QUANTIZATION FOR BIHARMONIC MAPS FROM NON-COLLAPSED DEGENERATING EINSTEIN 4-MANIFOLDS

YOU MIN CHEN AND MIAOMIAO ZHU

Abstract. For a sequence of extrinsic or intrinsic biharmonic maps $u_j : M_j \to N$ from a sequence of non-collapsed degenerating closed Einstein 4-manifolds $(M_j, g_j)$ with bounded Einstein constants, bounded diameters and bounded $L^2$ curvature energy into a compact Riemannian manifold $(N, h)$ with uniformly bounded biharmonic energy, we establish a compactness theory for finitely many bubbles, which are finite energy biharmonic maps from $\mathbb{R}^4$, or from $\mathbb{R}^4/\Gamma$ for some nontrivial finite group $\Gamma \subset SO(4)$, or from some complete, noncompact, Ricci flat, non-flat ALE 4-manifold (orbifold). To achieve this, we develop a sophisticated asymptotic analysis for solutions over degenerating neck regions.

1. Introduction

There are many geometric variational problems which are critical in 4 dimensions and the moduli space of solutions of the Euler-Lagrange equations (in particular, the minimizers of the action functional) possess rich structures when the underlying domain manifold is of dimension 4. Careful investigation of the moduli space usually leads to profound geometric and topological applications, for instance, the theories of Yang-Mills fields (in particular, ASD connections) and Seiberg-Witten equations, etc.

In the present paper, we shall develop a new scheme in geometric analysis, to lay the ground for investigating the compactness of moduli space of solutions of certain geometric PDEs defined over 4-manifolds equipped with varying geometric structures. We shall consider a 4th order nonlinear geometric system that is critical in dimension 4, namely, the biharmonic map system, which is analytically complicated and subtle to handle. For the underlying domains, we shall consider non-collapsed closed Einstein 4-manifolds with varying metrics. The scheme developed in this paper can be applied to many other systems arising from geometry and physics, for instance, some lower order systems like Yang-Mills fields (including ASD connections) and Seiberg-Witten equations, which will be investigated in future works.

The theory of biharmonic functions dates back to 1862, when Maxwell and Airy studied it to describe a mathematical model of elasticity. In 3 dimensions, biharmonic functions can also be used to represent the velocity and pressure fields in the 3 dimensional Stokes flow (also called creeping flow), which is a type of fluid flow where advective inertial forces are small compared with viscous forces (namely, the Reynolds number is low), see e.g. [24]. Biharmonic maps between Riemannian manifolds are natural generalizations of biharmonic
functions and harmonic maps, and it is expected that exploration of the solution space of biharmonic maps should lead to some geometric and topological applications. Here we recall some basic notations and some geometric analysis aspects for biharmonic maps.

Let \((M, g)\) be a smooth Riemannian manifold and let \((N, h)\) be a compact smooth Riemannian manifold which is embedded in some Euclidean space \(\mathbb{R}^K\). Consider the following functionals for a map \(u\) from \((M, g)\) to \((N, h)\),

\[
H(u) = \frac{1}{4} \int_M |\Delta_g u|^2 dV_g
\]

and

\[
T(u) = \frac{1}{4} \int_M |\tau(u)|^2 dV_g.
\]

Here \(\tau(u)\) is the tension field of \(u\), or equivalently, the tangential part of \(\Delta_g u\). Critical points of \(H(u)\) are called extrinsic biharmonic maps, because the functional \(H(u)\) depends on the specific embedding of \((N, h)\) into the Euclidean space. Critical points of \(T(u)\) are called intrinsic biharmonic maps. The corresponding Euler-Lagrange equations are 4th order nonlinear elliptic systems which are critical in dimension 4 (see Section 2). When the target is an Euclidean space, then biharmonic maps reduce to biharmonic functions. In the sequel, if not specified, we shall mainly consider extrinsic biharmonic maps and omit the word extrinsic. The case of intrinsic biharmonic maps will be discussed in Section 7.

When the domain is of dimension 4, the above two functionals and the corresponding Euler-Lagrange equations are scaling invariant, therefore, in general, strong compactness for a sequence of solutions with uniformly bounded energy fails, due to the possible energy concentration at isolated points in the domain. To investigate the compactness issue, one needs to apply blow-up analysis and get some bubbling solutions, i.e., finite energy nontrivial biharmonic maps from \(\mathbb{R}^4\) to \(N\), by rescaling at the energy concentration points, called blow-up points. This is a well known phenomenon in geometric variational problems that are invariant under a non-compact group of local symmetries, dating back to the work [44].

In analogy to two dimensional harmonic maps which is conformally invariant, when the domain is Euclidean and of dimension 4, the blow-up theory for biharmonic maps have been developed via various methods e.g. [54, 25, 57, 58, 30, 33, 34]. More precisely, a bubble tree convergence (including energy identity and no neck property) hold for a sequence of such maps with uniformly bounded energy. This gives compactness modular bubbles, a phenomenon which appears also in many other geometric variational problems, for instance, 2 dimensional and 1st order problems like J-holomorphic curves e.g. [22], 2 dimensional and 2nd order problems like harmonic maps from surfaces e.g. [44, 26, 13, 40, 42, 32], 4 dimensional problems like Yang-Mills fields e.g. [14, 15, 17, 43].

If we allow the domain to vary, then in general, bubble tree convergence may not hold. For instance, in two dimensions, when the domain Riemann surface varies in a noncompact region of the moduli space and hence degenerates to surfaces with nodes, bubble tree convergence still holds for some 1st order problems like J-holomorphic curves e.g. [22, 60, 41, 61], however, for 2nd order problems like harmonic maps, in general, energy identity fails as energy may get lost on some degenerating necks, see the example in [40]. In fact, the bubble tree
convergence holds true only when we impose some extra constraints on the Pohožaev constant for harmonic maps over degenerating collar regions [62].

More precisely, we shall investigate the compactness problem for the moduli space of biharmonic maps from varying non-collapsed Einstein 4-manifolds to a fixed compact Riemannian manifold. Thanks to the well developed and beautiful compactness theory for non-collapsed Einstein 4-manifolds e.g. [1, 38, 4, 3, 39, 52, 2], we may consider a sequence of biharmonic maps \( u_j : M_j \to N \) from a sequence of non-collapsed closed Einstein 4-manifolds \( (M_j, g_j) \) with bounded Einstein constants, bounded diameters and bounded \( L^2 \) curvature energy to a compact Riemannian manifold \( (N, h) \) with uniformly bounded biharmonic energy.

It is well known that, on one hand, biharmonic maps can blow-up at energy concentration points in the domain and on the other hand, the underlying Einstein 4-manifolds can also blow-up at curvature concentration points and hence degenerate to a union of a Einstein 4-orbifold and some ALE (asymptotically locally Euclidean) bubble manifolds (orbifolds). To study the compactness problem, there are several difficulties we need to overcome. The first difficulty is that the bubble-neck decomposition for biharmonic maps from varying domains is somewhat subtle, since there is a bubble tree convergence of the underlying Einstein 4-manifolds. In other words, the blow-up of biharmonic maps and the blow-up of the domain manifolds are entangled with each other. The second difficulty is that the degeneration of Einstein 4-manifolds is accompanied by the blow-up of Riemannian curvatures, while previous blow-up theory and neck analysis for biharmonic maps are developed for the cases of Euclidean domains in \( \mathbb{R}^4 \) or standard flat cylinder domains \([a, b] \times S^3\) and hence can not be applied to the case of degenerating neck regions, where the geometric data are not explicitly known. More precisely, in order to analysis the behavior of biharmonic maps over the neck regions appeared in the bubble-neck decomposition for the degeneration of Einstein 4-manifolds, we need firstly to construct good global coordinates for those degenerate neck regions and take care of the behaviors of the varying metrics in these coordinate. Fortunately, inspired by the methods of constructing coordinates at the infinity in [4, 3], we succeed in constructing two types of good global coordinates for the degenerate neck region, which allow us to perform subtle neck analysis for biharmonic maps over the degenerating neck regions. In the following, we will sketch the bubble-neck decomposition procedure for the domain manifolds and briefly describe the geometry of the degenerating neck regions which look like portions of \( \mathbb{R}^4/\Gamma \), and leave the details to the main text.

To investigate our problem, we shall firstly extend the energy identity and no neck property for biharmonic maps from a domain in \( \mathbb{R}^4 \) to the case that the domain is a fixed general Riemannian 4-manifold. Differently from harmonic maps from Riemann surfaces which is conformally invariant and hence the uniformization theorem for Riemann surfaces can be applied to reduce the local blow-up analysis problem to the case of Euclidean domains in \( \mathbb{R}^2 \) and standard flat cylinder domains \([a, b] \times S^1\), the neck analysis for biharmonic maps from general curved 4-manifolds is more subtle. In [33], the authors attempted to deal with neck analysis for biharmonic maps from a curved space and carried out detailed analysis for the special case of \( S^4 \) with round metric. However, the case of general curved domain manifolds remains open.

For a general Riemannian 4-manifold \((M, g)\), we shall always take the normal coordinates \((x)\) at some point \( p = 0\), so that the scaling \( u(\lambda x) \) is well defined for small \( \lambda > 0\). For
simplicity of notations, it suffices to consider the case that there is only one blow-up point in the domain and there is only one bubble map, as the general case of multiple bubbles follows by induction.

**Theorem 1.1.** Let \( u_i \) be a sequence of extrinsic (or intrinsic) biharmonic maps from a geodesic ball \( B_1 \subset (M, g) \) to \( N \) with uniformly bounded biharmonic energy

\[
E(u_i, B_1) \equiv \int_{B_1} |\nabla^2_g u_i|^2 + |\nabla_g u_i|^4 dV_g \leq \Lambda,
\]

for some \( \Lambda > 0 \). Assume that there is a sequence of positive numbers \( \lambda_i \to 0 \) such that

\[
u_i(\lambda_i x) \to \omega,
\]
on any compact set \( K \subset \mathbb{R}^4 \), \( u_i \) converges weakly in \( W^{2,2} \) to some \( u_\infty : B_1 \subset (M, g) \to N \), and moreover, the nontrivial biharmonic map \( \omega : \mathbb{R}^4 \to N \) is the only bubble. Then, we have

**Energy identity:**

\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{i \to \infty} \int_{B_3 \setminus B_{\lambda_i R}} |\nabla^2_g u_i|^2 + |\nabla_g u_i|^4 dV_g = 0
\]

**No neck property:**

\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{i \to \infty} \text{osc}_{B_3 \setminus B_{\lambda_i R}} u_i = 0.
\]

To prove the above theorem, we can not use the idea of rewriting the biharmonic map equations with respect to the Riemannian metric \( g \) as some approximate biharmonic map type equations on a Euclidean domain. This is because if we view \( u \) as approximate biharmonic maps from some Euclidean domain and put the effect of the domain metric \( g \) into some error terms which involve fourth derivatives of \( u \), then we can not get uniform \( L^p(B_1) \) (for some \( p > 4/3 \)) control of these error terms when we deal with a blow-up sequence of biharmonic maps from \( (M, g) \). So we can not apply the classical Pohožaev type arguments for approximate biharmonic maps from Euclidean domains and with error terms uniformly bounded in \( L^p(B_1) \) for some \( p > 4/3 \), which are developed in the previous works [58, 33]. Therefore we need to make new observations and develop new methods.

A key step for the proof of the above theorem is to establish the following Pohožaev identity for biharmonic maps from Riemannian manifolds (see Subsection 3.2):

\[
(1) \quad \int_{\partial B_r(p)} (r\partial_r u)(\partial_r \triangle_g u) + r |\triangle_g u|^2 \frac{1}{2} - \partial_r (r\partial_r u) \triangle_g u \, d\sigma_g = P(g, r)
\]

where

\[
P(g, r) = -\int_{B_r(p)} (\triangle_g x^k) e_k u(\triangle_g u) - 2 \int_{B_r(p)} (\nabla_g x^k) \nabla_g e_k u \triangle_g u - |\triangle_g u|^2 dV_g + \int_{B_r(p)} (\text{div}_g(r\partial_r) - 4) \frac{|\triangle_g u|^2}{2} dV_g - \int_{B_r(p)} \text{Ric}_g(\nabla_g u, x^k e_k) \triangle_g u dV_g,
\]

\((x)\) are the normal coordinates, \( \{e_i(q)\}, i = 1, \cdots, 4 \), are the orthonormal vector fields obtained by parallel transport along the geodesic curve from \( p \) to \( q \), and \( r\partial_r = \sum_{i=1}^4 x^i e_i \). It is easy
to see that the error term \(P(g, r)\) in the right hand side of the above Pohožaev identity is caused by the non-flatness of the metric and it involves at most second derivatives of \(u\). Later we will see that it is of order \(O(r^2)\). This decay order is sufficient to carry out the neck analysis for the case of a fixed Riemannian domain in Theorem 1.1.

