in  $C^{4,\alpha}$  norm as  $\varepsilon \rightarrow 0$ .  $\square$

**Remark 1.2.** It follows from the arguments (as explained in the sequel of this Introduction) that the critical points we construct are in fact *saddle points* for  $I_\varepsilon$ . Moreover, a standard *bumpy-metric argument* shows that (in case  $(\mathbb{S}^3, g_\varepsilon)$  *does not* have constant sectional curvature) these are *generically nondegenerate* of index exactly 4. To our

knowledge, this is the first existence result for embedded Willmore surfaces of saddle type in a compact (nonconstantly curved) ambient manifold.  $\square$

Let us now briefly describe the logical scheme of the proof and, correspondingly, the structure of this article. The first basic idea is, loosely speaking, to compare the functional  $I_\varepsilon$  to its unperturbed counterpart  $I_0$ : indeed, if  $\varepsilon = 0$ , then we endow  $\mathbb{S}^3$  with the round metric  $g_0$  and clearly the conformal Willmore functional admits a 4D manifold of minimum points (where  $I_0$  is identically null) which is made of all totally umbilical spheres  $(S_{p,\rho})$  for  $p \in \mathbb{S}^3$  and  $\rho \in (0, \pi)$ . In fact, thanks to the quantization results by Bryant [8] (see also the recent paper [19] of Lamm and Nguyen for the branched case) it is well known that we can *separate such manifold*: the full moduli space of smooth immersions of  $\mathbb{S}^2$  into  $\mathbb{S}^3$  which are critical points of  $I_0$  consists of a countable number of connected components, the one of minimal  $I_0$ -energy corresponding to the *totally umbilic* 2-spheres in  $\mathbb{S}^3$  (for which  $I_0 = 0$ ), followed by a second component with a gap of exactly  $12\pi$ . As a result, for our purposes we can neglect all higher energy components and so, from now onwards, let us denote by  $Z'$  the (closure of the) critical sub-manifold of  $I_0$  where such functional vanishes identically. The perturbation scheme we need to apply is based on the explicit knowledge of the global topology of the critical manifold  $Z'$ : unfortunately, if  $\Sigma$  is *not oriented*, then we *cannot* simply identify  $Z'$  with  $\mathbb{S}^3 \times [0, \pi]$  since clearly the couples  $(p, r)$  and  $(\hat{p}, \pi - r)$  correspond to *the same sphere* (where we have denoted by  $\hat{p}$  the antipodal point of  $p$  in  $\mathbb{S}^3$ ). Instead, the set  $Z'$  is diffeomorphic to  $\text{Bl}_O D^4$ , namely the *real blow-up* at the origin of the unit disk  $D^4 \subseteq \mathbb{R}^4$ . We recall that

$$\text{Bl}_O D^4 = \{(x, l) \in D^4 \times \mathbb{R}P^3 \mid x \in l\},$$

and that  $\text{Bl}_O D^4 \cong D^4 \# \mathbb{R}P^4$  so that, as a result,  $Z'$  is a closed smooth *nonorientable* manifold with boundary.

Given this fact and in order to avoid unnecessary complications in our proof, we will therefore consider the same Willmore functional  $I_\varepsilon$  defined on *oriented* isometrically embedded spheres, so that the critical manifold at minimal energy for  $I_0$  (which we denote by  $Z$ ) is diffeomorphic to  $\mathbb{S}^3 \times [0, \pi]$ . We remark that this choice is not at all necessary for our arguments to work, yet in the former setting the logical structure of our proof would considerably lose in terms of conceptual clarity and effectiveness. For the sake of brevity, we will sometimes refer at  $Z$  as a *cylinder* and its subsets  $\mathbb{S}^3 \times \{0\}$  and  $\mathbb{S}^3 \times \{\pi\}$  will be called *bases* of such cylinder.

Given such manifold  $Z$ , our strategy is based on applying a finite-dimensional reduction of our problem in the spirit of Ambrosetti–Badiale [2, 3] (which in turn is

based on the classical Lyapunov–Schmidt reduction; for a sketch of the abstract method see Subsection 2.4, for a more extensive discussion including applications, see [4]; let us also recall the seminal paper of Ye [39] about geometric applications of perturbative techniques): namely, we construct a small perturbation  $Z_\varepsilon$  of  $Z$  which plays the role, for the functional  $I_\varepsilon$ , of a *natural variational constraint* in the sense that (interior) critical points of  $I_\varepsilon$  on  $Z_\varepsilon$  are in fact critical points for the unconstrained functional  $I_\varepsilon$ . The manifold  $Z_\varepsilon$  corresponds, in our setting, to small *graphical* perturbations of the totally umbilic spheres  $S_{p,\rho}$ : we construct a map  $w_\varepsilon(p, \rho)$  that associates, to each point  $p \in \mathbb{S}^3$  and radius  $\rho$ , a (smooth) function defined over the 2-sphere  $\mathbb{S}^2$ , identified with  $S_{p,\rho}$ . In fact, in performing this reduction we face a number of technical obstacles. First of all, the general method by Ambrosetti–Badiale refers to a Hilbert space setting, while in our case it is convenient to work with  $w \in C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})$  (the equation we need to solve in order to find critical points of  $I_\varepsilon$  has order 4) and therefore we need to adapt the construction to our specific setting. The second issue (which is related to the first) is that the construction of the map  $w_\varepsilon(\cdot, \cdot)$  (see Section 3) requires the second derivative operator of  $I_\varepsilon$  to be *uniformly elliptic*, which is true (by Equation (8)) only on relatively compact subsets of  $\mathring{Z} \simeq \mathbb{S}^3 \times (0, \pi)$ . As a result, in order to solve our problem by means of a finite-dimensional reduction we first need to show that *we can get rid of suitably small neighborhoods of the bases of  $Z$* . More precisely, we will show that the functional  $I_\varepsilon$  is strictly increasing for  $\rho \simeq 0$  (hence, by symmetry, for  $\rho \simeq \pi$ ) so that we can apply the reduction scheme to a suitable closed subcylinder of  $Z$  of the form  $\mathbb{S}^3 \times [\delta, \pi - \delta]$  for some suitably small  $\delta > 0$ . A delicate aspect, in doing this, is that we need to find  $\delta$  *not depending on  $\varepsilon$* . To this aim, we need to construct (in Section 4) a sort of *second* finite-dimensional reduction map (which we will still denote by  $w_\varepsilon(\cdot, \cdot)$ ) in order to study the asymptotics for  $\rho \rightarrow 0$  of  $I_\varepsilon$  on graphs over the spheres  $S_{p,\rho}$  for very small values of their radii. In other terms, we show that if  $I_\varepsilon$  has a critical point  $\tilde{w}$  which is a graphical perturbation of a totally umbilic sphere  $S_{p,\rho}$  of small radius, then in fact  $\tilde{w} = w_\varepsilon(p, \rho)$  and at that point we study the behavior, both in  $\rho$  and in  $\varepsilon$ , of the corresponding reduced functional  $I_\varepsilon(S_{p,\rho}(w_\varepsilon(p, \rho)))$ . This is based on the work performed in [25, Section 3] (where it is proved that  $I_\varepsilon(S_{p,\rho}(w_\varepsilon(p, \rho))) \simeq \frac{\pi}{5}\rho^4\|\mathring{\text{Ric}}\|^2 + O_\varepsilon(\rho^5)$ ); yet in our setting we do not fix a given Riemannian metric on  $\mathbb{S}^3$  and it is crucial for us to obtain estimates that are uniform in  $\varepsilon$ , at least for small values of this parameter.

A distinctive feature of our Theorem 1.1, compared with other perturbative results (like, for instance, [24, 25]), is that the perturbation  $h$  is *completely arbitrary* in the sense that we do not add any sort of technical nondegeneracy condition. This is possible thanks to the fact that if the traceless Ricci tensor of  $g_\varepsilon$  vanishes identically

on  $\mathbb{S}^3$ , then  $g_\varepsilon$  is homothetic to the round metric; if this is not the case, we can exploit the fact that all curvature tensors of  $g_\varepsilon$  are in fact analytic in  $\varepsilon$  and hence  $\mathring{\text{Ric}}$  can be expanded in  $\varepsilon$  (with analytic coefficients) and the problem can be suitably reduced to the fully degenerate case thanks to a quantitative Schur-type argument (we remark that an *integral*-quantitative Schur Lemma was proved by De Lellis and Topping [10] in case of positive Ricci curvature, but our arguments are independent from theirs: indeed on one hand we work in a perturbative regime, on the other hand we get *pointwise* estimates). We believe that the method used in this part of the argument is rather new and interesting in itself.

When the analyticity assumption on  $g_\varepsilon$  (or, equivalently, on  $h$ ) is removed, our argument still works provided we require that the expansion (in  $\varepsilon$ ) of the traceless Ricci tensor in *nondegenerate* at least at some point. This amounts to requiring that the variation  $h$  is not in the kernel of the linearization at the metric  $g_0$  of the traceless Ricci operator. Therefore, we can state the following smooth counterpart of our main result.

**Theorem 1.3.** Let  $g_\varepsilon = g_0 + \varepsilon h$  be a Riemannian metric on  $\mathbb{S}^3$  for some  $C^\infty$  symmetric  $(0, 2)$ -tensor  $h$ . There exists  $\bar{\varepsilon} \in \mathbb{R}_{>0}$  such that if  $\varepsilon \in (-\bar{\varepsilon}, \bar{\varepsilon})$ , then there exist embedded critical points for the conformal Willmore functional in metric  $g_\varepsilon$  provided  $h$  does *not* satisfy the following equation:

$$\frac{1}{2}\Delta_L h + \frac{1}{2}\mathcal{L}_{\delta G(h)^\sharp} g_0 - \frac{1}{3}g_0(\text{Ric}_{g_0}, h) + \frac{1}{3}(\delta^2 h)g_0 - \frac{1}{3}\Delta(\text{tr}_{g_0} h)g_0 = 0.$$

Here  $\Delta_L$  denotes the Lichnerowicz Laplace operator,  $\delta$  is the divergence (with respect to the metric  $g_0$ ),  $^\sharp$  is the standard musical isomorphism  $\Gamma(T^*M) \rightarrow \Gamma(TM)$  determined by the metric  $g_0$  and we have set  $G(h) = h - \frac{1}{2}(\text{tr}_{g_0} h)g_0$ .

Moreover, every Willmore surface we construct is a normal graph over a totally umbilic sphere of a smooth function  $w_\varepsilon$  converging to 0 in  $C^{4,\alpha}$  norm as  $\varepsilon \rightarrow 0$ .  $\square$

Our argument also implies multiplicity results whenever  $\mathbb{S}^3$  is endowed with left-invariant metrics with respect to its Lie group structure. Indeed, the manifold  $\mathbb{S}^3$  can be identified with the algebraic group

$$\begin{aligned} \text{SU}(2) &= \{A \in M_{2 \times 2}(\mathbb{C}) : \det(A) = 1, A^* = A^{-1}\} \\ &= \left\{ \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} : |z|^2 + |w|^2 = 1 \right\}. \end{aligned}$$

Now, if  $G$  is an analytic Lie group (in fact we know that any  $C^0$  Lie group always admits a unique analytic structure), then every left-invariant Riemannian metric is itself analytic



and therefore we are always in position to apply our Theorem 1.1 to left-invariant metrics on  $SU(2)$  without any nondegeneracy constraint. If we combine this fact with the trivial remark that the action of  $G = SU(2)$  on itself is *transitive*, we obtain the following remarkable consequence.