From the identity (1), we get
\[
\int_{\partial B_r(p)} (\Box_1 + \Box_2) d\sigma_g = \frac{P(g, r)}{r},
\]
where
\[
\Box_1 = -\frac{1}{2} (\partial_r^2 u)^2 + \left(\frac{\tilde{m}^2}{2r^2}\right) (\partial_r u)^2 + \frac{1}{2r^4} (\tilde{\Delta} u)^2 + \left(\frac{\tilde{m}}{r^3}\right) (\partial_r u)(\tilde{\Delta} u),
\]
and
\[
\Box_2 = (\partial_r u) (\partial_r^2 u) + \frac{\tilde{m} - 1}{r} (\partial_r u) (\partial_r^2 u) + \left(\frac{\partial_r \tilde{m}}{r} - \frac{2\tilde{m}}{r^2}\right) (\partial_r u)^2 + \frac{1}{r^2} (\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3} (\partial_r u)(\tilde{\Delta} u).
\]
Here \(\tilde{\Delta}\) is the associated Laplace operator of \(\tilde{\partial} B_r(p)\), \(\tilde{\partial} B_r(p)\) is the sphere \(\tilde{\partial} B_r(p)\) with the metric \(g_r\), and \(g_r\) is the metric of \(\tilde{\partial} B_r(p)\) induced from \((M, g)\). The above identity together with some further observations then reveals the relation between the tangential part energy and the radial part energy.

Now, we turn to the most challenging part of the present paper, namely, the case of domain manifolds with varying geometric structures. It turns out that when domain manifolds are equipped with degenerating Einstein metrics, the situation becomes much more complicated and difficult to handle. In fact, we need to take extra care of the decay orders of both the maps and the metrics over degenerating neck regions involved in each step, which requires a sophisticated asymptotic analysis.

Let \((M_i, g_i)\) be a sequence of closed 4-dimensional Einstein manifolds with uniformly bounded Einstein constants \(\{\mu_i\}\) and satisfying
\[
diam(M_i, g_i) \leq D, \ vol(M_i, g_i) \geq V \text{ and } \int_{M_i} |R_{g_i}|^2 dV_{g_i} \leq R
\]
for some positive constants \(D > 0, V > 0\) and \(R > 0\), where we denote the curvature tensor of a metric \(g\) by \(R_g\). Let \(u_i\) be a sequence of extrinsic (or intrinsic) biharmonic maps from \((M_i, g_i)\) to \(N\) and satisfies the following uniform biharmonic energy bound condition:
\[
E(u_i, M_i) \equiv \int_{M_i} |\nabla_{g_i} u_i|^2 + |\nabla_{g_i} u_i|^4 dV_{g_i} \leq \Lambda,
\]
for some \(\Lambda > 0\).

By the compactness theory of non-collapsed Einstein 4-manifold, we can get bubble tree convergence for the underlying Einstein 4-manifolds (up to a subsequence). For simplicity of the statement, we assume at the moment that there is only one curvature blow-up point in the domain (i.e., the curvatures of \(M_i\) blow-up at only one point as \(i \to \infty\)), and that there
is only one bubble manifold, which is a Ricci flat, asymptotically locally Euclidean (ALE) manifold. The case of multiple bubble manifolds (orbifolds) will be discussed at the end of Section 4.

Suppose that \( M_j \) blow up at \( x_{a,j} \), and for any \( \delta > 0 \),

\[
(B(x_{a,j}, \delta), g_j) \subset M_j \to (B(x_a, \delta), g_\infty) \subset M_\infty
\]

in the sense of Gromov-Hausdorff convergence. Here \( M_\infty \) is an orbifold with isolated singularity \( x_a \). Then there is a sequence of \( r_j \to 0 \) such that \((M_j, r_j^{-2}g_j, x_{a,j})\) converge to \((M_a, h_a, x_{a,\infty})\) in the pointed Gromov-Hausdorff distance, where \((M_a, h_a)\) is a complete, noncompact, Ricci flat, non-flat ALE 4-manifold (4-orbifold). And the degenerate neck region \( B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R) \) with \( \delta \) being very small and \( R \) being very large is close to a portion of a flat cone \( \mathbb{R}^4/\Gamma \) for some discrete group \( \Gamma \subset SO(4) \), for large \( j \).

On the thick part of the domain manifold, which is away from the curvature blow-up point \( x_{a,j} \), we have nice convergence of the metrics (up to a subsequence) and hence the energy identity and the no neck property can be proved by slightly modifying the argument for the proof of Theorem 1.1, see Subsection 4.1.

At the curvature blow-up point \( x_{a,j} \), the situation becomes much more complicated. The blow-up of the maps \( u_j \) can occur inside the bubble manifold region \( B(x_{a,j}, r_j R) \), or inside the degenerate neck region \( B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R) \). In the former case, the blow-up scale can be some sequence \( \lambda_j^A \ll r_j \) (i.e., \( \lim_{j \to \infty} \lambda_j^A r_j = 0 \)) or \( r_j \). For the case of \( \lambda_j^A \ll r_j \) (assume that there is only one bubble map in this case, i.e. \( u_j \) blow up at only one scale), we have a bubble map

\[
\omega^A : \mathbb{R}^4 \to N.
\]

For the case of \( \{r_j\} \), we have a bubble map

\[
\omega^B : M_a \to N.
\]

In the degenerate neck region \( B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R) \), multiple bubble maps of different scales can possibly occur. Without loss of generality, suppose that there is only one bubble map in this case, then it takes the form

\[
\omega^C : \mathbb{R}^4/\Gamma \to N.
\]

Thus, to understand the compactness problem, we are left with two questions concerning the limiting behavior of the maps \( u_j \) over the following region

\[
(B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R)) \setminus B(x_{a,j}, \lambda_j^C R),
\]

where \( B(x_{a,j}^C, \lambda_j^C R) \) is the bubble map domain corresponding to \( \omega^C \). The first question is whether there is any energy of the maps lost over this region. The second question is whether the image of the maps over this region is converging (up to a subsequence) to a point and hence the weak limit of the maps and the bubbles are all connected.

In this paper, we shall give affirmative answers to these two questions and hence establish a compactness theory modular finitely many bubbles for biharmonic maps from non-collapsed Einstein 4-manifolds with varying metrics. Now we state our main result in this paper.
Theorem 1.2. Let \((M_j, g_j), N\) and \(u_j\) be as above. Without loss of generality, we assume that there is only one ALE bubble manifold and there is at most one bubble map in each case in the blow-up scheme discussed above, then we have (up to a subsequence)

**Energy identity:**

\[
\int_{B(x_{a,j}, \delta)} |\nabla_{g_j}^2 u_j|^2 + |\nabla_{g_j} u_j|^4 dV_{g_j} = \int_{B(x_{a,\delta_0})} |\nabla_{g_\infty}^2 u_\infty|^2 + |\nabla_{g_\infty} u_\infty|^4 dV_{g_\infty} + \sum_{a=A,B,C} E(\omega^a),
\]

where \(E(\cdot)\) is the biharmonic energy.

**No neck property:** The image of \(u_\infty\) and the images of all bubbles are connected, namely, both the distance between the images of any two connected bubbles and the distance between the image of \(u_\infty\) and the image of any bubble which is connected to it in \(N\) are zero.

The general case of multiple bubble manifolds and multiple bubble maps follows from induction arguments. This gives bubble tree convergence (including possibly finitely many ghost bubbles) for biharmonic maps from non-collapsed degenerating Einstein 4-manifolds. We remark that the biharmonic maps from Ricci flat ALE manifolds (orbifolds) may be trivial, since the number of nontrivial Ricci flat ALE bubble manifolds (orbifolds) is finite, so the induction process must be terminated in finite number of steps.

When the target \(N\) has trivial fourth fundamental group, namely, \(\pi_4(N) = \{e\}\), then it is easy to see that the two types of bubble maps \(\omega^A : \mathbb{R}^4 \to N\) and \(\omega^C : \mathbb{R}^4/\Gamma \to N\) have to be trivial, therefore, the maps \(u_j\) blow up only if the underlying Einstein 4-manifolds \((M_j, g_j)\) blow up. Note that the existence of biharmonic maps in each free homotopy class in \([M, N]\) is ensured by the method of heat flow in [56, 18] when \(\pi_4(N) = \{e\}\). In this case, it is interesting to know whether there are finite energy nontrivial biharmonic maps from a nontrivial Ricci flat ALE manifold (orbifold) \(V\) to \(N\).

To prove Theorem 1.2, we shall develop three circle type method and establish new Pohožaev type identity for biharmonic maps from non-collapsed degenerating Einstein 4-manifolds. Recall that three circle type method is an important and powerful technique in geometric analysis since the work [47]. For biharmonic maps from 4 dimensional Euclidean domains, such kind of method was developed in [33, 34] to get the decay estimates for the tangential part energy. Pohožaev type identities (or arguments) for biharmonic maps from a fixed domain in \(\mathbb{R}^4\) were derived in [54, 25, 57, 58, 33, 34, 30]. One can control the radial part energy by the tangential part energy in the neck analysis by detailed analysis of the relation between the radial part energy by the tangential part energy embodied in these identities. We remark that this relation is not as obvious as in the harmonic maps case.

However, when domains are degenerating Einstein 4-manifolds where the geometric data of the degenerating neck regions are not explicitly known, the situation becomes much more complicated and subtle, and there are several new difficulties we need to overcome. The first one is that we have to construct good global coordinates on the whole degenerating neck region \(B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R)\) in which the metric \(g_j\) is sufficiently regular for the neck analysis of biharmonic maps in order to prove Theorem 1.2. The second one is that the
non-flat metrics can effect the arguments throughout the proof, so we need to take care of the error terms appearing in each step.

To carry out delicate analysis on the degenerate neck region $B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R)$, we succeed in constructing two types of good global coordinates on this region: $(x)$ and $(y)$, see Section 5. In the coordinates $(x)$, $g_j$ is $C^0$ close to the flat metric. While, in the coordinates $(y)$, $g_j$ is $C^4$ close to the flat metric in some weighted function space. To develop the three circle type method for biharmonic maps over the degenerating neck region, we have to use the coordinate $(y)$. While for the Pohožaev type argument, we have to use the $(x)$ coordinates which share a lot of similarities with the usual normal coordinates at a point. Therefore, to complete the proofs, we need to take advantage of both two coordinate systems.

To prove the degenerating case in Theorem 1.2, we develop a new Pohožaev type argument which is more general than the case of a fixed Riemannian domain give in the proof of Theorem 1.1 and the cases of Euclidean domains derived in previous works. More precisely, we derive a new Pohožaev type argument in a more geometric way and it depends much less on the coordinate functions than the previous ones developed for a fixed domain with local geometric data being explicitly given and regular enough. In fact, to deal with the degenerating case, we have to treat it in this way, since on one hand the coordinate functions $x^k$ are not regular enough for the analysis, and on the other hand the geodesic spheres of radius $r$ are not equal to the coordinates spheres given by $|y| = r$. We believe that our methods are also of wider interest in the field.

Here, we briefly describe our argument, see Subsection 6.2 for more details. The Pohožaev identity used in the degenerating neck region takes the following form:

$$\int_{\partial A_{j,t,ct}} (r \partial_r u_j)(\partial_r \Delta_{g_j} u_j) + r \frac{|\Delta_{g_j} u_j|^2}{2} - \partial_r (r \partial_r u_j) \Delta_{g_j} u_j d\sigma_{g_j} = P(g_j, t),$$

where $r = |x| = d_{g_j}(\cdot, x_{a,j})$, $P(g_j, t)$ is an error term depending on the metric $g_j$, and the bi-harmonic energy of $u_j$ over a portion of the degenerating neck region, denoted by $A_{j,t,ct} \equiv B(x_{a,j}, ct) \setminus B(x_{a,j}, t)$. Then it follows that

$$c \int_{\partial B(x_{a,j}, ct)} \Box_1 + \Box_2 d\sigma_{g_j} - \int_{\partial B(x_{a,j}, t)} \Box_1 + \Box_2 d\sigma_{g_j} = \frac{P(g_j, t)}{t},$$

where $\Box_1, \Box_2$ have the same form as above. Later in Section 6, we will see that the following form

$$cA(ct) - A(t) = \frac{P(g_j, t)}{t}, \quad (c > 1)$$

is crucial to prove Theorem 1.2 even though we do not have an identity of the form (1).

A crucial and subtle fact in our argument is to show that the error term $P(g_j, t)$ has the same order as the closeness of $g_j$ to the flat metric, see Lemma 6.7. Another difficulty in our Pohožaev type argument used in the proof of Theorem 1.2 is that we can only work on some sub-annular regions with nice geometric pictures, but not on the whole degenerating neck region $B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R)$. This makes the whole argument more technical and also more interesting.
We remark that the geometry of the degenerating neck domains plays a vital role in the paper, especially the property
\[ |g_{j,kl} - \delta_{kl}| < \eta_j(r) \]
on \( A_{j;r_j R, \delta} \equiv B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R) \) in the coordinates \((x)\) and \((y)\), where
\[ \eta_j(r) = O\left(r_{\infty}^{\varepsilon_5} r^{\varepsilon_5} + (r_j R_0)^{\varepsilon_5} r^{-\varepsilon_5}\right), \]
\( r_{\infty} > 0, \varepsilon_5 > 0, r_j > 0 \) are the same as in the curvature estimate in Proposition 2.6, and \( R_0 > 0 \) is some fixed large number. The arguments in this proof reveal that Theorem 1.2 holds because of the important fact that there is no curvature concentration on the degenerating neck regions in the tree of ALE bubble manifolds (orbifolds) \( Tr_{ALE} \), which is a direct consequence of Proposition 2.6 (see Section 2).