**Corollary 1.4.** Let  $g_\varepsilon = g_0 + \varepsilon h$  be a left-invariant metric on  $SU(2) \cong \mathbb{S}^3$ . There exists  $\bar{\varepsilon} \in \mathbb{R}_{>0}$  such that if  $\varepsilon \in (-\bar{\varepsilon}, \bar{\varepsilon})$ , then for every  $p \in \mathbb{S}^3$  there exists an embedded critical 2-sphere for the conformal Willmore functional (in metric  $g_\varepsilon$ ) passing through  $p$ . As a result, under these assumptions the functional  $I_\varepsilon$  has uncountably many distinct critical points.  $\square$

**Remark 1.5.** We would like to stress that both Theorem 1.1 and Corollary 1.4 can easily be extended, with almost no change in the proofs, to the case when  $g_\varepsilon = g_0 + h_\varepsilon$  with  $h$  a perturbation which is analytic in all of its variables and such that  $h_{\varepsilon=0} = 0$ . Therefore, it is clear that Corollary 1.4 does in fact apply to *any* left-invariant metric  $g$  on  $SU(2)$  which is sufficiently close to  $g_0$ .  $\square$

**Remark 1.6.** It is also appropriate to remark that the Berger spheres are indeed a one parameter family of left-invariant metrics on  $\mathbb{S}^3 \cong SU(2)$ , so that Corollary 1.4 implies the existence for them of uncountably many *generalized* totally umbilic 2-spheres, in sharp contrast with the negative result of Souam and Toubiana asserting that there are no totally umbilic 2-spheres at all. More generally, our multiplicity result apply to the subclass of left-invariant metrics  $\mathcal{G}$  defined by requiring that

$$g \in \mathcal{G} \Rightarrow g(X_i, X_j) = \delta_j^i \lambda_i \quad \text{for some } \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_{>0},$$

where

$$X_1 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad X_3 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

These clearly form a basis of the Lie algebra  $\mathfrak{su}(2)$  of  $SU(2)$ . Note that the Berger spheres correspond to the 1-parameter family in  $\mathcal{G}$  given by choosing  $\lambda_1 = \lambda$ ,  $\lambda_2 = \lambda_3 = 1$  and in that case  $X_1$  is tangent to the orbits of the Hopf circle action.  $\square$

The paper is structured as follows: in Section 2, we collect some preliminary results (both concerning perturbation schemes and expansions of curvature tensors), in



Section 3 we construct the manifold  $Z_\varepsilon$ , namely the finite-dimensional reduction map  $w$ , in Section 4 we study the sharp asymptotics of the functional  $I_\varepsilon$  for small radii, and finally we give in Section 5 a detailed proof of Theorem 1.1. The first and second variation formulas for the conformal Willmore functional, which are recalled in Section 2, are proved in the appendix at the end of this work.

## 2 Notation and Preliminary Results

In order to make more concise and readable the key arguments in the proof of Theorem 1.1, we collect in this section a number of useful results. For each of them, we will either provide a proof or give the reader an appropriate reference.

### 2.1 Notation

As anticipated in Section 1, it is convenient for us to consider small perturbations of the totally umbilic spheres  $S_{p,\rho}$  in  $\mathbb{S}^3$  of center  $p \in \mathbb{S}^3$  and radius  $\rho \in [0, \pi]$ ; such perturbations are of the form of normal graphs defined over the unit sphere  $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$  (the identification using the exponential map of the appropriate metric  $g_\varepsilon$ ). Coherently with [24, 25], we will denote by  $\Theta_1$  and  $\Theta_2$  the corresponding coordinate vector-fields on  $\mathbb{S}^2$  (induced by the standard polar coordinates on the unit sphere of the Euclidean space  $\mathbb{R}^3$ ). Due to technical reasons (specifically: the need to apply suitable Schauder estimates), we will take  $w \in C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})$ , which is the Banach space of functions whose fourth-order derivatives with respect to  $\Theta_i$ ,  $i = 1, 2$ , are  $\alpha$ -Hölder, for some specific  $\alpha \in (0, 1)$ . Denoted by  $\Delta_{\mathbb{S}^2}$  the Laplace–Beltrami operator on  $\mathbb{S}^2$ , we will often work with the fourth-order operator  $\Delta_{\mathbb{S}^2}(\Delta_{\mathbb{S}^2} + 2)$ , which induces a splitting of the Hilbert space  $L^2(\mathbb{S}^2)$  as follows:

$$L^2(\mathbb{S}^2) = \text{Ker}[\Delta_{\mathbb{S}^2}(\Delta_{\mathbb{S}^2} + 2)] \oplus \text{Ker}[\Delta_{\mathbb{S}^2}(\Delta_{\mathbb{S}^2} + 2)]^\perp.$$

We shall then consider  $C^{4,\alpha}(\mathbb{S}^2)$  as a subspace of  $L^2(\mathbb{S}^2)$ , hence there is an induced splitting as above and we can set

$$C^{4,\alpha}(\mathbb{S}^2)^\perp = C^{4,\alpha}(\mathbb{S}^2) \cap \text{Ker}[\Delta_{\mathbb{S}^2}(\Delta_{\mathbb{S}^2} + 2)]^\perp. \quad (3)$$

We remark that  $K = \text{Ker}[\Delta_{\mathbb{S}^2}(\Delta_{\mathbb{S}^2} + 2)]$  is finite-dimensional therefore closed, and  $C^{4,\alpha}(\mathbb{S}^2)^\perp$  is itself a Banach space with (the restriction of) the  $C^{4,\alpha}$ -norm. Finally, it is also convenient to name  $P : L^2(\mathbb{S}^2; \mathbb{R}) \rightarrow K^\perp$  the  $L^2$ -orthogonal projector to  $K^\perp$ .

Given a point  $p \in \mathbb{S}^3$ ,  $\rho \in (0, \pi)$  and a function  $w \in C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})$  (of suitably small norm), we are then in position to define a *perturbed geodesic sphere*, denoted by  $S_{p,\rho}(w)$  as the image of the map  $\Psi_{p,\rho,w,\varepsilon} : \mathbb{S}^2 \rightarrow \mathbb{S}^3$

$$\Psi_{p,\rho,w,\varepsilon}(\Theta) = \exp_p((\rho + w(\Theta))\Theta), \tag{4}$$

which is in fact a normal graph over  $S_{p,\rho} \hookrightarrow \mathbb{S}^3$ . We stress that here  $\exp_p$  denotes the exponential map defined on  $T_p\mathbb{S}^3$  for a given metric  $g_\varepsilon = g + \varepsilon h$  on  $\mathbb{S}^3$ , not necessarily the round one (in which case, we will add an explicit remark to our discussion).

Given  $a \in \mathbb{N}$ , any expression of the form  $L_p^{(a)}(w)$  denotes a *linear* combination of  $w$  and its derivatives (with respect to  $\Theta_1$  and  $\Theta_2$ ) up to order  $a$ . We allow the coefficients of such combination to depend (smoothly) on  $p, \rho$ , and  $\varepsilon$ , but we require the existence of a constant  $C$  (independent of these) so that

$$\|L_p^{(a)}(w)\|_{C^{k,\alpha}(\mathbb{S}^2)} \leq C \|w\|_{C^{k+a,\alpha}(\mathbb{S}^2)}, \quad k \in \mathbb{N}.$$

More generally, for  $b \in \mathbb{N}$ , any expression of the form  $Q_p^{(b)(a)}(w)$  denotes a polynomial expression involving monomials of degree *at least*  $b$ , each of these involving  $w$  and its derivatives up to order  $a$ . Again, we allow the coefficients to depend on our parameters, yet we require the existence of absolute constants giving bounds as above and also of the form

$$\|Q_p^{(b)(a)}(w_2) - Q_p^{(b)(a)}(w_1)\|_{C^{k,\alpha}(\mathbb{S}^2)} \leq C (\|w_2\|_{C^{k+a,\alpha}(\mathbb{S}^2)} + \|w_1\|_{C^{k+a,\alpha}(\mathbb{S}^2)})^{b-1} \|w_2 - w_1\|_{C^{k+a,\alpha}(\mathbb{S}^2)},$$

provided  $\|w_l\|_{C^a(\mathbb{S}^2)} \leq 1$ ,  $l = 1, 2$ . If the numbers  $a, b$  are not specified, we agree that they equal 4 and 2, respectively.

If  $x$  is a real variable and  $f : I \rightarrow \mathbb{R}$  is a function of  $x$  defined at least on some neighborhood of zero, we will write  $f(x) = O(|x|^\beta)$  (for some  $\beta \in \mathbb{R}_{>0}$ ) in order to mean that

$$\limsup_{x \rightarrow 0} \frac{|f(x)|}{|x|^\beta} < \infty.$$

When  $f$  depends (smoothly enough) on some other variable, say  $z$ , we will use the notation  $O_z(|x|^\beta)$  in order to stress the dependence on  $z$  (and, more specifically, to stress the fact that the remainder might not be uniform in  $z$ ). In our problem, we need to consider functionals and functions depending on several parameters, typically  $p \in \mathbb{S}^3$ ,  $\rho \in [0, \pi]$  (or possibly in a smaller interval) and  $\varepsilon \in (-\varepsilon^*, \varepsilon^*)$  for some suitably small  $\varepsilon^*$  and therefore

we will often write  $O_{p,\rho}(\varepsilon^\beta)$  and  $O_{p,\varepsilon}(\rho^\beta)$  whenever an estimate is gotten by *freezing* some of the parameters (e.g.,  $p, \rho$  and  $p, \varepsilon$ , respectively) and considering the asymptotics with respect to the other ones.

## 2.2 Riemannian geometry preliminaries

Given a Riemannian metric  $g$  on  $\mathbb{S}^3$ , we will only make use of the associated Levi-Civita connection  $\nabla$  and concerning all the corresponding curvature tensors we will follow the conventions given, for instance, on the book by Petersen [29]. It is a trivial, yet crucial remark that if  $g = g_0 + \varepsilon h$ , then all curvature tensors are *analytic* in  $\varepsilon$ : thus, for  $|\varepsilon| < \bar{\varepsilon}$  they can be expanded in power series of  $\varepsilon$  with smooth coefficients if  $h$  is, or more generally of class  $C^k$  if  $h$  is a tensor of class  $C^{k+2}$ . Both these statements follow at once from the local expression of the curvature tensors. In this work, we will mostly be interested in the Ricci curvature tensor  $\text{Ric}_{g_\varepsilon}$  of  $g_\varepsilon$  and in its trace-free part  $\mathring{\text{Ric}}_{g_\varepsilon} := \text{Ric}_{g_\varepsilon} - \frac{1}{3}R_{g_\varepsilon}g_\varepsilon$  where  $R_{g_\varepsilon}$  is the scalar curvature of the same metric. Concerning the perturbative expansion of  $\mathring{\text{Ric}}_{g_\varepsilon}$ , observe that

$$\|\mathring{\text{Ric}}_{g_\varepsilon}\|^2 = \varepsilon^2 T_p^{(2)}(h) + o(\varepsilon^2),$$

where  $T_p^{(2)}(h)$  denotes a nonnegative quadratic expression in the second derivatives of  $h$  and namely (see the statement of Theorem 1.3)

$$T_p^{(2)}(h) = \left(\frac{1}{2}\Delta_L h + \frac{1}{2}\mathcal{L}_{\delta_G(h)}g_0 - \frac{1}{3}g_0(\text{Ric}_{g_0}, h) + \frac{1}{3}(\delta^2 h)g_0 - \frac{1}{3}\Delta(\text{tr}_{g_0} h)g_0\right)^2.$$

Moreover, if  $T_p^{(2)}(h) \equiv 0$  identically on  $\mathbb{S}^3$  (which is a nongeneric condition on the perturbation  $h$ ), then locally (around any given point)  $\|\mathring{\text{Ric}}_{g_\varepsilon}\|^2 = \sum_{k \geq k_0} \varepsilon^k T_p^{(k)}(h)$  for some  $k_0 \geq 4$  and with suitably strong convergence in a (possibly smaller) neighborhood.

## 2.3 First and second variation formulas

Given  $p \in \mathbb{S}^3$  and  $\rho \in (0, \pi)$ , we state here the first and second variations of the functional  $I_0$  on the totally umbilic spheres  $S_{p,\rho}$  (with the pullback metric  $\gamma_0$  given by the restriction of  $g_0$ ), the proof being postponed to the appendix.

**Lemma 2.1.** Let us consider an isometrically immersed surface  $(\Sigma, \gamma)$  and a deformation  $F : \Sigma \times (-\sigma, \sigma) \rightarrow \mathbb{S}^3$  such that  $F(\Sigma, 0) = \Sigma$  and  $\frac{\partial F}{\partial s}(\Sigma, 0) = \nu$  where  $\nu$  is the (co-)normal

vector field of  $\Sigma$  (which is oriented, by assumption) in  $\mathbb{S}^3$  and  $u \in C^{4,\alpha}(\Sigma, \mathbb{R})$ . Then, if we set  $L$  to be the Jacobi operator of  $\Sigma$ , namely

$$Lu = -\Delta_{\Sigma,\gamma}u - (\text{Ric}(v, v) + |A|^2)u,$$

we have that:

- (1) the first variation formula for  $I_0$  is given by

$$\delta I_0(\Sigma)[u] = \int_{\Sigma} u \left[ \frac{1}{2}LH + \left( \frac{H^3}{4} + H \right) \right] d\mu_{\gamma}, \tag{5}$$

hence the first derivative operator is  $I'_0(\Sigma) = \frac{1}{2}LH + (\frac{H^3}{4} + H)$ , and  $(\Sigma, \gamma)$  is Willmore if and only if it satisfies the fourth-order equation  $LH + (\frac{H^3}{2} + 2H) = 0$ ;

- (2) if  $(\Sigma, \gamma) = (S_{p,\rho}, \gamma_0)$  is a totally umbilic sphere with the corresponding pull-back metric, then the second variation formula is given by

$$\delta^2 I_0(\Sigma)[u_1, u_2] = (I''_0 u_1, u_2)_{L^2(\Sigma,\gamma)} \quad \text{for } I''_0(\Sigma)[u] = \frac{1}{2}\Delta_{\Sigma,\gamma_0} \left( \Delta_{\Sigma,\gamma_0} + \frac{H^2}{2} + 2 \right) u, \tag{6}$$

where  $H = \frac{\sin(2\rho)}{\sin^2(\rho)}$ . □

**Remark 2.2.** For the purpose of the present work, it is convenient to pull back  $u$  to the standard unit sphere  $\mathbb{S}^2$  (note that in (6) the operator  $\Delta_{\Sigma,\gamma}$  depends on the metric  $\gamma$  induced on the CMC 2-sphere  $\Sigma$ , while it would be much more convenient to work with a *normalized operator*, making the dependence on  $\rho$  explicit) and to this aim, we define the following correspondence:

$$u \in C^{4,\alpha}(S_{p,\rho}; \mathbb{R}) \rightsquigarrow w \in C^{4,\alpha}(\mathbb{S}^2; \mathbb{R}) : w(\Theta) = u(\exp_p(\rho\Theta)). \tag{7}$$

For the sake of clarity, let us set  $f = \exp_p(\rho \cdot)$  so that one simply has  $w = u \circ f$ . Then by the scaling properties of the Laplace–Beltrami operator (and the Gauss Lemma) we get

$$\Delta_{\Sigma,\gamma}u(q) = \frac{1}{\sin^2(\rho)} \Delta_{\mathbb{S}^2}w(f^{-1}(q)),$$

and so, as a result, our second derivative operator takes the final form

$$I''_0[w] = \frac{1}{2 \sin^4(\rho)} \Delta_{\mathbb{S}^2}^2 w + \frac{1}{\sin^2(\rho)} \Delta_{\mathbb{S}^2} w + \frac{\sin^2(2\rho)}{4 \sin^6(\rho)} \Delta_{\mathbb{S}^2} w;$$

by the well-known trigonometric identity  $\sin(2\rho) = 2 \sin(\rho) \cos(\rho)$ , we end up getting

$$I_0''(S_{p,\rho})[w] = \frac{1}{2 \sin^4(\rho)} \Delta_{\mathbb{S}^2}(\Delta_{\mathbb{S}^2} + 2)w. \quad (8)$$

Note that here we are identifying the spaces  $C^{4,\alpha}(S_{p,\rho}; \mathbb{R})$  and  $C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})$ , so that the functional  $I_0$  is in fact defined on the latter of these (coherently with [24, 25]) and we will always stick to this convention in the sequel. As a further remark, observe that the operator  $I_0''[w]$  has all the scaling and symmetry properties we might expect and, more specifically, it is invariant under the map  $\rho \mapsto \pi - \rho$ , as it must be.  $\square$

**Remark 2.3.** It is easily checked from (8) that the operator  $I_0''(S_{p,\rho})$  is Fredholm of index 0; moreover, its kernel  $K \subset L^2(\mathbb{S}^2, \mathbb{R})$  is given by the linear span  $\langle 1, x_1, x_2, x_3 \rangle$ , where  $x_1, x_2, x_3$  are the restrictions of the coordinate functions of  $\mathbb{R}^3$  to  $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ .  $\square$

#### 2.4 Perturbation methods: the Lyapunov–Schmidt reduction

The most basic idea behind our approach is to find critical points of the functional  $I_\varepsilon$  by applying a *finite-dimensional reduction*, after which our main theorem will follow by showing that a certain function of four variables defined on  $Z$  has an interior maximum point. The tool we need is a sort of *generalized implicit function theorem*, which is usually referred to as *Lyapunov–Schmidt reduction*. We recall here its general formulation (see [2–4] for a wider discussion of the method).

Let  $H$  be a Hilbert space and let us consider a suitably smooth functional  $J_\varepsilon : H \rightarrow \mathbb{R}$  of the form

$$J_\varepsilon(u) = J_0(u) + \varepsilon G(u),$$

for some  $J_0 \in C^2(H; \mathbb{R})$  which plays the role of the *leading term* (namely the unperturbed functional) and where  $G \in C^2(H; \mathbb{R})$  is an additive perturbation. Let us *assume* that  $J_0$  has a *finite-dimensional* smooth manifold of critical points:

$$\mathcal{E} = \{\xi \in H \mid J_0'(\xi) = 0\}.$$

The general idea behind the method is that *if  $J_0$  satisfies suitable nondegeneracy conditions, then for  $\varepsilon$  small enough the functional  $J_\varepsilon$  has a finite-dimensional natural constraint, namely there exists a smooth finite-dimensional manifold  $\mathcal{E}_\varepsilon$  such that the*

critical point of  $J_\varepsilon$  constrained to  $\mathcal{E}_\varepsilon$  is in fact stationary points for  $J_\varepsilon$ . Such nondegeneracy conditions are:

- (i) for all  $\xi \in \mathcal{E}$ , one has that  $T_\xi \mathcal{E} = \text{Ker}(J_0''(\xi))$ ;
- (ii) for all  $\xi \in \mathcal{E}$ , the second derivative operator  $J_0''(\xi)$  is a Fredholm operator of index 0;

and the precise statement is the following.

**Theorem 2.4.** Suppose that the functional  $J_0$  has a critical manifold of dimension  $d$  and satisfies conditions (i) and (ii) above. Given a compact subset  $\mathcal{E}_c$  of  $\mathcal{E}$ , there exists  $\varepsilon_0 > 0$  such that for all  $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$  there is a smooth function  $w_\varepsilon : \mathcal{E}_c \rightarrow H$  satisfying the following three properties:

- (1) for  $\varepsilon = 0$ , it results  $w_\varepsilon(\xi) = 0$ , for all  $\xi \in \mathcal{E}_c$ ;
- (2)  $w_\varepsilon(\xi)$  is orthogonal to  $T_\xi \mathcal{E}$  for all  $\xi \in \mathcal{E}_c$ ;
- (3) the manifold  $\mathcal{E}_\varepsilon = \{\xi + w_\varepsilon(\xi) : \xi \in \mathcal{E}_c\}$  is a natural constraint for  $J_\varepsilon$ , by which we mean that if  $\xi_\varepsilon$  is a critical point for the function  $\Phi_\varepsilon : \mathcal{E}_c \rightarrow \mathbb{R}$  given by  $\Phi_\varepsilon(\xi) = J_\varepsilon(\xi + w_\varepsilon(\xi))$ , then  $u_\varepsilon = \xi_\varepsilon + w_\varepsilon(\xi_\varepsilon)$  is a critical point of  $J_\varepsilon$ . □

**Remark 2.5.** When applying this method, it is often difficult to characterize the map  $w$ , so that it is in fact necessary to upgrade this scheme showing that under suitable regularity assumptions on  $J_\varepsilon$  the function  $w_\varepsilon(\xi)$  is of order  $O(\varepsilon)$  uniformly for  $\xi \in \mathcal{E}_c$  (possibly depending on some parameters), so that as a result  $\Phi(\xi) = J_\varepsilon(\xi) + o(\varepsilon)$  and so it is sufficient to study the perturbed functional  $J_\varepsilon$  on the critical manifold  $\mathcal{E}$  (which is typically known). □

### 3 Finite-Dimensional Reduction of the Problem

In this section, we state and prove the two key lemmas that allow the finite-dimensional reduction of our problem.

The goal of this work is to solve, for suitably small  $\varepsilon > 0$  the Willmore equation  $I'_\varepsilon(S_{p,\rho}(w)) = 0$  in  $w \in C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})$ . Our *ansatz* is that in fact we can split this into two problems, namely

$$\begin{cases} P I'_\varepsilon(S_{p,\rho}(w)) = 0, \\ (I_{C^{0,\alpha}(\mathbb{S}^2; \mathbb{R})} - P) I'_\varepsilon(S_{p,\rho}(w)) = 0, \end{cases} \tag{9}$$

where  $P : C^{0,\alpha}(\mathbb{S}^2; \mathbb{R}) \rightarrow C^{0,\alpha}(\mathbb{S}^2; \mathbb{R})$  is the projection operator defined in Section 2.1, after Equation (3). In this section, we are concerned with the first of the two, called *auxiliary equation*; as it will be clear in the sequel, this equation is somehow simpler even if infinite-dimensional. In the next lemma we show, using an implicit function-type argument, that such equation is solvable at least for suitably small perturbative parameter, that is,  $0 < \varepsilon \ll 1$ .

**Lemma 3.1.** For each suitably small  $\delta > 0$ , there exist  $\varepsilon_0 = \varepsilon_0(\delta)$  and  $r_0 > 0$  such that, for  $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$  and  $\rho \in [\delta, \pi - \delta]$ , the auxiliary equation  $PI'_\varepsilon(S_{p,\rho}(w)) = 0$  has a unique solution  $w_\varepsilon(p, \rho) \in B(0, r_0) \subseteq C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})^\perp$ . Moreover,

- (1) the map  $w_\varepsilon(\cdot, \cdot) : \mathbb{S}^3 \times [\delta, \pi - \delta] \rightarrow C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})^\perp$  is  $C^1$ ;
- (2)  $\|w_\varepsilon(p, \rho)\|_{C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})} = O(\varepsilon)$  uniformly in  $(p, \rho) \in \mathbb{S}^3 \times [\delta, \pi - \delta]$ . □

**Proof.** Given the first variation formula (5), which has been derived for a generic isometric immersion of  $(\Sigma, \gamma)$  into a Riemannian 3-manifold  $(M, g)$ , it is immediate to note that in the special case of  $M = \mathbb{S}^3$  and  $g_\varepsilon = g_0 + \varepsilon h$ , then one has

$$I'_\varepsilon(S_{p,\rho}(w)) = I'_0(S_{p,\rho}(w)) + \varepsilon G(\varepsilon, S_{p,\rho}(w)), \tag{10}$$

where  $G(\cdot, \cdot)$  is a smooth function which is uniformly bounded for  $\varepsilon$  suitably small. It should be remarked that even though  $G(\varepsilon, \cdot)$  is defined on the perturbed sphere  $S_{p,\rho}(w) \subseteq \mathbb{S}^3$ , it is convenient (with slight abuse of notation) to consider it defined on  $\mathbb{S}^2$  instead: this being said, we observe that  $G$  is a smooth function of  $\varepsilon, \Theta, w, Dw, D^2w, D^3w, D^4w$  where  $D^{(k)}w$  synthetically denotes the  $k$ th order covariant derivatives of the function  $w \in C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})$ . Let us now define the remainder term  $R_{p,\rho}(w) = I'_0(S_{p,\rho}(w)) - I''_0(S_{p,\rho})[w]$  so that, thanks to (10), the auxiliary equation we aim at solving takes the expanded form

$$PI''_0(S_{p,\rho})[w] + PR_{p,\rho}(w) + \varepsilon PG(\varepsilon, S_{p,\rho}(w)) = 0. \tag{11}$$

At this point, in order to turn our problem into a fixed point equation (to be solved by iterative schemes), we recall that the second derivative operator  $I''_0(S_{p,\rho})$  given in (8), takes the form

$$I''_0(S_{p,\rho})[w] = \frac{1}{2 \sin^4(\rho)} \Delta_{\mathbb{S}^2}(\Delta_{\mathbb{S}^2} + 2)w.$$

Observe that, if  $f \in C^{0,\alpha}(\mathbb{S}^2, \mathbb{R})^\perp$  and  $u \in C^{4,\alpha}(\mathbb{S}^2, \mathbb{R})^\perp$  solves  $I''_0(S_{p,\rho})[u] = f$ , then (looking at  $I''_0$  as a composition of two *linear* bounded second-order elliptic operators and applying



the Schauder estimates)

$$\|u\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})} \leq C \|f\|_{C^{0,\alpha}(\mathbb{S}^2;\mathbb{R})}, \tag{12}$$

for some constant  $C = C(\alpha, \delta)$ . Therefore,  $I_0''(S_{p,\rho})^{-1} : C^{0,\alpha}(\mathbb{S}^2; \mathbb{R})^\perp \rightarrow C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})^\perp$  is well defined as a bounded operator and we can turn (11) into the equivalent *fixed point problem*

$$w = F_{\varepsilon,p,\rho}(w) \quad \text{where } F_{\varepsilon,p,\rho}(w) = -I_0''(S_{p,\rho})^{-1}[\varepsilon PG(\varepsilon, S_{p,\rho}(w)) + PR_{p,\rho}(w)].$$

Henceforth, the only issue is to show that, given  $\delta$  as in our statement, we can find a threshold  $\varepsilon_0$  so that for any  $|\varepsilon| < \varepsilon_0$ ,  $p \in \mathbb{S}^3$ , and  $\rho \in [\delta, \pi - \delta]$ , the map  $F_{\varepsilon,p,\rho} : B(0, r) \subseteq C^{4,\alpha}(\mathbb{S}^2; \mathbb{R}) \rightarrow B(0, r)$  is a contraction (for some  $r > 0$  small, to be determined).