The paper is organized as follows. In Section 2, we recall some background knowledge about the bubble tree convergence of non-collapsed degenerating Einstein manifolds and some key results for the blow-up analysis of biharmonic maps. In Section 3, we prove Theorem 1.1 (for extrinsic biharmonic maps). In Section 4, we do the bubble-neck decomposition. In Section 5, we construct the two types of good coordinates. In Section 6, we prove Theorem 1.2 (for extrinsic biharmonic maps). In Section 7, we deal with the case of intrinsic biharmonic maps.

2. Preliminaries

2.1. Bubble tree convergence of non-collapsed Einstein manifolds. We recall the compactness theory of non-collapsed Einstein metrics developed in [1, 38, 4, 3, 52, 39, 2] etc. The definitions of (pointed) Gromov-Hausdorff convergence, Riemannian orbifold, and ALE (asymptotically locally Euclidean) manifold can be found in the above mentioned works and references therein. For more details about G-H convergence, we refer to Chapter 7 of [6]. Roughly speaking, a 4-dimensional orbifold locally looks like \( \mathbb{R}^4/\Gamma_0 \) for some finite group \( \Gamma_0 \subset SO(4) \) at the isolated singularities, an ALE 4-manifold is a complete noncompact manifold which is close to \( \mathbb{R}^4/\Gamma_\infty \) at the infinity for some finite group \( \Gamma_\infty \subset SO(4) \). We remark here that the results collected in this subsection in fact hold for all dimensions \( n \geq 4 \), as along as we replace the \( L^2 \) curvature condition given in Theorem 2.1 by the following
\[ \int_{M_i} |R_{g_i}|^2 dV_{g_i} \leq R. \]

**Theorem 2.1.** ([1, 38, 4]) Let \((M_i, g_i)\) be a sequence of 4-dimensional smooth manifolds and the Einstein metrics on them with uniformly bounded Einstein constants \( \{\mu_i\} \) satisfying
\[ \text{diam}(M_i, g_i) \leq D, \ vol(M_i, g_i) \geq V \text{ and } \int_{M_i} |R_{g_i}|^2 dV_{g_i} \leq R \]
for some positive constants \( D, V \) and \( R \), where the curvature tensor of a metric \( g \) is denoted by \( R_g \). Then there exist a subsequence \( \{j\} \subset \{i\} \) and a compact Einstein orbifold \((M_\infty, G_\infty)\) with a finite set (possibly empty) of orbifold singular points \( S = \{x_1, x_2, \cdots, x_s\} \subset M_\infty \) for which the following statements hold:
(1) \((M_j, g_j)\) converges to \((M_\infty, g_\infty)\) in the Gromov-Hausdorff distance.

(2) There exists an into diffeomorphism \(F_j : M_\infty \to M_j\) for each \(j\) such that \(F_j^* g_j\) converges to \(g_\infty\) in the \(C^\infty\)-topology on \(M_\infty \setminus S\).

(3) For every \(x_a \in S\) (\(a = 1, 2, \cdots, s\)) and \(j\), there exists \(x_{a,j} \in M_j\) and a positive number \(r_j\) such that

3.a) \(B(x_{a,j}, \delta)\) converges to \(B(x_a, \delta)\) in the Gromov-Hausdorff distance for all \(\delta > 0\).

3.b) \(\lim_{j \to \infty} r_j = 0\)

3.c) \((M_j, r_j^{-2} g_j, x_{a,j})\) converge to \((M_a, h_a, x_{a,\infty})\) in the pointed Gromov-Hausdorff distance, where \((M_a, h_a)\) is a complete, noncompact, Ricci flat, non-flat 4-manifold which is ALE of order 4.

3.d) There exists an into diffeomorphism \(G_j : M_a \to M_j\) such that \(G_j^* (r_j^{-2} g_j)\) converges to \(h_a\) in the \(C^\infty\)-topology on \(M_a\).

(4) It holds that

\[
\lim_{j \to \infty} \int_{M_j} |R_{g_j}|^2 dv_j \geq \int_{M_\infty} |R_{g_\infty}|^2 dv_\infty + \sum_a \int_{M_a} |R_{h_a}|^2 dv_{h_a}.
\]

Remark 2.2. Cheeger-Naber showed that non-collapsed 4-manifolds with bounded Ricci curvature have a priori \(L^2\) Riemannian curvature estimates, see [10].

Remark 2.3. By Theorem (5.1) of [4], the metric \(g_\infty\) on \(M_\infty\) extends smoothly across \(x_a\) as an orbifold Einstein metric. That is to say there is a (covering) map \(\Pi_a : B_1 \setminus \{0\} \to B(x_a, \delta) \setminus \{x_a\}\) such that \(\Pi_a^* g_\infty\) extends to a smooth Einstein metric on \(B_1 \subset \mathbb{R}^4\).

Later, the bubble tree convergence was proved in [3], showing that there is no curvature concentration in the neck regions (see also [39]). Here for the sake of convenience, we state these results using notations in [3]. Now we take a positive constant \(r_\infty\) sufficiently small so that for all

\[
\sup_{B(x_{a,j}, r_\infty)} |R_{g_j}|^2 = |R_{g_j}|^2(x_{a,j}) \to \infty \quad \text{as} \quad j \to \infty
\]

and

\[
\int_{B(x_a, r_\infty)} |R_{g_\infty}|^2 \leq \frac{\varepsilon}{2}
\]

with a small positive number \(\varepsilon \leq \varepsilon_4/2\) determined by the argument in [3], where \(\varepsilon_4\) is the number in Proposition 1 of [3] (a small energy regularity theorem). And for sufficiently large \(j\) we can find a positive number \(r_j\) so that

\[
\int_{B(x_{a,j}, r_\infty) \setminus B(x_{a,j}, r_j)} |R_{g_j}|^2 = \varepsilon.
\]

It is easy to see that

\(r_j \to 0\) \quad \text{as} \quad j \to \infty.

Proposition 2.4. ([3]) There is a subsequence \(\{k\} \subset \{j\}\) (denoted still by \(\{j\}\) ), such that \((M_j, r_j^{-2} g_j, x_{a,j})\) converges to \(((Y, h), y_\infty)\) (we call it an ALE bubble) in the pointed G-H distance, where \((Y, h)\) is a complete, non-compact, Ricci flat, non-flat ALE 4-manifold (orbifold) of order 4 with only finitely many isolated singular points. The convergence is actually smooth except at the singular points.
Remark 2.5. By repeating the same process, we can obtain a bubble tree, for more details we refer to Page 211 of [3]. In the following we shall call it ALE-tree, and denote it by $\text{Tr}_{\text{ALE}}$. And each one of the Ricci flat ALE manifolds (orbifolds) in the tree has only one end by the splitting theorems [3, 9] and the arguments in Theorem 3.5 of [1].

Proposition 2.6. (Proposition 3 and Proposition 4 in [3]) There exist positive constants $C_7 > 0$ and $\varepsilon_5 > 0$ such that for $4r_j \leq r < 4r \leq r_\infty$ it holds that

$$r^2|R_{g_j}| \leq C_7 \max \left\{ \frac{(r_j)^{\varepsilon_5}}{r}, \frac{(r_\infty)^{\varepsilon_5}}{r} \right\}.$$ 

And if one takes $1 < K_1 < K_2$ sufficiently large, then the subset $B(x_{a,j}, K_2^{-1}r_\infty) \setminus B(x_{a,j}, K_1r_j)$ (called the degenerate neck region) is close to a portion of some flat cone $\mathbb{R}^4/\Gamma$ for large $j$.

Remark 2.7. It was noticed in [3] that there exist certain coordinates on the degenerate neck region. However, no details of construction of coordinates were given and little was known about the properties of such coordinates before. In order to carry out refined analysis for solutions of geometric PDEs defined over degenerating domains, we need to firstly construct good coordinates on these neck regions. By adapting the methods in [4], we succeed in constructing coordinates on the degenerate neck region (see Theorem 5.3)

$$A_{K_1r_j, K_2^{-1}r_\infty} \equiv B(x_{a,j}, K_2^{-1}r_\infty) \setminus B(x_{a,j}, K_1r_j),$$

in which $g_j$ is close to the standard flat metric in certain sense. It indicates that the finite group $\Gamma$ of the neck is equal to the fundamental group at the infinity of the ALE orbifold $Y$ and the local fundamental group of $B(x_a, r_\infty)$.

Notice that $\varepsilon_5 > 0$, a direct consequence of Proposition 2.6 is a curvature energy identity of the following type, one can find more specific statements in Theorem 2.5 of [39].

Corollary 2.8. Let $r_j, r_\infty, x_{a,j}, K_1, K_2$ be as above, then it holds

$$\lim_{K_2 \to \infty} \lim_{K_1 \to \infty} \lim_{r_j \to 0} \int_{B(x_{a,j}, K_2^{-1}r_\infty) \setminus B(x_{a,j}, K_1r_j)} |R_{g_j}|^2 dV_j = 0.$$ 

2.2. Biharmonic maps. Recall that extrinsic and intrinsic biharmonic maps are the critical points of the following two functionals

$$H(u) = \frac{1}{4} \int_M |\Delta_g u|^2 dV_g$$

and

$$T(u) = \frac{1}{4} \int_M |\tau(u)|^2 dV_g.$$ 

The Euler-Lagrange equation for $H(u)$ is

$$\Delta_g^2 u = \Delta_g u (B(u)(\nabla_g u, \nabla_g u)) + 2 \nabla_g \cdot (\Delta_g u, \nabla_g (P(u))) - \langle \Delta_g (P(u)), \Delta_g u \rangle.$$

Here $B$ is the second fundamental form of $N \subset \mathbb{R}^K$ and $P(u)$ is the projection to the tangent space $T_u N$. The Euler-Lagrange equation for $T(u)$ is more complicated, we refer to [34]. In this paper we shall mainly focus on extrinsic biharmonic maps, the case of intrinsic biharmonic maps will be handled in Section 7.
The analytic aspects of biharmonic maps have been studied extensively in recent decades. The regularity theory of biharmonic maps are developed in [7, 51, 53, 55, 54, 36, 29, 50, 45] etc. The blow-up theory (including energy identity and the no neck property) of biharmonic maps from domains in $\mathbb{R}^4$ had been studied in [54, 25, 30, 57, 58, 34, 33] etc.

Now we recall some basic analytical tools for the blow-up analysis for biharmonic maps in dimension 4. In fact, these basic tools can be easily extended to the case of a Riemannian domain.

First of all, there is a small energy regularity theorem for biharmonic maps.

**Theorem 2.9.** ($\varepsilon$-regularity) Let $(B_1, g)$ be a ball in $\mathbb{R}^4$ equipped with a Riemannian metric $g$, for any $p > 1$, there exists $\varepsilon_0 > 0$ and $C_p > 0$ such that if $u \in W^{2,2}(B_1, \mathbb{R}^K)$ is an extrinsic (or intrinsic) biharmonic map into $N$ satisfying

$$\int_{B_1} |\nabla^2_g u|^2 + |\nabla_g u|^4 dV_g \leq \varepsilon_0,$$

then

$$\|u - \bar{u}\|_{W^{4,p}(B_{1/2})} \leq C_p (\|\nabla^2 u\|_{L^2(B_1)} + \|\nabla u\|_{L^2(B_1)}),$$

where $\bar{u}$ is the mean value of $u$ over the unit ball.

**Proof.** When the metric $g$ is Euclidean, the corresponding result has been proved in previous works, for example, [58, 30]. It is easy to see that the arguments in [30] remain true when we consider biharmonic maps on balls equipped with smooth Riemannian metrics. It follows that (4) holds. \qed

**Remark 2.10.** There is no need to distinguish the $L^p$ norm with respect to $g$ and the $L^p$ norm with respect to the Euclidean metric for our purpose, as they are locally equivalent to each other.

**Corollary 2.11.** Let $\varepsilon_0 > 0$ be as in Theorem 2.9, if

$$\int_{B_r} |\nabla^2_g u|^2 + |\nabla_g u|^4 dV_g \leq \varepsilon_0,$$

then for $l = 1, 2, \cdots$,

$$\sup_{B_{r/2}} |\nabla^l_g u| \leq C_{\varepsilon_0}^{1/2}.\frac{1}{r^l}.$$

It is easy to see that the regularity theory for weakly biharmonic maps from domains in $\mathbb{R}^4$ developed in e.g. [53, 29] can be extended to the case of a Riemannian domain $B_1 \subset (M^4, g)$. Moreover, one can apply an argument as Lemma 2.5 in [25] to show that a biharmonic map which is smooth in $B_1 \setminus \{0\} \subset (M^4, g)$ and with finite energy is in fact a weakly biharmonic map over the whole domain. Therefore, the removable singularity theorem of extrinsic (or intrinsic) biharmonic maps follows as a corollary.

**Theorem 2.12.** (removable singularity) Let $u$ be a smooth biharmonic map on $B_1 \setminus \{0\} \subset (M^4, g)$, if

$$\int_{B_1} |\nabla^2_g u|^2 + |\nabla_g u|^4 dV_g < \infty,$$
then $u$ can be extended to a smooth biharmonic map on the whole $B_1$.