Now, it is convenient to proceed in two steps:

*Step 1:* we show that there exist positive  $\varepsilon$  and  $r$  so small that  $F_{\varepsilon,p,\rho}(w)$  maps the ball  $B(0, r) \subseteq C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})$  into itself;

*Step 2:* we show that, possibly by further decreasing  $\varepsilon$  and  $r$  with respect to Step 1, we have that in fact that  $F_{\varepsilon,p,\rho}(w)$  acts as a contraction on such ball.

Concerning Step 1, we can make use of the argument given in [24, pp. 605–606] to show that given  $\delta$  as above there exists a constant  $C(\delta)$  such that

$$\|F_{\varepsilon,p,\rho}(w)\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})} \leq C(\delta) \|PI_0''(S_{p,\rho})^{-1}\|(\varepsilon + \|w\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})}^2),$$

and therefore, given the fact that the right-hand side of the previous inequality is quadratic in  $w$ , we can certainly choose  $r$  and  $\varepsilon$  so that  $F_{\varepsilon,p,\rho}$  is a selfmapping of  $B(0, r)$ . Concerning Step 2, we can make use of [24, Equation (A.7)] (its simple proof can be repeated verbatim) to get

$$\|F_{\varepsilon,p,\rho}(w_2) - F_{\varepsilon,p,\rho}(w_1)\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})} \leq C(\delta) \|PI_0''(S_{p,\rho})^{-1}\|(\varepsilon + 2r)\|w_2 - w_1\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})},$$

which implies the claim. As a result, we have shown that there exists  $\varepsilon_0$  such that for  $|\varepsilon| < \varepsilon_0$  the auxiliary equation  $PI_\varepsilon'(S_{p,\rho}(w)) = 0$  is uniquely solvable in  $w$ . Now, thanks to the well-known version of the Contraction Mapping Theorem in dependence of parameters (specifically: [5, pp. 447–449; 7, pp. 22–23]) one proves that  $w_\varepsilon(p, \rho)$  is  $C^0$  in  $(\varepsilon, p, \rho)$  and  $C^1$  in  $(p, \rho)$ . Clearly, this implies (just by uniform continuity) that

$$\lim_{\varepsilon \rightarrow 0} \|w_\varepsilon(p, \rho)\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})} = 0 \quad \text{uniformly for } p \in \mathbb{S}^3, \rho \in [\delta, \pi - \delta].$$

We need to improve this result to show that in fact  $\|w_\varepsilon(\mathbf{p}, \rho)\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})} = O(\varepsilon)$  uniformly on our domain, which would end the proof. To this aim, let us recall from (11) that the auxiliary equation takes the form

$$PI_0''(S_{\mathbf{p},\rho})[w] + \frac{1}{\sin^4(\rho)} Q^{(2)(4)}(w) = -\varepsilon PG(\varepsilon, S_{\mathbf{p},\rho}(w)),$$

which is more conveniently analyzed after dividing by  $\varepsilon$  namely in the form

$$PI_0''(S_{\mathbf{p},\rho})\left[\frac{w}{\varepsilon}\right] + \frac{1}{\varepsilon \sin^4(\rho)} Q^{(2)(4)}(w) = -PG(\varepsilon, S_{\mathbf{p},\rho}(w)).$$

Indeed, the right-hand side is uniformly bounded in  $C^{0,\alpha}$  norm for  $|\varepsilon| < \varepsilon_0$  and so the left-hand side has to be as well, but clearly the second summand is of order strictly higher than the first (since obviously  $\|Q^{(2)(4)}(w)\|_{C^{0,\alpha}(\mathbb{S}^2;\mathbb{R})} \leq C\|w\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})}^2$ ) and, as a result,  $PI_0''(S_{\mathbf{p},\rho})[\frac{w}{\varepsilon}]$  has to be uniformly bounded in  $C^{0,\alpha}$  for  $\varepsilon \rightarrow 0$ . Hence, it follows from Schauder estimates that the function  $\frac{w}{\varepsilon}$  is bounded in  $C^{4,\alpha}$ -norm, which is equivalent to  $\|w_\varepsilon(\mathbf{p}, \rho)\|_{C^{4,\alpha}(\mathbb{S}^2;\mathbb{R})} = O(\varepsilon)$ , as we claimed. ■

Following the general Lyapunov–Schmidt reduction as our model (as outlined in Section 2.4), we will now show that Lemma 3.1 determines a *natural constraint*, in the sense that the problem of proving existence of (conformal) Willmore surfaces is reduced to finding critical points of the  $C^1$ -function  $\Phi_\varepsilon : \mathbb{S}^3 \times [\delta, \pi - \delta] \rightarrow \mathbb{R}$  given by

$$\Phi_\varepsilon(\mathbf{p}, \rho) := I_\varepsilon(S_{\mathbf{p},\rho}(w_\varepsilon(\mathbf{p}, \rho))). \tag{13}$$

**Lemma 3.2.** Given  $\delta > 0$  suitably small, let  $\varepsilon_0$  and  $r_0$  be given as in the statement of Lemma 3.1 and let  $\Phi_\varepsilon$ , for  $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$  be the reduced functional defined above by Equation (13). Then there exists  $\varepsilon'_0 \in (0, \varepsilon_0)$  such that if  $|\varepsilon| < \varepsilon'_0$  and  $\Phi_\varepsilon$  has a critical point  $(\mathbf{p}_\varepsilon, \rho_\varepsilon) \in \mathbb{S}^3 \times (\delta, \pi - \delta)$ , then  $S_{\mathbf{p}_\varepsilon, \rho_\varepsilon}(w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon))$  is a critical point for  $I_\varepsilon$ . □

**Proof.** By construction, we already know that  $S_{\mathbf{p}_\varepsilon, \rho_\varepsilon}(w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon))$  solves the auxiliary equation

$$PI_\varepsilon'(S_{\mathbf{p}_\varepsilon, \rho_\varepsilon}(w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon))) = 0,$$

and so we only need to show that the orthogonal component of  $I_\varepsilon'(S_{\mathbf{p}_\varepsilon, \rho_\varepsilon}(w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon))) \in L^2(\mathbb{S}^2)$  vanishes as well, namely that  $(I_{L^2(\mathbb{S}^2;\mathbb{R})} - P)I_\varepsilon'(S_{\mathbf{p},\rho}(w)) = 0$ . To this aim, let us recall

from Remark 2.3 (based on the explicit formula (8)) that  $K = \text{Ker}[\Delta_{\mathbb{S}^2}(\Delta_{\mathbb{S}^2} + 2)]$  is spanned over  $\mathbb{R}$  by (the restriction to  $\mathbb{S}^2 \subset \mathbb{R}^3$  of) constant and affine functions; let us consider an orthonormal basis  $\{q_i^\varepsilon\}_{i=0,\dots,3}$  for  $K$  with respect to the  $L^2$ -inner product, obtained by normalizing such functions:

$$q_0 = \frac{1}{\sqrt{4\pi}}, \quad q_i = \sqrt{\frac{3}{4\pi}} x_i \text{ for } i = 1, 2, 3.$$

Let us then decompose  $I'_\varepsilon(S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon)))$  with respect to this basis:

$$I'_\varepsilon(S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon))) = \sum_{i=1}^4 A_{i,\varepsilon} q_i, \quad A_{i,\varepsilon} = (I'_\varepsilon(S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon))), q_i)_{L^2(\mathbb{S}^2)},$$

so that the assertion we need to prove reduces to showing that  $A_{i,\varepsilon} = 0$  for  $i = 0, 1, 2, 3$ . To that aim let us make explicit the condition that  $(p_\varepsilon, \rho_\varepsilon)$  is a stationary point for  $\Phi_\varepsilon$ :

$$\left[ \frac{\partial \Phi_\varepsilon}{\partial \rho} \right] (p_\varepsilon, \rho_\varepsilon) = 0, \quad \left[ \frac{\partial \Phi_\varepsilon}{\partial p_i} \right] (p_\varepsilon, \rho_\varepsilon) = 0,$$

where  $\frac{\partial}{\partial p_i}$  is computed in local coordinates around  $p_\varepsilon$  (and, specifically, by taking geodesic normal coordinates centered at  $p_\varepsilon$  so that  $\mathbb{S}^3 \setminus \hat{p}_\varepsilon$  is identified with  $\mathbb{R}^3$ ). For brevity, let us refer from now onwards to the variable  $\rho$  as  $p_0$  so that  $\frac{\partial}{\partial p_0}$  will stand for  $\frac{\partial}{\partial \rho}$  (this allows to use a unified notation). Moreover, let us set  $c_0 = \sqrt{4\pi}$  and  $c_i = \sqrt{\frac{4\pi}{3}}$  for  $i = 1, 2, 3$ . Therefore, we have

$$0 = dI_\varepsilon(S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon)))[X^{(i)}] \quad \text{for } i = 0, 1, 2, 3,$$

where  $X^{(i)}$  is the variation vector field (with respect to  $p_i$ ) of  $(S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon)))$ . We can decompose each of these vector fields (which are sections, defined over  $S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon))$  of the tangent bundle  $T\mathbb{S}^3$ ) into their tangential and normal component: namely, if  $\nu$  is an (outward-pointing) co-normal vector field along  $S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon))$  and  $\tau_1, \tau_2$  are a local orthonormal basis to the tangent space of the same sphere, we can write  $X^{(i)} = X_1^{(i)} \tau_1 + X_2^{(i)} \tau_2 + X_n^{(i)} \nu$ . Now clearly  $dI_\varepsilon(S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon)))[X_l^{(i)} \tau_l] = 0$ , for  $i = 0, 1, 2, 3$  and  $l = 1, 2$  because the tangential components of  $X^{(i)}$  only determine a re-parameterization of the same sphere  $S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon))$ . Therefore, we can reduce our analysis to the normal component, for which we have

$$(I'_\varepsilon(S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon))), X_n^{(i)})_{L^2(S_{p_\varepsilon, \rho_\varepsilon}(w_\varepsilon(p_\varepsilon, \rho_\varepsilon)))} = 0 \quad \text{for } i = 0, 1, 2, 3. \tag{14}$$

At this stage, using the fact that  $\|w_\varepsilon(\mathbf{p}, \rho)\|_{C^{4,\alpha}(\mathbb{S}^2)} = O_\rho(\varepsilon)$  for  $\varepsilon \rightarrow 0$  (which was proved in Lemma 3.1) and recalling Equation (4), namely  $\Psi_{\mathbf{p},\rho,w,\varepsilon}(\Theta) = \exp_p((\rho + w(\Theta))\Theta)$ , we get that

$$X_n^{(i)}(\Psi_{\mathbf{p},\rho,w,\varepsilon}(\Theta)) = c_i q_i + \frac{\partial}{\partial p_i}(w_\varepsilon(\mathbf{p}, \rho)(\Theta)) + O_\rho(\varepsilon), \quad i = 0, 1, 2, 3, \tag{15}$$

with the remainder term uniformly bounded for  $\rho \in [\delta, \pi - \delta]$ , as it is in our case. In order to check this, one needs to proceed quite differently for the  $i = 0$  and the  $i = 1, 2, 3$  cases. For the former, one starts by noticing that (thanks to the chain-rule and by linearity of the differential):

$$\begin{aligned} \frac{\partial \Psi}{\partial p_0} &= \frac{\partial \Psi}{\partial \rho} = (D \exp_p)_{(\rho+w(\Theta))\Theta} \left( \Theta + \frac{\partial w}{\partial \rho} \Theta \right) \\ &= (D \exp_p)_{(\rho+w(\Theta))\Theta}(\Theta) + \frac{\partial w}{\partial \rho} (D \exp_p)_{(\rho+w(\Theta))\Theta}(\Theta), \end{aligned}$$

and the claim easily follows combining the Gauss Lemma together with the fact that  $\Theta$  belongs, by definition, to the *unit* sphere in  $T_p\mathbb{S}^3$ . When  $i \geq 1$ , one has to consider the exponential map as a function of the basepoint as well: the conclusion (Equation (15)) is achieved by observing that, for the round metric on  $\mathbb{S}^3$ , the normal component of  $\frac{\partial \Psi}{\partial p_i}$  trivially equals  $x_i$  and recalling that  $\|w_\varepsilon(\mathbf{p}, \rho)\|_{C^{4,\alpha}(\mathbb{S}^2)} = O_\rho(\varepsilon)$  for  $\varepsilon \rightarrow 0$ .

Hence, going back to (14) and using (15), it follows that

$$0 = \sin^2(\rho_\varepsilon) \left[ c_i A_{i,\varepsilon} + \sum_{j=0}^3 A_{j,\varepsilon} \left( q_j, \frac{\partial}{\partial p_i}(w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon)) \right)_{L^2(\mathbb{S}^2)} \right] + O_\rho(\varepsilon), \quad i = 0, 1, 2, 3. \tag{16}$$

Since the system above for  $A_{i,\varepsilon}$ ,  $i = 0, 1, 2, 3$  is *homogeneous*, the claim will follow from showing that the  $4 \times 4$  matrix given by  $\sin^2(\rho_\varepsilon)[c_i \delta_{ij} + (q_j, \frac{\partial}{\partial p_i}(w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon)))_{L^2(\mathbb{S}^2)}] + O_\rho(\varepsilon)$  is nonsingular. From the conditions  $(w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon), q_i)_{L^2(\mathbb{S}^2)} = 0$ , by differentiating with respect to  $\rho$  and  $p_i$  for  $i = 1, 2, 3$ , we get that

$$\left( \frac{\partial}{\partial p_i} w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon), q_j \right)_{L^2(\mathbb{S}^2)} + \left( w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon), \frac{\partial q_j}{\partial p_i} \right)_{L^2(\mathbb{S}^2)} = 0,$$

and hence, recalling once again that  $\|w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon)\|_{C^{4,\alpha}(\mathbb{S}^2)} = O_\rho(\varepsilon)$  we obtain that also  $(\frac{\partial}{\partial p_i} w_\varepsilon(\mathbf{p}_\varepsilon, \rho_\varepsilon), q_j)_{L^2(\mathbb{S}^2)} = O_\rho(\varepsilon)$ . As a result, possibly taking  $\varepsilon'_0$  smaller than  $\varepsilon_0$  if  $\varepsilon \in (-\varepsilon'_0, \varepsilon'_0)$  the determinant of this matrix is not zero and so system (16) forces  $A_{i,\varepsilon} = 0$  for  $i = 0, 1, 2, 3$  which is what we had to prove. ■

#### 4 Small Radii Sharp Asymptotics

Recall that the reduction performed in the previous section, in particular in Lemma 3.1, was done for technical reasons on  $\mathbb{S}^3 \times [\delta, \pi - \delta]$  for any  $\delta > 0$ . The goal of the present section is to understand what happens at the boundary of the critical manifold  $Z = \mathbb{S}^3 \times [0, \pi]$ ; as anticipated in Section 1, we need a different method to extend the reduction suitably till the bases of such *cylinder*, where the spheres degenerate to points. To this aim, the strategy is essentially to *fix* the parameter  $\varepsilon$  (in a suitably small neighborhood of zero) and to use the radius  $\rho$  as perturbative parameter. This construction follows directly from some results proved in [25] about the expansion of the functional  $I$  on suitably small perturbed geodesic spheres in a given Riemannian 3-manifold  $(M, g)$  (which happens to be, in our case  $(\mathbb{S}^3, g_\varepsilon = g + \varepsilon h)$ ). The real issue is then to show that the reduction corresponding map  $w_\varepsilon(p, \rho)$  is sufficiently smooth in all its parameters (and, specifically, in  $\varepsilon$ ), since this is needed to give an upper bound (involving both  $\varepsilon$  and  $\rho$  together) for  $I_\varepsilon$  on such perturbed spheres. Note that, due to the *uniqueness* part in the statement of both Lemmas 3.1 and 4.1, the two construction can be glued together and give a global reduction map. Concerning the functional analytic setting, and specifically concerning the definition of the subspace  $C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})^\perp$  the reader is referred to Section 2.1.

**Lemma 4.1.** There exist  $\varepsilon_1, \rho_1 > 0$ ,  $r_1 > 0$  and a  $C^1$  map  $w_\varepsilon(\cdot, \cdot) : [-\varepsilon_1, \varepsilon_1] \times \mathbb{S}^3 \times [0, \rho_1] \rightarrow C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})^\perp$ ,  $(\varepsilon, p, \rho, \cdot) \mapsto w_\varepsilon(p, \rho)$  such that if  $S_{p,\rho}(w)$  is a critical point of the conformal Willmore functional  $I_\varepsilon$  (for some  $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ ) with  $(p, \rho, w) \in \mathbb{S}^3 \times [0, \rho_1] \times B(0, r_1)$ , then  $w = w_\varepsilon(p, \rho)$ . Moreover, the following properties are satisfied:

- (1) for any  $p \in \mathbb{S}^3$ , the map  $(\varepsilon, \rho) \mapsto w_\varepsilon(p, \rho)$  is  $C^\infty$ ;
- (2)  $\|w_\varepsilon(p, \rho)\|_{C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})} = O_\varepsilon(\rho^3)$  as  $\rho \rightarrow 0$  uniformly for  $p \in \mathbb{S}^3$ ;
- (3)  $\|\frac{\partial}{\partial \rho} w_\varepsilon(p, \rho)\|_{L^2(\mathbb{S}^2; \mathbb{R})} = O_\varepsilon(\rho^2)$  uniformly for  $p \in \mathbb{S}^3$ ;
- (4) one has that

$$\left\| w_\varepsilon(p, \rho) + \left( -\frac{1}{12} \rho^3 \operatorname{Ric}_p(\Theta, \Theta) + \frac{1}{36} \rho^3 R(p) \right) \right\|_{C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})} = O_\varepsilon(\rho^4) \quad \text{as } \rho \rightarrow 0;$$

- (5) one has that  $\|w_\varepsilon\|_{C^{4,\alpha}(\mathbb{S}^2)} = O(\varepsilon)$  uniformly for  $(p, \rho) \in \mathbb{S}^3 \times [0, \rho_1]$ , namely there exists a constant  $C = C(g_0, \varepsilon_1, \rho_1)$  such that  $\|w_\varepsilon\|_{C^{4,\alpha}(\mathbb{S}^2)} \leq C\varepsilon$ .  $\square$

**Proof.** The construction of this map was performed, for a *fixed* Riemannian metric  $g$  in [25, Lemma 3.10]. Here, we need to show that the map is smooth in the couple

$(\varepsilon, \rho)$  and that estimate (5) holds. Concerning the first assertion, let us recall from [25, Proposition 3.9] the explicit expansion of the first derivative of the conformal Willmore functional (the same result could be easily deduced from Section 2.3):

$$I'_\varepsilon(S_{p,\rho}(w)) = \frac{1}{2\rho^4} \Delta_{\mathbb{S}^2} (\Delta_{\mathbb{S}^2} + 2)w + \frac{1}{\rho} \left[ O_{p,\varepsilon}(\rho^0) + \frac{1}{\rho^2} L_p^{(4)}(w) + \frac{1}{\rho^4} Q_p^{(2)(4)}(w) \right].$$

It is convenient to perform a change of variable calling  $\tilde{w} := \rho^{-3}w$ , so that the equation above can be rewritten as

$$\rho I'_\varepsilon(S_{p,\rho}(\rho^3 \tilde{w})) = \frac{1}{2} \Delta_{\mathbb{S}^2} (\Delta_{\mathbb{S}^2} + 2)\tilde{w} + O_{p,\varepsilon}(\rho^0) + L_p^{(4)}(\tilde{w}) + \rho Q_p^{(2)(4)}(\tilde{w}). \tag{17}$$

It is easily checked from the (relatively straightforward) computation leading to (17), that all terms on the right-hand side are smooth (i.e.,  $C^\infty$ ) in the triple  $(\varepsilon, p, \rho)$ . Let us then fix a point  $p \in \mathbb{S}^3$  and consider the map  $F_p$  given by

$$F_p : [-\varepsilon_1, \varepsilon_1] \times [0, \rho_1] \times C^{4,\alpha}(\mathbb{S}^2)^\perp \rightarrow C^{0,\alpha}(\mathbb{S}^2)^\perp, \quad F_p(\varepsilon, \rho, \tilde{w}) = \rho P I'_\varepsilon(S_{p,\rho}(\rho^3 \tilde{w})),$$

so that we want  $\tilde{w} = \tilde{w}_\varepsilon(p, \rho)$  to be the function implicitly defined by the equation  $F_p(\varepsilon, \rho, \tilde{w}) = 0$ . Indeed, for  $\varepsilon = \rho = 0$  we get that (directly from (17) by projecting through  $P$ )  $F_p(0, 0, 0) = 0$  and moreover if we take the derivative  $\frac{\partial F_p}{\partial \tilde{w}}(0, 0, 0)$  (in the appropriate sense of Banach Calculus), we obtain

$$\frac{\partial F_p}{\partial \tilde{w}}(0, 0, 0) = \frac{1}{2} \Delta_{\mathbb{S}^2} (\Delta_{\mathbb{S}^2} + 2),$$

which is invertible from  $C^{4,\alpha}(\mathbb{S}^2)^\perp$  to  $C^{0,\alpha}(\mathbb{S}^2)^\perp$ . As a result, we get that (for that fixed point  $p \in \mathbb{S}^3$ ) there exist positive constants  $\varepsilon_p, \rho_p$  so that  $(\varepsilon, \rho) \rightarrow \tilde{w}_\varepsilon(p, \rho)$  is smooth, for  $\varepsilon \in (-\varepsilon_p, \varepsilon_p)$  and  $\rho \in (-\rho_p, \rho_p)$ . Thanks to the Implicit Function Theorem for functional depending on parameters (see, for instance, [7]) and the compactness of  $\mathbb{S}^3$  we get a global  $C^1$  map  $w := \rho^3 \tilde{w}$ , which coincides (thanks to the local uniqueness property in both constructions) with the map defined in [25, Lemma 3.10]. Assertions (2)–(4) follows then directly from [25, Lemma 3.10] (note the difference in the notation of  $w$ : what we call here  $w$  was called  $-\rho w$  in [25]; this explains the apparently different statements).