**Remark 2.13.** Usually, removable singularity theorems can be proved by applying Pohožaev type arguments as in [44]. See for example the proof of removable singularity theorem for biharmonic maps in Section 6 of [34] for the case of Euclidean domains. In fact, one can modify the method in [34] to give a different proof of Theorem 2.12. We will use similar arguments to show that the oscillation of a biharmonic map at the infinity is arbitrarily small (see Lemma 4.8) in Appendix C.

**Theorem 2.14.** (energy gap) Let $u$ be a biharmonic map from $\mathbb{R}^4/\Gamma$ ($\Gamma$ is some finite group in $SO(4)$) to $N$. There exists $\varepsilon(N, \Gamma) > 0$ depending only on $N$ and $\Gamma$, such that if $u$ satisfies

\begin{equation}
\int_{\mathbb{R}^4/\Gamma} |\nabla^2_g u|^2 + |\nabla_g u|^4 dV_g \leq \varepsilon(N, \Gamma),
\end{equation}

then $u$ is a constant map.

**Proof.** For $\Gamma = \{e\}$, it is a direct consequence of Theorem 2.11. For a nontrivial group $\Gamma$, we can lift the map $u$ to be a map $\tilde{u} : \mathbb{R}^4 \setminus \{0\} \to N$. And then by Theorem 2.12, it can be extended to be a biharmonic map from $\mathbb{R}^4$ to $N$. We may take $\varepsilon(N, \Gamma) = \varepsilon_0/|\Gamma|$, where $|\Gamma|$ is the number of elements in $\Gamma$. $\square$

With the above analytical tools in hand, now we are ready to analyse the blow-up procedure for biharmonic maps from Riemannian manifolds. We refer to [30, 58] for the bubble-neck decomposition.

For a fixed Riemannian manifold $(M, g)$, we always take the normal coordinates $x$, so that the scaling $u(\lambda x)$ is well defined for small $\lambda > 0$. By Gauss Lemma, for small $r > 0$ the geodesic ball $B_r$ is the same as the ball of radius $r$ in normal coordinates.

Now we state the energy identity and no neck property theorem for biharmonic maps defined on general Riemannian manifolds.

**Theorem 2.15.** (Energy identity and no neck property) Let $u_i$ be a sequence of extrinsic biharmonic maps from $B_1 \subset (M, g)$ to $N$ satisfying

\[
E(u_i, B_1) \equiv \int_{B_1} |\nabla^2_g u_i|^2 + |\nabla_g u_i|^4 dV_g \leq \Lambda,
\]

for some $\Lambda > 0$. Assume that there is a sequence positive $\lambda_i \to 0$ such that

$u_i(\lambda_i x) \to \omega$,

on any compact set $K \subset \mathbb{R}^4$, $u_i$ converges weakly in $W^{2,2}$ to $u_\infty$, and moreover $\omega$ is the only bubble. Then,

\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{i \to \infty} \int_{B_\delta \setminus B_{\lambda_i R}} |\nabla^2_g u_i|^2 + |\nabla_g u_i|^4 dV_g = 0
\]

and

\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{i \to \infty} \text{osc}_{B_\delta \setminus B_{\lambda_i R}} u_i = 0.
\]

**Remark 2.16.** By arguments in Section 7, we know that the same results hold for intrinsic biharmonic maps.

A simple consequence of Theorem 2.15 is that the following limit
\[
\lim_{|x| \to \infty} \omega(x)
\]
exists (see the end of Section 2 of [34]). In this theorem, we assume that there is only one bubble. The general case of multiple bubble maps can be handled by induction.

In order to prove Theorem 2.15, we shall extend the methods developed for the case of Euclidean domains in [30, 34, 33] to a more general setting. There are two main ingredients in [34, 33], one is a three circle theorem for approximate biharmonic functions which is used to get the exponential decay for the tangential part energy on the neck regions, the other is a Pohožaev type argument. The main difficulty is to establish a Pohožaev type identity for biharmonic maps defined on geodesic balls in general Riemannian manifolds. Fortunately, the Pohožaev type identities on Euclidean domains derived in [25, 58, 30] provides the key clue. We shall put the proof in the next section. We remark that the proof have something in common with the proof for the uniqueness of tangent cone of (bi)harmonic maps [47, 48, 12]. Both of them need to prove some decay estimates for the maps by subtle analysis of the structure of equations.

3. Neck analysis on a fixed Riemannian manifold

In this section, we shall prove the energy identity and no neck property in Theorem 2.15 by establishing a Pohožaev identity for biharmonic maps on general Riemannian manifolds.

3.1. Decay of the tangential part energy. For the sake of readers’ convenience, we shall firstly recall some results in [33, 34]. We start from the definition of η-approximate biharmonic functions.

Definition 3.1. (Definition 1 of [33]) Let \( u \) be a smooth function defined on \( B_{r_2} \setminus B_{r_1} \), \( u \) is called an \( \eta \)-approximate biharmonic function, if it satisfies
\[
\Delta^2_g u(r, \theta) = a_1 \nabla_g \Delta_g u + a_2 \nabla^2_g u + a_3 \nabla_g u + a_4 u \\
+ \frac{1}{|\partial B_r|} \int_{\partial B_r} b_1 \nabla_g \Delta_g u + b_2 \nabla^2_g u + b_3 \nabla_g u + b_4 u \, d\sigma + h(x),
\]
where \( a_i \), \( b_i \) and \( h \) are smooth functions satisfying the followings: for any \( \rho \in [r_1, r_2/2] \),
(a) \( ||g_{ij}(\rho x) - \delta_{ij}||_{C^4(B_\rho \setminus B_1)} < \eta \). Namely, the metric after scaling to \( B_2 \setminus B_1 \) is close to the flat metric in \( C^4 \) norm.
(b) \( |||x|^{4(1-1/p)}h||_{L^p(B_{r_2} \setminus B_{r_1})} \leq \eta \)
(c) \( \sum_{i=1}^{4} ||a_i||_{\tilde{W}^{4-i,p}(B_{2\rho} \setminus B_\rho)} + ||b_1||_{\tilde{W}^{4-i,p}(B_{2\rho} \setminus B_\rho)} \leq \eta \).

Here the \( \tilde{W}^{4-i,p} \) norm is defined by
\[
|||w||_{\tilde{W}^{4-i,p}(B_{2\rho} \setminus B_\rho)} = |||\rho^i w(\rho x)||_{W^{4-i,p}(B_2 \setminus B_1)}|
\]
One can check that if \( u \) is an \( \eta \)-approximate biharmonic function on \( B_{r_2} \setminus B_{r_1} \), then \( w(x) = u(x/\lambda) \) is another \( \eta \)-approximate biharmonic function on \( B_{\lambda r_2} \setminus B_{\lambda r_1} \).

The next is an interior \( L^p \) estimate for approximate biharmonic functions.

**Lemma 3.2.** (Lemma 1 of [33]) Suppose that \( u : B_4 \setminus B_1 \rightarrow \mathbb{R} \) is an \( \eta \)-approximate biharmonic function (for small \( \eta > 0 \)) with

\[
\sum_{i=1}^{4} ||a_i||_{W^{4-i,p}(B_4\setminus B_1)} + ||b_i||_{W^{4-i,p}(B_4\setminus B_1)} \leq \eta.
\]

and

\[
||h||_{L^p(B_4\setminus B_1)} \leq C.
\]

Then, for any \( p > 1 \), we have

\[
||u||_{W^{4,p}(B_3\setminus B_2)} \leq C(||u||_{L^p(B_4\setminus B_1)} + ||h||_{L^p(B_4\setminus B_1)}).
\]

Now we study the behaviour of the biharmonic map (with respect to the metric \( g \)) \( u_i \) on the neck region \( \Sigma = B_{\delta} \setminus B_{\lambda R} \) as in Section 4 of [34]. Assume without loss of generality that

\[ \Sigma = \bigcup_{l=_{0}}^{l_i} A_l, \]

for \( A_l = B_{e^{-(l-1)L}} \setminus B_{e^{-lL}} \) and \( l_0 < l_i \). Moreover we may assume as in [34] that for any \( \varepsilon > 0 \),

\[
\int_{A_l} |\nabla^2 g u_i|^2 + |\nabla g u_i|^4 dV_g < \varepsilon^2 < \varepsilon_0
\]

holds for \( l = l_0, \ldots, l_i \), when \( i \) and \( R \) are large enough and \( \delta > 0 \) is small enough. In the above, \( \log \delta = -l_0 L \) and \( \log \lambda_i R = -l_i L \).

Our aim in this subsection is to prove the following exponential decay estimates:

**Lemma 3.3.** After setting \( r = e^t \), we have for arbitrarily small \( \varepsilon > 0 \), if \( \delta > 0 \) is sufficiently small and \( i \) and \( R \) are large enough, then

\[
\int_{(-l,L)-(l-1)L)} (|\Delta g^3 u_i|^2 + |\nabla g^3 u_i|^4 + |\partial_t \nabla g^3 u_i|^2 + |\partial_t \Delta g^3 u_i|^2) \, dt \, d\theta \leq C \varepsilon^2 (e^{-(l-1)L} + e^{-((l_i-1)L)}).
\]

Or equivalently,

\[
\int_{B_{\delta \lambda_i R} \setminus B_r} \frac{1}{r^4} (|\Delta g^3 u_i|^2 + |\nabla g^3 u_i|^4) + \frac{1}{r^2} |\partial_r \nabla g^3 u_i|^2 + \frac{1}{r^2} |\partial_r \Delta g^3 u_i|^2 \, dx \leq C \varepsilon^2 \left( \frac{r}{\delta} + \frac{\lambda_i R}{r} \right).
\]

The main tool to prove the above lemma is the following three circle theorem for approximate biharmonic functions:
Theorem 3.4. (Theorem 4 of [33]) There is some constant \( \eta_0 > 0 \) such that the following is true. Assume that \( v : A_{l-1} \cup A_l \cup A_{l+1} \to \mathbb{R}^N \) is an \( \eta_0 \)-approximate biharmonic function in the sense of Definition 3.1. Suppose

\[
\max_{l-1, l, l+1} \| \partial^{4(1-1/p)} v \|_{L^p(A_l)}^2 \leq \eta_0 F_l(v)
\]

and

\[
\int_{\partial B_r(p)} v d\sigma_g = 0,
\]

then we have the following alternatives:

(a) if \( F_{l+1}(v) \leq e^{-L} F_l(v) \), then \( F_l(v) \leq e^{-L} F_{l-1}(v) \);

(b) if \( F_{l-1}(v) \leq e^{-L} F_l(v) \), then \( F_l(v) \leq e^{-L} F_{l+1}(v) \);

(c) either \( F_l(v) \leq e^{-L} F_{l-1}(v) \), or \( F_l(v) \leq e^{-L} F_{l+1}(v) \).

Here \( r = d_g(x, p) \), \( A_l = B_{e^{-(l-1)L}} \setminus B_{e^{-lL}}(p) \subset (M, g) \),

\[
F_l(v) = \int_{A_l} \frac{v^2}{r^4} dV_g.
\]

In this paper, we shall apply the above theorem only in the flat metric situation. Correspondingly,

\[
F_l(v) := \int_{A_l} \frac{v^2}{|x|^4} dx.
\]

Corollary 3.5. If \( u_i - (u_i)^* \equiv v_i \) is an \( \eta_0 \)-approximate biharmonic function on \( \Sigma = B_\delta \setminus B_{\lambda, R} \) in the sense of Definition 3.1 with respect to the Euclidean metric, and it satisfies (8) with the condition

\[
\| \partial^{4(1-1/p)} v_i \|_{L^p(A_l)} \leq \eta(r) \varepsilon.
\]

Here \( (u_i)^*(r) = \frac{1}{|\partial B_r|} \int_{\partial B_r(0)} u_id\sigma \) and \( \eta(r) = O \left( \left( \frac{\Delta R}{r} \right)^\kappa + \left( \frac{r}{s} \right)^\kappa \right) \), \( \kappa > 0 \). Then we have

\[
F_l(v_i) \leq C \varepsilon^2 (e^{-\hat{\kappa}(l-b_l)L} + e^{-\hat{\kappa}(l_i-l)L}),
\]

and

\[
\int_{(-L, -(l-1)L) \times S^3} (|\Delta_{S^3} u_i|^2 + |\nabla_{S^3} u_i|^4 + |\partial_t \nabla_{S^3} u_i|^2 + |\partial_t \Delta_{S^3} u_i|^2) dt d\theta \leq C \varepsilon^2 (e^{-\hat{\kappa}(l-b_l)L} + e^{-\hat{\kappa}(l_i-l)L}),
\]

where \( \hat{\kappa} = \min(1, 2\kappa) \). In particular, if \( \kappa \geq 1/2 \), then \( u_i \) satisfies the estimate in Lemma 3.3.

Remark 3.6. Corollary 3.5 is almost a direct consequence of Theorem 3.4 and Lemma 3.2. Firstly, it is easy to see that \( v_i = u_i - (u_i)^* \) satisfies (11). Secondly, (9) says that \( F_l(v_i) \leq C \varepsilon^2 \) for all \( l_0 < l < l_i \). Thirdly, when condition (10) in Theorem 3.4 is not satisfied, (12) ensures that we can get the desired decay estimate in (13). Notice that \( \nabla_{S^3} u_i = \nabla_{S^3} v_i \), by using
Lemma 3.2 and Sobolev embedding theorem we obtain the decay estimate for the tangential part energy of $u_i$ in (14). For more details, we refer to the proofs of Lemma 3 and Lemma 4 in [33].

By using Corollary 3.5, to prove Lemma 3.3, the key is to rewrite the biharmonic map equation for $u_i$ (w.r.t the metric $g$) to be some approximate biharmonic map equation (w.r.t the flat metric), and then apply similar arguments as in [33] to show that $u_i - (u_i)^*$ is an approximate biharmonic function.

**Proof of Lemma 3.3.** Recall that the extrinsic biharmonic map equation (2) can be written as (see [34])

$$
\triangle_g^2 u = \alpha_1(u) \nabla_g \triangle_g u^\flat \nabla_g u + \alpha_2(u) \nabla_g^2 u^\flat \nabla_g^2 u + \alpha_3(u) \nabla_g u^\flat \nabla_g u^\flat \nabla_g u + \alpha_4(u) \nabla_g u^\flat \nabla_g u^\flat \nabla_g u^\flat \nabla_g u.
$$

(15)

Here $\alpha_i$ is a smooth function of $u$ and $\cdot^\flat$ is the contraction of tensors (with respect to the metric $g$), for which we have, for example,

$$
|\nabla_g \triangle_g u^\flat \nabla_g u| \leq C |\nabla_g \triangle_g u| |\nabla_g u|.
$$

(16)

Let $\rho \in [e^{-(l-1)L}, e^{-lL}]$, we can assume

$$
\|g_{ij}(\rho x) - \delta_{ij}\|_{C^4} < \eta(\rho),
$$

(17)

where $\eta(\rho) \to 0$ as $\rho \to 0$. Namely, the metric after scaling to $B_2 \setminus B_1$ is close to the flat metric in $C^4$ norm.

We make the following:

**Claim 0:** Actually we can take $\eta(\rho) = O(\rho^2)$.

In local coordinates on $B_2 \setminus B_1$, we have

$$(\nabla_g u)^i = g^{ij} \partial_j u = g^\flat \partial u$$

and

$$
\triangle_g u = \frac{1}{\sqrt{g}} \partial_i (\sqrt{g} g^{ij} \partial_j u) = \frac{1}{\sqrt{g}} (\partial \sqrt{g}) g^\flat \partial u + \partial g^\flat \partial u + g^\flat \partial \partial u,
$$

where $g \equiv \text{det} [g_{ij}]$ and $\partial g$ denotes the first derivatives of the coefficients of the metric tensor. Note that $\cdot^\flat$ in the last two formulas denote Einstein summation in local coordinates, we use the same notation here since both tensor contraction and Einstein summation satisfy estimates of type (16).

Now by the above two formula we rewrite (15) in local coordinates. We will not give the full details here, since it is just a direct computation which is long but not difficult. We shall
compute the term $\nabla_g \triangle_g u^2 \nabla_g u$, the others follows in a similar way.

$$\nabla_g \triangle_g u^2 \nabla_g u = g^g \partial (\triangle_g u)^2 g^\partial u$$

$$= g^g g^\partial \partial \partial u^2 g^\partial u + \left( \frac{1}{\sqrt{g}} g^g \partial \sqrt{g} \partial g + 2g^g \partial g \right) \partial \partial u^2 g^\partial u$$

$$+ \left( g^g \partial \frac{1}{\sqrt{g}} \sqrt{g} g + \frac{1}{\sqrt{g}} g^g \partial \sqrt{g} \partial g + g^g \partial g \right) \partial \partial u^2 g^\partial u$$

So by using (17), we have

$$|\nabla_g \triangle_g u^2 \nabla_g u - \nabla \triangle u^2 \nabla u| \leq C \eta (|\nabla^3 u + \nabla^2 u + \nabla u|^2)$$

on $B_{2} \setminus B_{1}$, where $g_{kl} \equiv g_{kl}(\rho x)$ and $\nabla, \triangle$ are the gradient operator and Laplacian operator with respect to the flat metric on $B_{2} \setminus B_{1}$.

By repeated computation (with the use of (17)) as above, it is easy to see that (15) (for $u_i(\rho \tilde{x})$) in local coordinates $(\tilde{x})$ on $B_{2} \setminus B_{1}$ is equivalent to

$$\triangle^2 u = \alpha_1(u) \nabla \triangle u^2 \nabla u + \alpha_2(u) \nabla^2 u^2 \nabla u$$

$$+ \alpha_3(u) \nabla^2 u^2 \nabla u^2 \nabla u + \alpha_4(u) \nabla u^2 \nabla u^2 \nabla u + \tilde{\tau}(u),$$

where $\tilde{\tau}(u)$ satisfies

$$|\tilde{\tau}(u)| \leq C \eta (|\nabla^3 u| + \ldots + |\nabla u|) + C \eta (|\nabla^3 u| + |\nabla^2 u| + |\nabla u|)|\nabla u|$$

$$+ C \eta (|\nabla^2 u| + |\nabla u|)(|\nabla^2 u| + |\nabla u|) + C \eta |\nabla u|^{4}.$$

Therefore from (9) and the $\varepsilon$-regularity theorem (Theorem 2.9) together with the use of Sobolev embedding theorems and Hölder inequality, we can obtain that for any $p > 1$,

$$||\tilde{\tau}(u_i(\rho \tilde{x}))||_{L^p(B_{2} \setminus B_{1})} \leq C \varepsilon \eta (\rho).$$

After scaling back to the scale of $A_i = B_{e^{-(1-\kappa)L} \setminus B_{e^{-L}}}$, the equation (15) for $u_i(x)$ in local coordinates is equivalent to

$$\triangle^2 u = \alpha_1(u) \nabla \triangle u^2 \nabla u + \alpha_2(u) \nabla^2 u^2 \nabla u$$

$$+ \alpha_3(u) \nabla^2 u^2 \nabla u^2 \nabla u + \alpha_4(u) \nabla u^2 \nabla u^2 \nabla u + \tau(u).$$

It is easy to see that $\tau(u_i(x)) = \rho^{-3} \tilde{\tau}(u_i(\rho \tilde{x}))$. By using (18), we have that

$$|||x|^{d(1-1/p)} \tau(u_i(x))||_{L^p(A_i)} \leq C \varepsilon \eta (e^{-4L}).$$

Next, by applying an argument as in the proof of Lemma 2 of [33], we know that $v_i = u_i - u_i^*$ is an $\eta_0$-approximate biharmonic function (w.r.t. the Euclidean metric) in the sense of Definition 3.1 defined on $B_{h} \setminus B_{A_i R}$, if $i$ and $R$ are large, $\delta$ and $\varepsilon$ are small. That is

$$\triangle^2 v_i(r, \theta) = a_1 \nabla \triangle v_i + a_2 \nabla v_i + a_3 \nabla v_i + a_4 v_i$$

$$+ \frac{1}{|\partial B_r|} \int_{\partial B_r} b_1 \nabla \triangle v_i + b_2 \nabla v_i + b_3 \nabla v_i + b_4 \nabla v_i d\sigma + h_i(x),$$

where the coefficients satisfy the conditions in Definition 3.1 and

$$h_i = \tau(u_i) - \frac{1}{|\partial B_r|} \int_{\partial B_r(0)} \tau(u_i) d\sigma.$$

Here $\eta_0 > 0$ is the small constant in Theorem 3.4.
Now, we are able to show the decay on the neck for the tangential part energy. Thanks to Claim 0, we have that
\[
|\| |x|^{4(1-1/p)}h_i| |_{L^p(A)} \leq C|\| \tilde{T}(u_i)| |_{L^p(B_2 \setminus B_1)} = O(\varepsilon e^{-2L}) = O(\varepsilon |x|^2).
\]
In particular, \( h_i \) satisfies the condition (b) in Definition 3.1 when \( \varepsilon > 0 \) is small enough. Therefore, by applying Corollary 3.5, we have for any \( \varepsilon > 0 \), if \( i \) and \( R \) are large enough and \( \delta > 0 \) is sufficiently small, then
\[
\int_{-l_0 L < \mathbf{t} < (l_1 - 1)L} (|\Delta_S u_i|^2 + |\partial_t \Delta_S u_i|^2 + |\partial_t \Delta_S u_i|^2) dt d\theta \leq C\varepsilon^2 \left( e^{-(l_1 - 1)L} + e^{-(l_1 - 1)L} \right),
\]
where \( \log \delta = -l_0 L \) and log \( \lambda_i R = -l_i L \).

To complete the above proof, now we prove Claim 0.

**Proof of Claim 0.** Since the metric \( g \) is fixed, we can assume that the metric on \( B_{2\rho} \) has bounded geometry, namely, the norm of the Riemannian curvature tensor \( |R_g|_g \) and the norms of its covariant derivatives \( |\nabla g R_g|_g, |\nabla^2 g R_g|_g \) are bounded. Then by the scaling property of these tensors, we have that the metric \( \tilde{g} = \rho^2 g \) on \( B_2 \) (w.r.t \( \tilde{g} \)) satisfies
\[
|R_{\tilde{g}}|_{\tilde{g}} < C\rho^2, \quad |\nabla_{\tilde{g}} R_{\tilde{g}}|_{\tilde{g}} < C\rho^3, \quad |\nabla^2_{\tilde{g}} R_{\tilde{g}}|_{\tilde{g}} < C\rho^4.
\]
Then in some coordinates \( \tilde{x} \) (see Chapter 5 of [46] or Section 5 of [31])
\[
\tilde{g}_{pq}(\tilde{x}) = \delta_{pq} + 1/3R_{\tilde{g}pqij}\tilde{x}^i\tilde{x}^j + 1/6R_{\tilde{g}pqij,k}\tilde{x}^i\tilde{x}^j\tilde{x}^k + (1/20R_{\tilde{g}pqij,kl} + 2/45R_{\tilde{g}pqijm}\tilde{R}_{\tilde{g}ijklm})\tilde{x}^i\tilde{x}^j\tilde{x}^k\tilde{x}^l + O(\rho^5),
\]
Notice that if we set \( x = \rho \tilde{x} \), then \( \tilde{g}_{ij}(\tilde{x}) = g_{ij}(x) \), hence
\[
g_{ij}(\rho \tilde{x}) - \delta_{ij} \leq C\rho^2,
\]
\[
\partial_k g_{ij}(\rho \tilde{x}), \partial_k \partial_l g_{ij}(\rho \tilde{x}) \leq C\rho^2,
\]
\[
\partial^3 g_{ij}(\rho \tilde{x}), \partial^4 g_{ij}(\rho \tilde{x}) \leq C\rho^4,
\]
where \( \partial u \equiv \frac{\partial}{\partial x^i} u \) and \( \partial^3 \) represents all the third derivatives (w.r.t. \( \tilde{x} \)).

[3.2. Pohožaev identity on Riemannian manifolds. Recall that Pohožaev type estimates or identities for biharmonic maps on \( B_0(0) \subset \mathbb{R}^4 \) were already derived in [54, 25, 57, 58, 33, 34, 30].

In this subsection, we shall establish the Pohožaev identity for extrinsic biharmonic maps defined on general Riemannian manifolds. Let \( (B_1(p), q) \) be a geodesic ball in \( (M, g) \). Let \( (x) \) be the normal coordinates centered at the point \( p \), and set \( e_i = (\frac{\partial}{\partial x^i})_p, i = 1, \cdots, 4 \). For any \( q \in B_1(p) \), we can obtain an orthogonal frame \( \{ e_i(q) \}, i = 1, \cdots, 4 \), by parallel transport along the geodesic line from \( p \) to \( q \). Set \( r^2 = \sum_{i=1}^{4} x_i^2 \), then we have \( r \partial_r = \sum_{i=1}^{4} x_i^2 e_i \).