Concerning estimate (5), we can argue as follows. Given  $p \in \mathbb{S}^3$  and  $\rho \in [0, \rho_1]$ , let us consider the first-order Taylor expansion

$$w_\varepsilon(p, \rho) = w_0(p, \rho) + \varepsilon \frac{\partial w_\varepsilon}{\partial \varepsilon} \Big|_{\varepsilon=\xi_\varepsilon}(p, \rho),$$

for some  $\xi_\varepsilon \in (0, \varepsilon)$ ; hence, again by a local uniqueness argument we must have  $w_0(p, \rho) = 0$  (since totally umbilic spheres are trivially critical points for the conformal Willmore functional). It follows that

$$\|w_\varepsilon(p, \rho)\|_{C^{4,\alpha}(\mathbb{S}^2)} \leq \varepsilon \max_{\varepsilon \in [-\varepsilon_1, \varepsilon_1]} \left\| \frac{\partial w_\varepsilon(p, \rho)}{\partial \varepsilon} \right\|_{C^{4,\alpha}(\mathbb{S}^2)},$$

but now it is enough to observe that the composite map  $[-\varepsilon_1, \varepsilon_1] \times \mathbb{S}^3 \times [0, \rho_1] \rightarrow C^{4,\alpha}(\mathbb{S}^2; \mathbb{R})^\perp \rightarrow \mathbb{R}$  given by  $\| \frac{\partial w_\varepsilon(p, \rho)}{\partial \varepsilon} \Big|_{\varepsilon=0} \|_{C^{4,\alpha}(\mathbb{S}^2)}$  is  $C^0$  and is defined on a compact space, so that it attains a finite maximum value and this implies the claim. ■

These results being given, we analyze the asymptotics (in  $\rho$ , but depending on the parameter  $\varepsilon$ ) of the conformal Willmore functional.

**Lemma 4.2** ([25, Proposition 3.11], improved). Let  $\varepsilon_1, \rho_1$  be given by Lemma 4.1 and let  $p \in \mathbb{S}^3$ ,  $\rho \in [0, \rho_1]$ , and  $\varepsilon \in [-\varepsilon_1, \varepsilon_1]$ . For  $g_\varepsilon = g_0 + \varepsilon h$ , the expansion of the conformal Willmore functional on perturbed geodesics spheres  $S_{p,\rho}(w_\varepsilon(p, \rho))$  (determined by the previous Lemma 4.1) is

$$I_\varepsilon(S_{p,\rho}(w_\varepsilon(p, \rho))) = \frac{\pi}{5} \|\mathring{\text{Ric}}_{g_\varepsilon}(p)\|^2 \rho^4 + \Omega(\varepsilon, \rho),$$

with

$$|\Omega(\varepsilon, \rho)| \leq C \varepsilon^2 \rho^5 \quad \text{for } \varepsilon < \varepsilon_1 \text{ and } \rho < \rho_0,$$

for some constant  $C \in \mathbb{R}_{>0}$  that can be chosen independently of  $\varepsilon$  and  $\rho$ . □

From Subsection 2.2, we know that  $\|\mathring{\text{Ric}}_{g_\varepsilon}(p)\|^2 = \varepsilon^2 T_p^{(2)}(h) + o(\varepsilon^2)$  (where  $T_p^{(2)}(h)$  is a nonnegative quadratic function in the second derivatives of  $h$ , as we specified above), so we get that

$$I_\varepsilon(S_{p,\rho}(w_\varepsilon(p, \rho))) = \frac{\pi}{5} T_p^{(2)}(h) \varepsilon^2 \rho^4 + o(\varepsilon^2) \rho^4 + \Omega(\varepsilon, \rho), \tag{18}$$

which will be crucial in the sequel of this work.

**Proof.** Thanks to statement (5) of Lemma 4.1, we obtain that

$$I_\varepsilon(S_{p,\rho}(w_\varepsilon(p, \rho))) = O_{p,\rho}(\varepsilon^2) \quad \text{uniformly for } \rho < \rho_0, \quad p \in \mathbb{S}^3. \tag{19}$$

Indeed, let us first Taylor expand in the perturbative parameter  $\varepsilon$  to get

$$I_\varepsilon(S_{p,\rho}(w_\varepsilon(p, \rho))) = I_0(S_{p,\rho}(w_\varepsilon(p, \rho))) + \varepsilon G_1(S_{p,\rho}(w_\varepsilon(p, \rho))) + \varepsilon^2 G_2(S_{p,\rho}(w_\varepsilon(p, \rho))) + o_{p,\rho}(\varepsilon^2),$$



and then let us expand in  $w = w_\varepsilon(p, \rho)$  the first two summands above

$$\begin{aligned} I_0(S_{p,\rho}(w_\varepsilon(p, \rho))) &= I_0(S_{p,\rho}) + I'_0(S_{p,\rho})[w_\varepsilon(p, \rho)] + \frac{1}{2}I''_0(S_{p,\rho})[w_\varepsilon(p, \rho), w_\varepsilon(p, \rho)] + o_{p,\rho}(\varepsilon^2), \\ G_1(S_{p,\rho}(w_\varepsilon(p, \rho))) &= G_1(S_{p,\rho}) + G'_1(S_{p,\rho})[w_\varepsilon(p, \rho)] + o_{p,\rho}(\varepsilon). \end{aligned}$$

Now, using statement (5) of Lemma 4.1 repeatedly we see that all terms are (at least) uniformly quadratic in  $\varepsilon$  apart from  $I_0(S_{p,\rho})$ ,  $I'_0(S_{p,\rho})[w_\varepsilon(p, \rho)]$ , and  $G_1(S_{p,\rho})$  which are all exactly zero and so (19) follows ( $I_0(S_{p,\rho})$  and  $I'_0(S_{p,\rho})$  are clearly null on the totally umbilic spheres, the computation of  $G_1(S_{p,\rho})$  is analogous to the proof of [25, Lemma 4.6]). At this point, let us recall from [25, Proposition 3.1.1] that in fact we already know that

$$I_\varepsilon(S_{p,\rho}(w_\varepsilon(p, \rho))) = \frac{\pi}{5}T_p^{(2)}(h)\varepsilon^2\rho^4 + o_p(\varepsilon^2)\rho^4 + \Omega(\varepsilon, \rho), \quad \Omega(\varepsilon, \rho) = O_{p,\varepsilon}(\rho^5), \quad (20)$$

so that, by comparison with (19) we obtain that

$$\Omega(\varepsilon, \rho) = O_{p,\varepsilon}(\rho^5) = O_{p,\rho}(\varepsilon^2) - \frac{\pi}{5}T_p^{(2)}(h)\varepsilon^2\rho^4 - o_p(\varepsilon^2)\rho^4.$$

At this point, let us first observe that  $\Omega(\varepsilon, \rho) \in C^\infty((-\varepsilon_1, \varepsilon_1) \times (-\rho_1, \rho_1))$  because *all* other terms in (20) have this degree of regularity (the left-hand side as a consequence of the previous Lemma 4.1, and the term  $o_p(\varepsilon)$  because it is a remainder term in the expansion of the traceless Ricci tensor at  $p$ , hence it does not depend on  $\rho$ , while it depends analytically on  $\varepsilon$ ). Therefore, our claim comes from the following elementary lemma, whose easy proof (based on a Taylor expansion) is omitted.

**Lemma 4.3.** Given  $x^*, y^* \in \mathbb{R}_{>0}$ , let  $f \in (C^{h+k}([-x^*, x^*] \times [-y^*, y^*]); \mathbb{R})$  such that

$$f(x, y) = O_x(|y|^k) \quad \text{for } y \rightarrow 0$$

and

$$f(x, y) = O_y(|x|^h) \quad \text{for } x \rightarrow 0.$$

Then there exists a constant  $M \in \mathbb{R}_{>0}$  such that  $|f(x, y)| \leq M|x|^h|y|^k$  for all  $x \in [-x^*, x^*]$  and  $y \in [-y^*, y^*]$ . ■

## 5 Proof of Theorem 1.1

In this section, we give a short and direct proof of the main theorem stated in Section 1 which makes use of the various auxiliary tools developed above.

**Proof.** Let  $\mathring{\text{Ric}}_{g_\varepsilon}(p)$  be the traceless Ricci tensor of the Riemannian manifold  $(\mathbb{S}^3, g_\varepsilon = g_0 + \varepsilon h)$  and let us recall from Subsection 2.2 that, due to the analyticity in  $\varepsilon$  of all curvature tensors, we have a local expansion of the form

$$\|\mathring{\text{Ric}}_{g_\varepsilon}(p)\|^2 = \sum_{k \geq k_0} \varepsilon^k T_p^{(k)}(h),$$

for some  $k_0 \geq 2$  depending on  $p$ . There are three distinct (and disjoint) cases we need to consider:

- (I) there exists a point  $\bar{p} \in \mathbb{S}^3$  such that  $k_0(\bar{p}) = 2$ , namely  $T_{\bar{p}}^{(2)}(h) \neq 0$ ;
- (II) for all points  $p \in \mathbb{S}^3$ , one has  $T_p^{(2)}(h) = 0$  but there exists  $\bar{p} \in \mathbb{S}^3$  such that  $T_{\bar{p}}^{(k)}(h) \neq 0$ , for some  $k \geq 3$ ;
- (III) the traceless Ricci tensor vanishes identically on  $(\mathbb{S}^3, g_\varepsilon)$ .

We now develop the proof separately for each of these three cases.

*Case III: the fully degenerate case.* First of all, note that for any  $h$  as in the statement of the theorem we can find a constant  $\varepsilon_{\text{III}}$  such that for  $\varepsilon \in (-\varepsilon_{\text{III}}, \varepsilon_{\text{III}})$  the  $(0, 2)$  tensor  $g_\varepsilon = g_0 + \varepsilon h$  actually defines a Riemannian metric. The assumption  $\mathring{\text{Ric}}_{g_\varepsilon} \equiv 0$  can be restated as

$$\text{Ric}_{g_\varepsilon} = \frac{1}{3} R_{g_\varepsilon} g_\varepsilon,$$

so the metric  $g_\varepsilon$  is Einstein and hence, by the Schur Lemma, it has constant scalar curvature. As a result, the Ricci tensor is parallel and  $g_\varepsilon$  has *constant sectional curvature* (see, for instance, [29, p. 38]). It follows that  $(\mathbb{S}^3, g_\varepsilon)$  is homothetic (namely: isometric modulo scaling) to  $\mathbb{S}^3$  and so the result is trivial (in fact, in this case there is a 4D manifold of critical points for the conformal Willmore functional given by the totally umbilic spheres in  $\mathbb{S}^3$ ).

*Case I: the nondegenerate case.* Let  $\varepsilon_1, \rho_1$  be given by Lemma 4.1. Using the assumption  $T_{\bar{p}}^{(2)} \neq 0$  together with Lemma 4.2, observe that we can choose  $\varepsilon_2 \in (0, \varepsilon_1]$  and  $\rho_2 \in (0, \rho_1]$  such that for every  $\varepsilon \in [-\varepsilon_2, \varepsilon_2]$  we have

$$\max_{\mathbb{S}^3 \times ([0, \rho_2] \cup [\pi - \rho_2, \pi])} I_\varepsilon(S_{p, \rho}(w_\varepsilon(p, \rho))) < \frac{\max_{\mathbb{S}^3 \times [0, \pi]} I_\varepsilon(S_{p, \rho}(w_\varepsilon(p, \rho)))}{2}. \tag{21}$$

As a second step, let us set  $\varepsilon_I = \min\{\varepsilon'_0, \varepsilon_2\}$  (where  $\varepsilon'_0$  is given by Lemma 3.2, applied with  $\delta = \rho_2$ ) and  $\bar{\rho} = \rho_2/2$ . Now, let  $\Phi_\varepsilon$  be the corresponding finite-dimensional reduced functional, namely let us set for a *fixed value*  $\varepsilon < \bar{\varepsilon}$

$$\Phi_\varepsilon : \mathbb{S}^3 \times [\bar{\rho}, \pi - \bar{\rho}] \rightarrow \mathbb{R}, \quad \Phi_\varepsilon(p, \rho) = I_\varepsilon(S_{p, \rho}(w_\varepsilon(p, \rho))).$$

By compactness,  $I_\varepsilon$  has (at least) a maximum point  $(\bar{p}, \bar{r}) \in \mathbb{S}^3 \times [\bar{\rho}, \pi - \bar{\rho}]$  and, from (21) and the definition of  $\bar{\rho}$  as  $\rho_2/2$ , this must be an *interior* maximum point. As a result, by the *naturality of the constraint* (Lemma 3.2), we conclude that correspondingly  $S_{p,\rho}(w_\varepsilon(\bar{p}, \bar{\rho}))$  is a critical point for the functional  $I_\varepsilon$ . Namely, this graph gives a conformal Willmore surface in the perturbed metric  $g_\varepsilon$ . By construction, such submanifold is a saddle point of  $I_\varepsilon$  of index 4, yet a possibly degenerate one.

*Case II: the degenerate case.* This is the most delicate case, and our strategy is to reduce ourselves to Case III by applying a *quantitative* version of the (classical) Schur Lemma. Indeed, let us suppose that the traceless Ricci tensor has no second-order term in the  $\varepsilon$ -expansion at any point and let  $k_0 > 2$  be the minimum integer such that  $T_p^{(k_0)}(h) \neq 0$  for some  $p \in \mathbb{S}^3$ . Incidentally, observe that  $k_0$  must be even and hence, since  $\|\mathring{\text{Ric}}_{g_\varepsilon}\|^2$  is positive definite, at least equal to 4. Let us also denote by  $\bar{p}$  a point (which will be fixed from now onwards) where  $T_{\bar{p}}^{(k_0)}(h) \neq 0$ , so that

$$\|\mathring{\text{Ric}}_{g_\varepsilon}\|^2 = \varepsilon^{k_0} \left( T^{(k_0)}(h) + \sum_{j=1}^{\infty} \varepsilon^j T^{(k_0+j)}(h) \right) \quad \text{at all points.} \tag{22}$$

Starting from the identity

$$\text{Ric} = \mathring{\text{Ric}} + \frac{1}{3} Rg$$

(which we are going to apply for  $g = g_\varepsilon$ ) and taking the divergence of both left- and right-hand sides, we get

$$\delta(\text{Ric}) = \delta(\mathring{\text{Ric}}) + \frac{1}{3} dR,$$

so that, by means of the *contracted Bianchi identity*  $dR = 2 \delta(\text{Ric})$ , we have

$$dR = 6 \delta(\mathring{\text{Ric}}). \tag{23}$$

Now, given any point  $q \in \mathbb{S}^3$  let us choose a length-minimizing geodesic connecting  $\bar{p}$  to  $q$  (in the corresponding Riemannian metric  $g = g_\varepsilon$ ): if we integrate Equation (23) along that path, using our assumption (22), we obtain

$$R_\varepsilon(q) = R_\varepsilon(\bar{p}) + \varepsilon^{k_0/2} \Lambda_\varepsilon(q), \tag{24}$$

for some smooth function  $\Lambda_\varepsilon$  on  $\mathbb{S}^3$ . At this point, we can exploit (24) to get information on the full curvature tensor of  $(\mathbb{S}^3, g_\varepsilon)$ . Indeed, let us recall the *Ricci decomposition*

$$\text{Riem} = \frac{R}{2n(n-1)}g \cdot g + \frac{1}{n-2} \left( \text{Ric} - \frac{R}{n}g \right) \cdot g + W,$$

where  $\cdot$  stands for the *Kulkarni–Nomizu product* of two symmetric 2-tensors and  $W$  is the Weyl tensor (see, for instance, [11, p. 182]); for  $n=3$ , the Weyl tensor vanishes and therefore the previous reduces to

$$\text{Riem} = \frac{R}{12}g \cdot g + \overset{\circ}{\text{Ric}} \cdot g. \tag{25}$$

Making use of our assumption (22) and its consequence (24) into (25), we get that

$$\text{Riem}_{g_\varepsilon} = \frac{R_{g_\varepsilon}(\bar{p})}{12}g_\varepsilon \cdot g_\varepsilon + \varepsilon^{k_0/2} \widetilde{\text{Riem}}_\varepsilon, \tag{26}$$

for a suitable (0,4) *curvature-type* tensor  $\widetilde{\text{Riem}}_\varepsilon$ . Set  $r(\varepsilon) = \sqrt{\frac{6}{R_\varepsilon(\bar{p})}}$ , observe that  $(\mathbb{S}^3, g_\varepsilon)$  is locally isometric (then *globally isometric* since they are diffeomorphic) to an  $\varepsilon^{k_0/2}$ -perturbation of the round sphere of radius  $r(\varepsilon)$ , namely we can write

$$g_\varepsilon = r^2(\varepsilon)g_0 + \varepsilon^{k_0/2} \tilde{h}, \tag{27}$$

for some analytic, symmetric (0,2)-tensor  $\tilde{h}$ . Indeed, it is well known (see, for instance, [38, pp. 90–92]) that given a point  $q \in M = \mathbb{S}^3$  and denoted by  $x^1, x^2, x^3$  normal coordinates centered at  $q$  we can express an analytic metric  $g$  as a convergent power series with coefficients only depending on the curvature tensor and its covariant derivatives *at the point  $q$* :

$$g_{rs}(x) = \delta_{rs} + \sum_{n=2}^{\infty} \sum_{i_1, i_2, \dots, i_n=1}^3 E_{r,s,i_1 i_2 \dots i_n}^{(n)}(\text{Riem}) x^{i_1} x^{i_2} \dots x^{i_n}, \tag{28}$$

where for each  $n \geq 2$  the coefficient  $E^{(n)}$  consists of a finite number of summands, each one being the evaluation at  $q$  of a term of the form

$$\nabla^{\alpha_{i_1}} \text{Riem} * \nabla^{\alpha_{i_2}} \text{Riem} * \dots * \nabla^{\alpha_{i_{m_n}}} \text{Riem} \quad \text{for some } \alpha_{i_j} \in \mathbb{N}_{\geq 0} \text{ satisfying } \sum_{j=1}^{m_n} (\alpha_{i_j} + 2) = n. \tag{29}$$

The first terms in such expansion are well known:

$$g_{rs} = \delta_{rs} - \frac{1}{3} \sum_{i,j=1}^3 R_{irjs}(q) x^i x^j - \frac{1}{6} \sum_{i,j,k=1}^3 \nabla_i R_{jrks} x^i x^j x^k \\ + \frac{1}{120} \sum_{i,j,k,l=1}^3 \left\{ -6 \nabla_{ij}^2 R_{krls} + \frac{16}{3} \sum_{f=1}^3 R_{irjf} R_{kslf} \right\} (q) x^i x^j x^k x^l + \dots$$

Now, if we compute such expansion for a metric homothetic to  $g_0$  (specifically of the form  $r^2(\varepsilon)g_0$ ) and for our metric  $g_\varepsilon$  (just based on the estimate (26) and compare them, we immediately get that relation (27) holds provided we set

$$\tilde{h}_{rs} = \sum_{n=2}^{\infty} \sum_{i_1, i_2, \dots, i_n=1}^3 E_{r,s,i_1 i_2 \dots i_n}^{(n)}(\widetilde{\text{Riem}}_\varepsilon) x^{i_1} x^{i_2} \dots x^{i_n} \quad (30)$$

$$= -\frac{1}{3} \sum_{i,j=1}^3 (\widetilde{\text{Riem}}_\varepsilon)_{irjs} x^i x^j - \frac{1}{6} \sum_{i,j,k=1}^3 \nabla_i^{g_\varepsilon} (\widetilde{\text{Riem}}_\varepsilon)_{jrks} x^i x^j x^k + \dots \quad (31)$$

By our analyticity assumption, this series converges on a ball of suitable radius and hence it is easy to check that it determines a well-defined symmetric (0,2)-tensor  $\tilde{h}$ .

At this point, let us consider Equation (27): if we expand in  $\varepsilon$  both left-hand side ( $g_\varepsilon = g_0 + \varepsilon h$ ) and right-hand side and compare the two we get that it implies the existence of two real numbers  $c, \tilde{c} \in \mathbb{R}$  (independent of  $\varepsilon$ ) so that

$$h = cg_0, \quad \tilde{h} = \tilde{c}g_0.$$

In fact,  $c$  and  $-\tilde{c}$  are just the coefficients of order 1 and  $k_0/2$ , respectively, in the  $\varepsilon$ -expansion of the function  $r^2(\varepsilon)$ . Clearly, the first of these two relations imply that in fact such  $g_\varepsilon$  should be *totally degenerate* (in the sense of Case III), which is a contradiction.

Therefore, the proof of the assertion follows from the arguments we gave for Case I and Case III. ■

**Remark 5.1.** Concerning our Remark 1.5 at the end of Section 1, we need to indicate how to modify the argument above in order to treat the more general case when  $g_\varepsilon = g_0 + h_\varepsilon$  with  $h_\varepsilon$  analytic in all of its variables. In that case the proof is exactly the same, with the only substantial difference that Case II cannot, in general, be reduced to Case III based on equation (27). Instead, after the deduction of (27) we simply need to observe that Case II can be reduced to Case I: we can recover all our auxiliary estimates (and, specifically,

Lemma 4.2) replacing  $\varepsilon$  by  $\varepsilon^{k_0/2}/r(\varepsilon)^2$  (recall that  $r(\varepsilon) \simeq 1$  and  $k_0 \geq 4$ ) and we can complete the proof following the very same argument used in the nondegenerate Case I treated above.  $\square$

## Acknowledgements

The authors would like to thank Andrea Malchiodi for introducing them to variational and perturbative methods in Nonlinear Analysis. Gratitude is also expressed to Otis Chodosh for his careful reading of the preliminary manuscript and to the referee for his constructive comments.

## Funding

This work was supported by NSF grant DMS/0604960 (to A.C.) and by the ETH-Fellowship (to A.M.).

## Appendix

We give here the proof of Lemma 2.1 stated in Section 2.3.

Let us briefly recall the setting:  $(\Sigma, \gamma)$  is an isometrically immersed surface, and  $F : \Sigma \times (-\sigma, \sigma) \rightarrow \mathbb{S}^3$  is a smooth variation such that  $F(\Sigma, 0) = \Sigma$  and  $\frac{\partial F}{\partial s}(\Sigma, 0) = u\nu$ , where  $\nu$  is the (co-)normal vector field of  $\Sigma$  in  $\mathbb{S}^3$  and  $u \in C^{4,\alpha}(\Sigma)$  (concerning the geometric quantities we follow the notation of [13, 18]).

**Proof.** Concerning the *first variation*, we just need to recall the well-known formulas:

$$\frac{d\mu_\gamma}{ds} = uHd\mu_\gamma, \quad \frac{\partial H}{\partial s} = Lu, \quad (\text{A.1})$$

where evaluation at  $s = 0$  is tacitly assumed and  $L$  is the Jacobi operator of  $\Sigma$ , namely

$$Lu = -\Delta_{\Sigma,\gamma}u - (\text{Ric}(\nu, \nu) + |A|^2)u \quad (\text{A.2})$$

(note the sign convention, which might be not entirely conventional). As a result

$$\delta I_0(\Sigma) = \frac{\partial}{\partial s_{s=0}} \int_{\Sigma_s} \left( \frac{H^2}{4} + 1 \right) d\mu_\gamma = \int_{\Sigma} \left[ \frac{1}{2} HLu d\mu_\gamma + \left( \frac{H^2}{4} + 1 \right) uH \right] d\mu_\gamma,$$

and hence, integrating by parts

$$\delta I_0(\Sigma) = \int_{\Sigma} u \left[ \frac{1}{2} LH + \left( \frac{H^3}{4} + H \right) \right] d\mu_{\gamma}.$$

Correspondingly, (conformal) Willmore surfaces are defined by (weakly) satisfying the fourth-order equation

$$LH + 2 \left( \frac{H^3}{4} + H \right) = 0. \tag{A.3}$$

Concerning the second variation, we have

$$\begin{aligned} \delta^2 I_0(\Sigma) &= \int_{\Sigma} \frac{\partial}{\partial s_{s=0}} \left[ \frac{1}{2} LH + \left( \frac{H^3}{4} + H \right) \right] u d\mu_{\gamma} \\ &\quad + \int_{\Sigma} \left[ \frac{1}{2} LH + \left( \frac{H^3}{4} + H \right) \right] \left( \frac{\partial u}{\partial s} + H u^2 \right)_{s=0} d\mu_{\gamma}, \end{aligned}$$

which reduces, for any critical point of the conformal Willmore functional  $I_0$  to

$$\delta^2 I_0(\Sigma) = \int_{\Sigma} \frac{\partial}{\partial s_{s=0}} \left[ \frac{1}{2} LH + \left( \frac{H^3}{4} + H \right) \right] u d\mu_{\gamma}.$$

In order to proceed further, let us set

$$\mathcal{I}u = \frac{\partial}{\partial s_{s=0}} \left[ \frac{1}{2} LH + \left( \frac{H^3}{4} + H \right) \right],$$

and observe that

$$\mathcal{I}(u) = -\frac{1}{2} \left[ \frac{\partial}{\partial s}, \Delta_{\Sigma, \gamma} \right] H + \frac{1}{2} LLu - \frac{1}{2} H \frac{\partial}{\partial s} \text{Ric}(v, v) - \frac{1}{2} H \frac{\partial |A|^2}{\partial s} + \left( \frac{3}{4} H^2 + 1 \right) Lu, \tag{A.4}$$

where  $[T_1, T_2]$  denotes the commutator of two (suitably regular) scalar operators. At this point, we can make use of the computations done in [18, Section 3] namely

$$\frac{\partial}{\partial s} \text{Ric}(v, v) = u \nabla_v \text{Ric}(v, v) - 2 \text{Ric}(\nabla u, v), \tag{A.5}$$

$$\frac{\partial}{\partial s} |A|^2 = -2u \text{tr} A^3 - 2A_{ij} \nabla^i \nabla^j u - 2u A^{ij} T_{ij}, \tag{A.6}$$

$$\left[ \frac{\partial}{\partial s}, \Delta_{\Sigma, \gamma} \right] z = Hg(\nabla u, \nabla z) - ug(\nabla z, \nabla H) - 2A(\nabla u, \nabla z) - 2u \text{Ric}(\nabla z, v) - 2ug(A, \nabla^2 z), \tag{A.7}$$

where we set  $T_{ij} = \text{Ric}_{ij} + G(v, v)\gamma_{ij}$ , with  $G = \text{Ric} - (R/2)g$  the (ambient) Einstein tensor.