Firstly, we multiply the extrinsic biharmonic map equation by \( x^k e_k(u) \) and then we do integration by part. Thanks to the fact that \( \Delta^2_g u \perp T_u N \) almost everywhere for an extrinsic
biharmonic map \( u \), we have

\[
0 = \int_{B_r(p)} (x^k e_k(u)) (\Delta_g^2 u) dV_g \\
= -\int_{B_r(p)} \left\langle \nabla_g (x^k) e_k u, \nabla_g (\Delta_g u) \right\rangle dV_g - \int_{\partial B_r(p)} \left\langle x^k \nabla_g e_k u, \nabla_g \Delta u \right\rangle d\sigma_g \\
+ \int_{\partial B_r(p)} (x^k e_k(u)) (\partial_r \Delta_g u) d\sigma_g \\
= \int_{B_r(p)} \Delta_g x^k e_k u (\Delta_g u) dV_g + 2 \int_{B_r(p)} \left\langle \nabla_g x^k \nabla_g (e_k u) \Delta_g u dV_g \\
- \int_{\partial B_r(p)} \partial_r (x^k e_k u) \Delta_g u d\sigma_g + \int_{B_r(p)} x^k \Delta_g (e_k u) \Delta_g u dV_g \\
+ \int_{\partial B_r(p)} (x^k e_k(u)) (\partial_r \Delta_g u) d\sigma_g.
\]

Next, we show that

\[
\int_{B_r(p)} x^k \Delta_g (e_k u) \Delta_g u dV_g
\]

is closely associated with the following quantity

\[
r \int_{\partial B_r(p)} \frac{|\Delta_g u|^2}{2} d\sigma_g - \int_{B_r(p)} \text{div}_g (r \partial_r) \frac{|\Delta_g u|^2}{2} dV_g.
\]

Direct calculations show

\[
r \int_{\partial B_r(p)} \frac{|\Delta_g u|^2}{2} d\sigma_g = \int_{B_r(p)} \text{div} \left( \frac{|\Delta_g u|^2}{2} x^k e_k \right) dV_g \\
= \int_{B_r(p)} x^k e_k (\Delta_g u) \Delta_g u dV_g + \int_{B_r(p)} \text{div}_g (r \partial_r) \frac{|\Delta_g u|^2}{2} dV_g.
\]

By the Ricci identity,

\[
x^k \Delta_g e_k u = x^k e_k (\Delta_g u) + R_{icg}(\nabla_g u, x^k e_k).
\]

It follows that

\[
\int_{B_r(p)} x^k (\Delta_g e_k u) \Delta_g u dV_g \\
= \int_{B_r(p)} [x^k e_k (\Delta_g u) + R_{icg}(\nabla_g u, x^k e_k)] \Delta_g u dV_g \\
= r \int_{\partial B_r(p)} \frac{|\Delta_g u|^2}{2} d\sigma_g + \int_{B_r(p)} R_{icg}(\nabla_g u, x^k e_k) \Delta_g u - \text{div}_g (r \partial_r) \frac{|\Delta_g u|^2}{2} dV_g.
\]
Therefore, by putting the above formulas together, we have the following Pohožaev identity for biharmonic maps over Riemannian geodesic balls:

\[
0 = \int_{B_r(p)} (\triangle_g x^k) e_k u (\triangle_g u) dV_g + 2 \int_{\partial B_r(p)} (\nabla_g x^k) \nabla_g (e_k u) \triangle_g u d\sigma_g - \int_{\partial B_r(p)} \frac{|\triangle_g u|^2}{2} dV_g - \int_{\partial B_r(p)} \partial_r (x^k e_k u) \triangle_g u d\sigma_g \\
+ \int_{\partial B_r(p)} (x^k e_k(u))(\partial_r \triangle_g u) d\sigma_g + r \int_{\partial B_r(p)} \frac{|\triangle_g u|^2}{2} d\sigma_g \\
+ \int_{B_r(p)} Ric_g(\nabla_g u, x^k e_k) \triangle_g u dV_g - \int_{B_r(p)} (\text{div}_g(r \partial_r) - 4) \frac{|\triangle_g u|^2}{2} dV_g.
\]

Finally, by putting the boundary terms to the left hand side, we have that

\[
\int_{\partial B_r(p)} (r \partial_r u)(\partial_r \triangle_g u) + r \frac{|\triangle_g u|^2}{2} - \partial_r (r \partial_r u) \triangle_g u d\sigma_g \\
= - \int_{B_r(p)} (\triangle_g x^k) e_k u (\triangle_g u) - 2 \int_{B_r(p)} (\nabla_g x^k) \nabla_g (e_k u) \triangle_g u - |\triangle_g u|^2 dV_g \\
+ \int_{B_r(p)} (\text{div}_g(r \partial_r ) - 4) \frac{|\triangle_g u|^2}{2} dV_g - \int_{B_r(p)} Ric_g(\nabla_g u, x^k e_k) \triangle_g u dV_g.
\]  

(20)

We shall denote the integrand in the left hand side of (20) by \(\Box\), and denote the whole right hand side integration term by \(P(g, r)\).

To make better use of (20), we need to decompose \(\triangle_g u\) into the radial part and the tangential part. Let \(\partial B_r(p)\) be the sphere \(\partial B_r(p)\) with the metric \(\frac{g}{r}\), where \(g_r\) is the metric of \(\partial B_r(p)\) induced from \(M\). Then we have

\[
\triangle_g u = \partial_r^2 u + \frac{\tilde{m}}{r} \partial_r u + \frac{1}{r^2} \tilde{\Delta} u,
\]

where \(\tilde{m}\) is the mean curvature, \(\tilde{\nabla}\) is the gradient operator, and \(\tilde{\Delta}\) is the associated Laplace operator of \(\partial B_r(p)\). And direct computations yield

\[
(\triangle_g u)^2 = (\partial_r^2 u)^2 + (\frac{\tilde{m}}{r})^2 (\partial_r u)^2 + \frac{1}{r^4} (\tilde{\Delta} u)^2 \\
+ (\frac{2\tilde{m}}{r})(\partial_r u)(\partial_r^2 u)^2 + \frac{2}{r^2} (\tilde{\Delta} u)(\partial_r^2 u) + \frac{2\tilde{m}}{r^3} (\partial_r u)(\tilde{\Delta} u).
\]

The integrand term in the left hand side of (20)

\[
\Box = (r \partial_r u)(\partial_r \triangle_g u) + r \frac{|\triangle_g u|^2}{2} - \partial_r (r \partial_r u) \triangle_g u
\]

consists of the following two terms denoted by \(r \Box_1\) and \(r \Box_2\):
\( \Box_1 \equiv \frac{1}{2}(\Delta_g u)^2 - (\partial_r^2)u \Delta_g u \)

\begin{equation}
\Box_1 = -1/2(\partial_r^2)u^2 + \left( \frac{\tilde{m}^2}{2r^2} \right) (\partial_r u)^2 + \left( \frac{\tilde{m}}{r^3} \right) (\partial_r u) (\tilde{\Delta} u),
\end{equation}

\( \Box_2 \equiv (\partial_r u) (\partial_r \Delta_g u) - \frac{1}{r} (\partial_r u) (\Delta_g u) \)

\begin{equation}
\Box_2 = (\partial_r u) (\partial_r^2 u) + \frac{\tilde{m}}{r} (\partial_r u) (\partial_r^2 u) + \left( \frac{\partial_r \tilde{m}}{r} - \frac{2\tilde{m}}{r^2} \right) (\partial_r u)^2 \\
+ \frac{1}{r^2} (\partial_r u) (\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3} (\partial_r u) (\tilde{\Delta} u).
\end{equation}

So, the identity (20) can be written briefly as

\begin{equation}
\int_{\partial B_r(p)} r (\Box_1 + \Box_2) d\sigma_g = P(g, r).
\end{equation}

Next, we shall estimate \( P(g, r) \) in (23). It turns out that this estimate is crucial in our proof of Theorem 2.15. A key observation is that the quantity \( P(g, r) \) involves at most the second derivatives of \( u \). We remark that the idea of putting the effect of the metric \( g \) into the bi-tension field (or error term) does not work for the Pohožaev type argument.

For a fixed Riemannian manifold \((M, g)\), we can assume that \((\tilde{x} = \rho^{-1} x)\)

\[ \left\| \bar{g}_{ij}(\rho\tilde{x}) - \delta_{ij} \right\|_{C^4(B_{\tilde{r}})} < C\rho^2. \]

Namely, the metric after scaling is close to the flat metric in \( C^4 \) norm.

**Lemma 3.7.** Let \( x^i, e_i, \tilde{\Delta} \) and \( \tilde{m} \) be as above, then it is not hard to check that for all \( q \in B_{2\rho}(p) \), the followings hold:

1) \[ |e_i(x^j) - \delta_i^j| < C\rho^2; \]

2) for \( i = 1, \cdots, 4 \), \[ |\Delta_g x^i| < C\rho; \]

3) on \( \partial B_{\rho}(p) \), for any smooth function \( f \),

\[ |\nabla_{S^3} f - \tilde{\nabla} f| < C\rho^2 |\nabla_{S^3} f|, \]

\[ |\Delta_{S^3} f - \tilde{\Delta} f| < C\rho^2 (|\Delta_{S^3} f| + |\nabla_{S^3} f|), \]

where the norm is dependent on the standard metric on \( S^3 \) and it is equivalent to the norm defined by \( \frac{\partial}{\partial \rho} \), and

\[ |\partial_r \Delta_{S^3} f - \partial_r \tilde{\Delta} f| < C\rho^2 (|\partial_r \Delta_{S^3} f| + |\partial_r \nabla_{S^3} f|) + C\rho (|\Delta_{S^3} f| + |\nabla_{S^3} f|); \]

4) on \( \partial B_{\rho}(p) \),

\[ \Delta_g r = \frac{3 + O(r^2)}{r}, \quad \tilde{m} = r\Delta_g r = 3 + O(r^2); \]
5) on $\partial B_\rho(p)$,
\[ |\partial_r \tilde{m}| < C\rho; \]

6)
\[ \text{div}_g(r \partial_r) - 4 = r \triangle_g r + 1 - 4 = O(r^2). \]

We put the proof of Lemma 3.7 in Appendix A.

Recall that
\[ P(g, r) = -\int_{B_r(p)} (\triangle_g x^k)e_k u(\triangle_g u) - 2 \int_{B_r(p)} (\nabla_g x^k)\nabla_g (e_k u) \triangle_g u - |\triangle_g u|^2 dV_g \]
\[ + \int_{B_r(p)} (\text{div}_g(r \partial_r) - 4) \frac{|\triangle_g u|^2}{2} dV_g - \int_{B_r(p)} \text{Ric}_g(\nabla_g u, x^k e_k) \triangle_g u dV_g. \]

With the help of Lemma 3.7, we can easily derive the following error estimate:

**Lemma 3.8.**

(24) \[ P(g, r) = O(r^2). \]

**Proof.** Indeed, by 1) of Lemma 3.7,
\[ (\nabla_g x^k)\nabla_g (e_k u) = \sum_{i=1}^4 e_i x^k \nabla_i (e_k u) = e_k x^k \nabla_k (e_k u) + \sum_{l \neq k} e_l x^k \nabla_l (e_k u) \]
\[ = O(r^2) |\nabla_g^2 u| + \triangle_g u. \]

So thanks to
\[ \int_{B_1} |\nabla_g^2 u|^2 + |\nabla_g u|^4 dV_g < C, \]

it follows that
\[ \int_{B_r(p)} (\nabla_g x^k)\nabla_g (e_k u) \triangle_g u - |\triangle_g u|^2 dV_g \]
\[ = \int_{B_r(p)} O(r^2) |\nabla_g^2 u|^2 dV_g = O(r^2). \]

By using Hölder inequality, we have
\[ \int_{B_r(p)} |\triangle_g u| |\nabla_g u| dV_g \]
\[ \leq \left( \int_{B_r(p)} |\nabla_g^2 u|^2 dV_g \right)^{1/2} \left( \int_{B_r(p)} |\nabla_g u|^4 dV_g \right)^{1/4} \left( \int_{B_r(p)} dV_g \right)^{1/4} \leq Cr. \]

Notice that $\text{Ric}_g$ is bounded, $\triangle_g x^i = O(r)$ and $\text{div}_g(r \partial_r) - 4 = O(r^2)$ by Lemma 3.7, then (24) follows from simple computations. \[ \square \]
Now we shall reveal from (23) the relationship between the tangential part energy and the radial part energy, see (31) and (32) in the following arguments.

Integrating by part, we get

\[
\int_{B_R \setminus B_r} (\partial_r u)(\partial_r^2 u) dV_g = \int_{r}^{R} ds \int_{\partial B_s} (\partial_s u)(\partial_s^2 u) d\sigma_g
\]

\[
(25) \quad = \int_{r}^{R} ds \int_{S^3} \left( s^3 \sqrt{g(s, \theta)} \right) (\partial_s u)(\partial_s^2 u) d\theta
\]

\[
= \int_{r}^{R} ds \frac{d}{ds} \int_{S^3} \left( s^3 \sqrt{g(s, \theta)} \right) (\partial_s u)(\partial_s^2 u) d\theta
\]

\[
- \int_{r}^{R} ds \int_{S^3} \left( \frac{d}{ds} \left( s^3 \sqrt{g(s, \theta)} \right) \right) (\partial_s u)(\partial_s^2 u) d\theta - \int_{r}^{R} ds \int_{S^3} \left( s^3 \sqrt{g(s, \theta)} \right) (\partial_s^2 u)^2 d\theta.
\]

Here $\sqrt{g}$ is the square root of the determinant of the metric $g$. Set $A(r, \theta) = r^3 \sqrt{g(r, \theta)}$, notice that

\[
\frac{\partial_r A(r, \theta)}{A(r, \theta)} = \Delta_g r = \frac{3}{r} + O(r),
\]

we have

\[
(26) \quad \int_{B_R \setminus B_r} (\partial_r u)(\partial_r^2 u) dV_g = \int_{\partial(B_R \setminus B_r)} (\partial_r u)(\partial_r^2 u)
\]

\[
- \int_{B_R \setminus B_r} (\partial_r^2 u)^2 + \left( \frac{3}{r} + O(r) \right) (\partial_r u)(\partial_r^2 u) dV_g.
\]

Recall that

\[
\Box_2 = (\partial_r u)(\partial_r^2 u) + \frac{\tilde{m} - 1}{r} (\partial_r u)(\partial_r^2 u) + \left( \frac{\partial_r \tilde{m}}{r} - \frac{2\tilde{m}}{r^2} \right) (\partial_r u)^2
\]

\[
+ \frac{1}{r^2} (\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3} (\partial_r u)(\tilde{\Delta} u).
\]