Since we are interested in computing the second variation on totally umbilic spheres  $S_{p,\rho}$  in  $(\mathbb{S}^3, g_0)$ , we have the following simplifications:  $\nabla \text{Ric}(v, v) = 0$ ,  $A_{ij}^c \equiv 0$ , and the mean curvature is constant (depending on  $\rho$ ). Plugging (A.5)–(A.7) in (A.4), we obtain

$$\mathcal{I}u = \frac{1}{2}LLu + H\text{Ric}(\nabla u, v) + H[\text{utr}A^3 + g(A, \nabla^2 u) + ug(A, T)] + \left(\frac{3}{4}H^2 + 1\right)Lu.$$

Writing  $A = A^0 + \frac{1}{2}H\gamma$  and using the formulas computed at [18, p. 14], we finally get

$$\mathcal{I}u = \frac{1}{2}LLu + \left(\frac{H^2}{4} + 1\right)Lu + H, \text{Ric}(v, \nabla u).$$

Moreover, thanks to the identity  $\text{div}A^0 = \frac{1}{2}\nabla H + \text{Ric}(v, \cdot)^\sharp$  it follows that  $\text{Ric}(v, \nabla u) = 0$  for any variation  $u \in C^{4,\alpha}(S_{p,\rho})$  so that the previous formula simplifies and, as a result, the *second derivative operator* for  $I_0$  at  $S_{p,\rho}$  is given by (6), namely

$$I_0''[u] = \frac{1}{2}L^2u + \left(\frac{H^2}{4} + 1\right)Lu. \tag{A.8}$$

Observing that  $I_0''$  is  $L^2$ -selfadjoint we can write the associated second variation as  $d^2I_0(u_1, u_2) = (I_0''u_1, u_2)_{L^2(\Sigma, \gamma)}$ .

In order to make these formulas totally explicit, we need to compute the terms  $\text{Ric}(v, v)$  and  $|A|^2$  for any totally umbilic 2-sphere  $S_{p,\rho}$ . Clearly, the round metric  $g_0$  on  $\mathbb{S}^3$  is Einstein, so that obviously  $\text{Ric} = \frac{S}{3}g_0$  where  $S$  is the corresponding scalar curvature, which is exactly equal to 6 for the *unit sphere*, so that  $\text{Ric}(v, v) = 2$  at all points of each of the spheres  $S_{p,\rho}$ . Concerning the other terms, in a principal orthonormal frame at a given point one has

$$A = \begin{pmatrix} H/2 & 0 \\ 0 & H/2 \end{pmatrix},$$

so that  $|A|^2 = \frac{H^2}{2}$  and the computation reduces to determining the mean curvature  $H$  of  $S_{p,\rho}$ . To that aim, one can then use the first variation formula for the area functional, and easy computations give

$$H = \frac{\sin(2\rho)}{\sin^2(\rho)} \quad \text{so that } |A|^2 = \frac{\sin^2(2\rho)}{2\sin^4(\rho)}. \tag{A.9}$$

Note that for  $\rho \rightarrow 0^+$ , one has  $H \simeq 2/\rho$  and  $|A|^2 \simeq 2/\rho^2$ .

At this point, starting from (A.8) we can perform simple algebraic computations to write the second variation as

$$I_0''[u] = \frac{1}{2} \Delta_{\Sigma, \gamma}^2 u + \frac{1}{4} H^2 \Delta_{\Sigma, \gamma} u + \Delta_{\Sigma, \gamma} u = \frac{1}{2} \Delta_{\Sigma, \gamma} \left( \Delta_{\Sigma, \gamma} + \frac{H^2}{2} + 2 \right) u.$$

We complete the proof of Lemma 2.1 by replacing  $H$  with its explicit expression for  $S_{p, \rho}$  given in (A.9). ■

## References

- [1] Alexakis, S. and R. Mazzeo. "The Willmore functional on complete minimal surfaces in  $H^3$ : boundary regularity and bubbling." (2012): preprint arXiv:1204.4955.
- [2] Ambrosetti, A. and M. Badiale. "Homoclinics: Poincaré–Melnikov type results via a variational approach." *Annales de l'Institut Henri Poincaré. Analyse Non Linéaire* 15, no. 2 (1998): 233–52.
- [3] Ambrosetti, A. and M. Badiale. "Variational perturbative methods and bifurcation of bound states from the essential spectrum." *Proceedings of the Royal Society of Edinburgh Section A* 128, no. 6 (1998): 1131–61.
- [4] Ambrosetti, A. and A. Malchiodi. *Perturbation Methods and Semilinear Elliptic Problems in  $\mathbb{R}^n$* . Progress in mathematics. Boston: Birkhauser, 2006.
- [5] Ambrosetti, A., A. Malchiodi, and W. M. Ni. "Singularly perturbed elliptic equations with symmetry: existence of solutions concentrating on spheres." *Communications in Mathematical Physics* 235, no. 3 (2003): 427–66.
- [6] Barros, M. and A. Ferrández. "Willmore energy estimates in conformal Berger spheres." *Chaos, Solitons & Fractals* 44, no. 7 (2011): 515–21.
- [7] Bressan, A. *Hyperbolic Systems of Conservation Laws, The One-Dimensional Cauchy Problem*. Oxford: Oxford University Press, 2000.
- [8] Bryant, R. L. "A duality theorem for Willmore surfaces." *Journal of Differential Geometry* 20, no. 1 (1984): 23–53.
- [9] Chen, J. and Y. Li. "Bubble tree of a class of conformal mappings and applications to Willmore functional." (2011): preprint arXiv:1112.1818.
- [10] De Lellis, C. and P. Topping. "Almost-Schur lemma." *Calculus of Variations and Partial Differential Equations* 43, no. 3–4 (2012): 347–54.
- [11] Gallot, S., D. Hulin, and J. Lafontaine. *Riemannian Geometry*. Universitext Series. Berlin: Springer, 2004.
- [12] Hu, Z. and H. Song. "On Otsuki tori and their Willmore energy." *Journal of Mathematical Analysis and Applications* 395, no. 2 (2012): 465–72.
- [13] Huisken, G. and A. Polden. "Geometric Evolution Equations for Hypersurfaces." *Calculus of Variations and Geometric Evolution Problems*, 45–84. Lecture Notes in Mathematics 1713. Berlin: Springer, 1999.
- [14] Kuwert, E., A. Mondino and J. Schygulla. "Existence of immersed spheres minimizing curvature functionals in compact 3-manifolds." (2011): preprint arXiv:1111.4893.

- [15] Kuwert, E. and R. Schätzle. "Removability of isolated singularities of Willmore surfaces." *Annals of Mathematics* 160, no. 1 (2004): 315–57.
- [16] Lamm, T. and J. Metzger. "Small surfaces of Willmore type in Riemannian manifolds." *International Mathematics Research Notices* 2010, no. 19 (2010): 3786–813.
- [17] Lamm, T. and J. Metzger. "Minimizers of the Willmore functional with a small area constraint." *Annales de l'Institut Henri Poincaré. Analyse Non Linéaire* 30, no. 3 (2013): 497–518.
- [18] Lamm, T., J. Metzger, and F. Schulze. "Foliations of asymptotically flat manifolds by surfaces of Willmore type." *Mathematische Annalen* 350, no. 1 (2011): 1–78.
- [19] Lamm, T. and Nguyen, H. T. "Branched Willmore spheres." *Journal für die reine und angewandte Mathematik* (2011): preprint arXiv:1112.2877.
- [20] Li, H. and L. Vrancken. "New examples of Willmore surfaces in  $S^n$ ." *Annals of Global Analysis and Geometry* 23, no. 3 (2003): 205–25.
- [21] Li, P. and S. T. Yau. "A new conformal invariant and its applications to the Willmore conjecture and the first eigenvalue of compact surfaces." *Inventiones Mathematicae* 69, no. 2 (1982): 269–91.
- [22] Manzano, J. M. and R. Souam. "The classification of totally umbilical surfaces in homogeneous 3-manifolds." (2013): preprint arXiv:1302.2876.
- [23] Marques, F. and A. Neves. "Min-max theory and the Willmore conjecture." *Annals of Mathematics*, to appear.
- [24] Mondino, A. "Some results about the existence of critical points for the Willmore functional." *Mathematische Zeitschrift* 266, no. 3 (2010): 583–622.
- [25] Mondino, A. "The conformal Willmore Functional: a perturbative approach." *The Journal of Geometric Analysis* 23, no. 2 (2013): 764–811.
- [26] Mondino, A. and T. Rivière. "Immersed spheres of finite total curvature into manifolds." *Advances in Calculus of Variations*, to appear.
- [27] Mondino, A. and T. Rivière. "Willmore spheres in compact Riemannian manifolds." *Advances in Mathematics* 232, no. 1 (2013): 608–76.
- [28] Mondino, A. and J. Schygulla. "Existence of immersed spheres minimizing curvature functionals in non-compact 3-manifolds." *Annales de l'Institut Henri Poincaré. Analyse Non Linéaire*, to appear.
- [29] Petersen, P. *Riemannian Geometry*, 2nd ed. Graduate Texts in Mathematics 171. Berlin: Springer, 2006.
- [30] Ros, A. "The Willmore conjecture in the real projective space." *Mathematical Research Letters* 6, no. 5 (1999): 487–93.
- [31] Rivière, T. "Analysis aspects of Willmore surfaces." *Inventiones Mathematicae* 174, no. 1 (2008): 1–45.
- [32] Rivière, T. "Variational principles for immersed surfaces with  $L^2$ -bounded second fundamental form." *Journal für die Reine und Angewandte Mathematik (Crelle's Journal)*, to appear.

- [33] Simon, L. "Existence of surfaces minimizing the Willmore functional." *Communications in Analysis and Geometry* 1, no. 2 (1993): 281–325.
- [34] Souam, R. and E. Toubiana. "Totally umbilic surfaces in homogeneous 3-manifolds." *Commentarii Mathematici Helvetici* 84, no. 3 (2009): 673–704.
- [35] Urbano, F. "Minimal surfaces with low index in the three-dimensional sphere." *Proceedings of the American Mathematical Society* 108, no. 4 (1990): 989–92.
- [36] Wang, P. "On the Willmore functional of 2-tori in some product Riemannian manifolds." *Glasgow Mathematical Journal* 54, no. 3 (2012): 517–28.
- [37] Weiner, J. L. "On a problem of Chen, Willmore, et al." *Indiana University Mathematics Journal* 27, no. 1 (1978): 19–35.
- [38] Willmore, T. J. *Riemannian Geometry*. Oxford Science Publications. Oxford: Oxford University Press, 1993.
- [39] Ye, R. "Foliation by constant mean curvature spheres." *Pacific Journal of Mathematics* 147, no. 2 (1991): 381–96.