Therefore we have

\[
\int_{B_R \setminus B_r} \Box_2 dV_g
\]

\[
= \int_{\partial(B_R \setminus B_r)} (\partial_r u)(\partial_r^2 u) d\sigma_g - \int_{B_R \setminus B_r} (\partial_r^2 u)^2 + \left( \frac{3}{r} + O(r) \right) (\partial_r u)(\partial_r^2 u) dV_g
\]

\[
+ \int_{B_R \setminus B_r} \frac{\tilde{m} - 1}{r} (\partial_r u)(\partial_r^2 u) + \left( \frac{\partial_r \tilde{m}}{r} - \frac{2\tilde{m}}{r^2} \right) (\partial_r u)^2 dV_g
\]

\[
+ \int_{B_R \setminus B_r} \frac{1}{r^2} (\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3} (\partial_r u)(\tilde{\Delta} u) dV_g.
\]

In the above $\int_{\partial(B_R \setminus B_r)} f d\sigma_g \equiv \int_{\partial B_R} f d\sigma_g - \int_{\partial B_r} f d\sigma_g$. 
Then by direct calculations using 4) and 5) of Lemma 3.7, if we choose $R = cr$ for some $c > 1$, we have

$$
\int_{B_R \setminus B_r} \Delta_2 dV_g =
\int_{\partial(B_R \setminus B_r)} (\partial_r u)(\partial_r^2 u) d\sigma_g + \int_{B_R \setminus B_r} \left( O(1) - \frac{6 + O(r^2)}{r^2} \right) (\partial_r u)^2 - (\partial_r^2 u)^2 dV_g 
+ \int_{B_R \setminus B_r} \left( \frac{2 + O(r^2)}{r} \right) (\partial_r u)(\partial_r^2 u) - \left( \frac{3}{r} + O(r) \right) (\partial_r u)(\partial_r^2 u) dV_g 
+ \int_{B_R \setminus B_r} \frac{1}{r^2} (\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{m}{r^3} (\partial_r u)(\tilde{\Delta} u) dV_g.
$$

Hence

$$
\int_{B_R \setminus B_r} \Delta_2 dV_g =
\int_{\partial(B_R \setminus B_r)} (\partial_r u)(\partial_r^2 u) d\sigma_g + \int_{B_R \setminus B_r} \left( -\frac{5}{r^2} + O(1) \right) (\partial_r u)^2 - (\partial_r^2 u)^2 dV_g 
+ \int_{B_R \setminus B_r} \left( -\frac{1}{2r} + O(r) \right) \partial_r (\partial_r u)^2 - \frac{1}{r^2} (\partial_r u)^2 dV_g 
+ \int_{B_R \setminus B_r} \frac{1}{r^2} (\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{m}{r^3} (\partial_r u)(\tilde{\Delta} u) dV_g.
$$

Next by the use of the formula $\frac{\partial_r A(r, \theta)}{A(r, \theta)} = \Delta_g r = \frac{3}{r} + O(r)$ again, we get

$$
\int_{\partial(B_R \setminus B_r)} \frac{1}{2s} (\partial_s u)^2 d\sigma_g =
\int_r^R \frac{d}{ds} \int_{\partial Bs} \frac{1}{2s} (\partial_s u)^2 d\sigma_g ds 
= \int_r^R \frac{d}{ds} \int_{S^1} \frac{1}{2s} (\partial_s u)^2 A(s, \theta) d\theta ds 
= \int_r^R \int_{\partial B_x} \frac{1}{2s} (\partial_s u)^2 - \frac{1}{2s^2} (\partial_s u)^2 + \frac{1}{2s} (\Delta_g r)(\partial_s u)^2 d\sigma_g ds 
= \int_r^R \frac{1}{2s} (\partial_s u)^2 + \frac{1}{s^2} (\partial_s u)^2 + O(1)(\partial_s u)^2 d\sigma_g 
= \int_{B_R \setminus B_r} \frac{1}{2s} (\partial_s u)^2 + \frac{1}{s^2} (\partial_s u)^2 + O(1)(\partial_s u)^2 dV_g.
$$
Therefore
\[
\int_{B_R \setminus B_r} \square_2 dV_g
= \int_{\partial (B_R \setminus B_r)} (\partial_r u)(\partial_r^2 u) d\sigma_g + \int_{B_R \setminus B_r} \left( O(1) - \frac{5}{r^2} \right) (\partial_r u)^2 - (\partial_r^2 u)^2 dV_g
\]
\[
- \int_{\partial (B_R \setminus B_r)} \frac{1}{2r} (\partial_r u)^2 d\sigma_g + \int_{B_R \setminus B_r} O(r) \partial_r (\partial_r u)^2 dV_g
\]
\[
+ \int_{B_R \setminus B_r} \frac{1}{r^2} (\partial_r u)(\partial_r \triangle u) - \frac{m}{r^3} (\partial_r u)(\triangle u) dV_g.
\]

By putting the boundary terms together, we get
\[
\int_{B_R \setminus B_r} \square_2 dV_g
= \int_{\partial (B_R \setminus B_r)} (\partial_r u)(\partial_r^2 u) - \frac{1}{2r} (\partial_r u)^2 d\sigma_g
\]
\[
+ \int_{B_R \setminus B_r} \left( O(1) - \frac{5}{r^2} \right) (\partial_r u)^2 - (\partial_r^2 u)^2 + O(r) \partial_r (\partial_r u)^2 dV_g
\]
\[
+ \int_{B_R \setminus B_r} \frac{1}{r^2} (\partial_r u)(\partial_r \triangle u) - \frac{m}{r^3} (\partial_r u)(\triangle u) dV_g.
\]

By integrating (23) we get
\[
(29) \quad \int_{B_R \setminus B_r} (\square_1 + \square_2) dV_g = \int_r^{R} \int_{\partial B_r} (\square_1 + \square_2) d\sigma_g ds = \int_r^{R} \frac{1}{s} P(g, s) ds.
\]

Recall
\[
\square_1 = -1/2(\partial_r^2 u)^2 + (\frac{\tilde{m}^2}{2r^2})(\partial_r u)^2 + \frac{1}{2r^2}(\triangle u)^2 + (\frac{\tilde{m}}{r^3})(\partial_r u)(\triangle u).
\]

Then by combining (29), (28) and (24) (namely, \(P(g, r) = O(r^2)\)) together, we have
\[
(30) \quad \leq \int_{B_R \setminus B_r} \frac{3}{2}(\partial_r^2 u)^2 + \frac{(10 - \tilde{m}^2)}{2r^2} + O(1) (\partial_r u)^2 dV_g
\]
\[
- \int_{\partial (B_R \setminus B_r)} \frac{1}{2r^4}(\triangle u)^2 + \frac{1}{r^2}(\partial_r u)(\partial_r \triangle u) dV_g
\]
\[
+ \int_{B_R \setminus B_r} O(r)|\partial_r (\partial_r u)^2| dV_g + \int_r^{R} O(s) ds.
\]

Next we estimate \(\int_{B_R \setminus B_r} O(r)|\partial_r (\partial_r u)^2| dV_g\). Since
\[
\int_{B_R} |\nabla g u|^2 + |\nabla g u|^4 dV_g
\]
is bounded, by Hölder inequality, we have
\[ \int_{B_R \setminus B_r} |\partial_r (\partial_r u)^2| dV_g \leq Cr. \]

And Cauchy-Schwarz inequality gives us
\[ \int_{B_{cr} \setminus B_r} \frac{1}{r^2} (\partial_r u)(\partial_r \Delta u) dV_g \leq \int_{B_{cr} \setminus B_r} \frac{1}{r^2} \left( \frac{3}{32} (\partial_r u)^2 + 8 (\partial_r \Delta u)^2 \right) dV_g. \]

So (30) and 4) of Lemma 3.7 imply that
\[ \int_{B_r} \frac{3}{2} (\partial_r^2 u)^2 + \frac{1}{3r^2} (\partial_r u)^2 dV_g \leq \int_{B_r} \frac{1}{2r^4} (\Delta u)^2 + \frac{8}{r^2} (\partial_r \Delta u)^2 dV_g \]
\[ + \int_{\partial (B_{cr} \setminus B_r)} (\partial_r u)(\partial_r^2 u) - \frac{1}{2r} (\partial_r u)^2 d\sigma_g + Cr^2. \]

Set \( r_2(s) = e^{t_0+s} \) and \( r_1(s) = e^{t_0-s} \), here \( \log(\lambda_i R) < t_0 < \log \delta \), and
\[ 0 \leq s \leq \min \left( -t_0 + \log \delta, -\log(\lambda_i R) + t_0 \right). \]

Since we can chose \( c \) such that \( B_{r_2} \setminus B_{r_1} = \bigcup B_{e^{c^2 r_1}} \setminus B_{r_1} \), by adding up inequalities as (31), we have
\[ \int_{B_{r_2} \setminus B_{r_1}} \frac{3}{2} (\partial_r^2 u)^2 + \frac{1}{3r^2} (\partial_r u)^2 dV_g \leq \int_{B_{r_2} \setminus B_{r_1}} \frac{1}{2r^4} (\Delta u)^2 + \frac{8}{r^2} (\partial_r \Delta u)^2 dV_g \]
\[ + \int_{\partial (B_{r_2} \setminus B_{r_1})} (\partial_r u)(\partial_r^2 u) - \frac{1}{2r} (\partial_r u)^2 d\sigma_g + \int_{r_1}^{r_2} O(r) dr. \]

Through the above efforts, now we are ready to prove the energy identity and the no neck identity by the relationship between the tangential part energy and the radial part energy reflected in (31) and (32).

**Proof of Theorem 2.15.** If we set \( c = e^L, r = e^t \), then from the estimate
\[ \int_{B_{e^L} \setminus B_r} \frac{1}{r^4} \left( (\Delta S^3 u_j)^2 + |\nabla S^3 u_j|^4 \right) + \frac{1}{r^2} |\partial_r \nabla S^3 u_j|^2 + \frac{1}{r^2} |\partial_r \Delta S^3 u_j|^2 dx \]
\[ \leq C \varepsilon^2 \left( \frac{r}{\delta} + \frac{\lambda_i R}{r} \right) \]
in Lemma 3.3 and 3) in Lemma 3.7 which indicates that the operators $\tilde{\nabla}$ and $\tilde{\Delta}$ are almost equal to $\nabla_{S^3}$ and $\Delta_{S^3}$, we get by Cauchy-Schwarz inequality that

$$\int_{B_{2r} \setminus B_r} \frac{1}{2r^4} (\tilde{\Delta} u_i)^2 + \frac{8}{r^2} |(\partial_r \tilde{\Delta} u_i)|^2 dV_g$$

$$\leq C \varepsilon^2 \left( \frac{r}{\delta} + \frac{\lambda_i R}{r} \right).$$

Recall that $A_t = B_{e^{-i(L-t)}} \setminus B_{e^{-iL}}$, log $\delta = -l_0 L$ and $\log \lambda_i R = -l_i L$. Then by (31) and (33) we have

$$\int_{B_{3/2} \setminus B_{\lambda_i R}} \frac{3}{2} (\partial^2 u_i)^2 + \frac{1}{3r^2} (\partial_r u_i)^2 dV_g$$

$$\leq \sum_{i_0}^{l_i} \int_{A_t} \frac{1}{2r^4} (\tilde{\Delta} u_i)^2 + \frac{8}{r^2} |(\partial_r \tilde{\Delta} u_i)|^2 dV_g$$

$$+ \int_{\partial (B_{3/2} \setminus B_{\lambda_i R})} |(\partial_r u_i)(\partial^2 u_i)| + \frac{1}{2} (\partial_r u_i)^2 d\sigma_g + C \sum_{i_0}^{l_i} e^{-2iL}$$

$$\leq C \sum_{i_0}^{l_i} \varepsilon^2 \left( \frac{e^{-iL}}{\delta} + \frac{\lambda_i R}{e^{-iL}} \right) + C \sum_{i_0}^{l_i} e^{-2iL}$$

$$+ \int_{\partial (B_{3/2} \setminus B_{\lambda_i R})} |(\partial_r u_i)(\partial^2 u_i)| + \frac{1}{2r^2} (\partial_r u_i)^2 d\sigma_g.$$
Hence
\[ \partial_t F(s) = e^{t_0 - s} \int_{\partial B_{r_2}} \frac{3}{2} (\partial_r^2 u_i)^2 + \frac{1}{3r^2} (\partial_r u_i)^2 d\sigma_g + e^{t_0 - s} \int_{\partial B_{r_1}} \frac{3}{2} (\partial_r^2 u_i)^2 + \frac{1}{3r^2} (\partial_r u_i)^2 d\sigma_g. \]

By applying Cauchy-Schwarz inequality \((\sqrt{3} \leq 1/2(3 + 1))\), we get
\[
\int_{\partial(B_{r_2} \setminus B_{r_1})} (\partial_t u_i)(\partial_r^2 u_i) d\sigma_g \\
\leq \frac{1}{\sqrt{3}} e^{t_0 - s} \int_{\partial B_{r_2}} \frac{3}{2} (\partial_r^2 u_i)^2 + \frac{1}{2r^2} (\partial_r u_i)^2 d\sigma_g + \frac{1}{\sqrt{3}} e^{t_0 - s} \int_{\partial B_{r_1}} \frac{3}{2} (\partial_r^2 u_i)^2 + \frac{1}{2r^2} (\partial_r u_i)^2 d\sigma_g.
\]

By (33), we get
\[
\int_{B_{r_2} \setminus B_{r_1}} \frac{1}{2r^4} (\Delta u)^2 + \frac{8}{r^2} (\partial_r \Delta u)^2 dV_g \leq C \varepsilon^2 (e^{t_0 - \log \delta} + e^{\log(\lambda_i R) - t_0}) e^s.
\]

So from (32), we can deduce that \((1/\sqrt{3} < 2/3)\)
\[(35) \quad F(s) \leq \frac{2}{3} \partial_s F(s) + C e^{2(t_0 + s)} + C \varepsilon^2 (e^{t_0 - \log \delta} + e^{\log(\lambda_i R) - t_0}) e^s.\]

Recall \(\log(\lambda_i R) < t_0 < \log \delta\), and \(0 \leq s \leq \min(-t_0 + \log \delta, -\log(\lambda_i R) + t_0)\). Multiplying \(e^{-\frac{3}{2}s}\) to both sides of (35), if we choose \(\delta\) small such that \(\delta < \varepsilon^2\), then we have
\[ \partial_s (e^{-\frac{3}{2}s} F) \geq -C \varepsilon^2 (e^{t_0 - \log \delta} + e^{\log(\lambda_i R) - t_0}) e^{-\frac{1}{2}s}, \]

since \(e^{2(t_0 + s)} \leq e^{(t_0 + s)} = \delta e^{(t_0 - \log \delta)} e^s \leq \varepsilon^2 e^{(t_0 - \log \delta)} e^s\). Integrating the above inequality from \(s = 1\) to \(s = \min(-t_0 + \log \delta, -\log(\lambda_i R) + t_0)\), we have from the fact that \(F(s)\) is uniformly bounded by \(C \varepsilon^2\) (see (34)), that
\[ F(1) = \int_{B_{r_0 + 1} \setminus B_{r_0 - 1}} \frac{3}{2} (\partial_r^2 u_i)^2 + \frac{1}{3r^2} (\partial_r u_i)^2 dV_g \leq C \varepsilon^2 (e^{t_0 - \log \delta} + e^{\log(\lambda_i R) - t_0}). \]

Now we can apply similar arguments as in Section 5 of [34] to show that the decay of the radial part energy and the tangential part energy (see (33)) is sufficient to guarantee that the following no neck property hold
\[ \lim_{\delta \to 0} \lim_{R \to \infty} \lim_{i \to \infty} \text{osc}_{B_{\delta} \setminus B_{\lambda_i R}} u_i = 0. \]

The idea is to derive the decay of first derivative of \(u_i\) from the energy decay by the \(\varepsilon\)-regularity and the Sobolev embedding theorem. The proof of Theorem 2.15 is finished. \(\square\)
4. Blow-up analysis for maps from degenerating Einstein manifolds

Let \((M_i, g_i)\) be a sequence of 4-dimensional Einstein manifolds as in Theorem 2.1. And let \(u_i\) be a sequence of biharmonic maps from \((M_i, g_i)\) (with uniformly bounded Einstein constants) to \(N\), and satisfy

\[
E(u_i, M_i) \equiv \int_{M_i} |\nabla^2 g_i u_i|^2 + |\nabla g_i u_i|^4 dV_{g_i} \leq \Lambda,
\]

for some \(\Lambda > 0\). Without loss of generality, we can assume by Theorem 2.1 that there exists a subsequence \(\{j\} \subset \{i\}\) such that \((M_j, g_j)\) converges to \((M_\infty, g_\infty)\) which is an Einstein orbifold with only one singularity \(x_a\).

**Remark 4.1.** To study the convergence of maps from a sequence of manifolds, we recall here some necessary background knowledge about the convergence of functions defined on a sequence of metric spaces, see e.g. [21, 23]. Suppose \(X = \lim_{G.-H} X_i\). Extend the metrics on \(X, X_i\) to a metric on \(X \coprod X_i\) so that \(X_i \to X\) inside \(X \coprod X_i\) in the classical Hausdorff sense. We say that \(f_i : X_i \to \mathbb{R}^N\) converges to \(f : X \to \mathbb{R}^N\) provided \(f_i(x_i) \to f(x)\) whenever \(x_i \to x\).

**4.1. Blow-up away from orbifold singularities.** In this subsection, we shall consider the simper case that the maps \(u_j\) blow up away from orbifold singularities.

Away from the orbifold singularity \(x_a\), \((M_j, g_j)\) converges to \((M_\infty, g_\infty)\) smoothly, and

\[
u_j : M_j \to N\]

(sub)converges weakly to a map

\[
u_\infty : M_\infty \to N\]

via some diffeomorphism

\[G_j : M_\infty \setminus \{x_a\} \to M_j.\]

That is

\[u_j \circ G_j \to u_\infty.\]

Now we shall firstly examine how the limit map \(u_\infty : M_\infty \to N\) behave around the singularity \(x_a\). By Remark 2.3, \(\Pi_a^* g_\infty\) extends to a smooth metric on \(B_1\), and on account of the fact that \(u_\infty\) is a smooth map on \(B(x_a, \delta) \setminus \{x_a\}\), we can consider the map \(u_\infty(\Pi_a)\) as a smooth map on \((B_1, \Pi_a^* g_\infty) \setminus \{0\}\). It is easy to check that \(u_\infty(\Pi_a)\) satisfies the biharmonic map equation with respect to the metric \(\Pi_a^* g_\infty\) on \(B_1 \setminus \{0\}\), and the energy of \(u_\infty(\Pi_a)\) over \((B_1, \Pi_a^* g_\infty) \setminus \{0\}\) is finite. By the removable singularity Theorem 2.12, \(u_\infty(\Pi_a)\) extends to a smooth map on \(B_1\). Therefore, \(u_\infty\) is a continuous map on \(B(x_a, \delta) \subset M_\infty\). So we conclude the following:

**Lemma 4.2.** The limit map \(u_\infty\) is a continuous map from \(M_\infty\) to \(N\).

If \(u_j\) blow-up at non-degenerating points \(x_{b,j}\), namely, \((B(x_{b,j}, r), g_j) \to (B(x_b, r), g_\infty)\) smoothly in the Cheeger-Gromov sense, then we can pull the metric of \((B(x_{b,j}, r), g_j)\) back to \(M_\infty\) via a smooth map \(F_j : B(x_b, r) \to M_j\). Therefore locally we can consider \(u_j\) as a sequence of biharmonic maps defined on \(B(x_b, r)\). Without loss of generality, we suppose that away from the orbifold singularity \(x_a, u_j\) blow-up only at \(x_b\), and there is only one bubble map \(\omega^b\).
In local coordinates around the blow-up point $x_b$, it is easy to see that there is a sequence of positive numbers $\lambda_j \to 0$ such that $u_j(\lambda_j x) \to \omega^b$, where
\[
\omega^b : (\mathbb{R}^4, g_{\text{euc}}) \to N
\]
is a biharmonic map and $g_{\text{euc}}$ is the Euclidean metric.

Since $(M_j, g_j)$ converges to $(M_{\infty}, g_{\infty})$ smoothly away from $x_a$, without loss of generality, we can assume that in local coordinates around $x_b$, the metrics $g_j$ satisfy the following: for sufficiently small $\rho > 0$,
\[
\| (g_j)_{kl}(\rho x) - \delta_{kl} \|_{C^4} < C\rho^2
\]
on $B_2 \setminus B_1$ uniformly for large enough $j$. Then by similar arguments as in the proof of Theorem 2.15, we have the following bubble tree convergence away from the orbifold singularity:

**Theorem 4.3.** Let $u_i$ be a sequence of biharmonic maps from $(B(x_b, 1), g_i)$ to $N$ satisfying
\[
E(u_i, B_1) \equiv \int_{B_1} |\nabla^2_{g_i} u_i|^2 + |\nabla_{g_i} u_i|^4 dV_{g_i} \leq \Lambda,
\]
for some $\Lambda > 0$. If $\{u_i\}$ as above blow-up at $x_b$ (away from the orbifold singularity $x_a$ such that $d_{g_{\text{euc}}}(x_a, x_b) > 1$) with only one (nontrivial) bubble $\omega^b : (\mathbb{R}^4, g_{\text{euc}}) \to N$, then
\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{i \to \infty} \int_{B_\delta \setminus B_*} |\nabla^2_{g_i} u_i|^2 + |\nabla_{g_i} u_i|^4 dV_{g_i} = 0
\]
and
\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{i \to \infty} \text{osc}_{B_\delta \setminus B_*} u_i = 0.
\]

In particular, the above theorem indicates that the image in $N$ of the limit map $u_{\infty}$ is connected to the image of the bubble map $\omega^b$.

### 4.2. Blow-up at the orbifold singularity I: single ALE bubble case.

In this subsection, we shall focus on the more subtle case that the maps $u_j$ blow up at the orbifold singularity $x_a \in M_{\infty}$. The analysis involved in this case becomes much more complicated.

For simplicity, we shall assume that there is only one ALE bubble orbifold $M_a$ at $x_a$. Note that $M_a$ is smooth in this case. Later, we will study the more general cases in which we may encounter ALE bubble orbifolds.

First of all, we need to study the domain decomposition of the manifolds $M_j$ near the orbifold singularity $x_a$. Note that we have assumed that $x_a$ is the only singularity and there is only one bubble manifold.

By (3) of Theorem 2.1, for every $j$, there exists $x_{a,j} \in M_j$ and a positive number $r_j$ such that the followings hold:

(a) $B(x_{a,j}, \delta)$ converges to $B(x_a, \delta)$ in the Gromov-Hausdorff distance for all $\delta > 0$.

(b) $\lim_{j \to \infty} r_j = 0$. 
(c) \((M_j, r_j^{-2}g_j, x_{a,j})\) converge to \((M_a, h_a), x_{a,\infty}\) in the Gromov-Hausdorff distance, where \((M_a, h_a)\) is a complete, noncompact, Ricci flat, non-flat 4-manifold which is ALE of order 4.

(d) There exists an into diffeomorphism \(G_j : M_a \to M_j\) such that \(G_j^* (r_j^{-2}g_j)\) converges to \(h_a\) in the \(C^\infty\)-topology on \(M_a\). According to this, for any \(R > 0\), we shall call

\[ A_{r_j, R, \delta}(x_{a,j}) = B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R) \]

the degenerating neck region, and \(B(x_{a,j}, r_j R)\) the ALE bubble domain, and \(M_j \setminus B(x_{a,j}, \delta)\) the base domain.

In general, it is possible that there are several bubble maps \(\omega_k^A\) \((k = 1, \cdots, m_a)\) occurring at the orbifold singularity \(x_a\). Suppose that the maps

\[ u_j : (M_j, (\lambda_j^k)^{-2}g_j, x_{a,k,j}) \to N \]

converge to \(\omega_a^k\) as \(j \to \infty\). Here

\[ \lim_{j \to \infty} d_{g_j}(x_{a,k,j}, x_{a,j}) = 0. \]

To handle the complicated situation considered in this subsection, we shall investigate the blow-up process in the following three cases:

**Case A:** \(\lim_{j \to \infty} \frac{d_{g_j}(x_{a,k,j}, x_{a,j})}{r_j} < \infty\) and \(\lim_{j \to \infty} \frac{\lambda_j^k}{r_j} = 0\);

**Case B:** \(\lim_{j \to \infty} \frac{d_{g_j}(x_{a,k,j}, x_{a,j})}{r_j} < \infty\) and \(0 < \lim_{j \to \infty} \frac{\lambda_j^k}{r_j} < \infty\);

**Case C:** others.

Roughly speaking, the domain of the bubble map of Case B is (almost) the ALE bubble domain, bubbles of case A are on the bubble of Case B, the bubble of Case B is on some bubble of Case C or is away from bubbles of Case C. We remark here that we can assume \(x_{a,k,j} = x_{a,j}\) in Case B without loss of generality.

In the rest of the paper, bubbles of case A, Case B, Case C will be indicated by \(\omega^A\), \(\omega^B\), \(\omega^C\) respectively. We shall denote the centers of those three types of bubbles by \(x_{a,j}^A\), \(x_{a,j}^B\), \(x_{a,j}^C\) respectively, and denote those three type of bubble scales by \(\lambda_j^A\), \(\lambda_j^B\), \(\lambda_j^C\) respectively.

In case A, it is easy to see that

\[ [u_j : (M_j, (\lambda_j^A)^{-2}g_j, x_{a,j}^A) \to N] \longrightarrow [\omega^A : (\mathbb{R}^4, g_{\text{euc}}, 0) \to N]. \]

In Case B, by an adjustment of the bubble scale, we may assume \(r_j = \lambda_j^B\). By our assumption in the above, we have

\[ [u_j : (M_j, (\lambda_j^B)^{-2}g_j, x_{a,j}^B) \to N] \longrightarrow [\omega^B : (M_a, h_a, x_{a,\infty}) \to N] \]

on any compact subset of the ALE bubble manifold \(M_a\).
Remark 4.4. By our construction, bubble $\omega^A$ lies on the bubble $\omega^B$, and by applying the same arguments for the proof of Theorem 4.3, we know that there is no energy loss in the neck region

$$A_{\lambda_j^A R, \delta_j}(x_{a,j}^A) \equiv B(x_{a,j}^A, \delta_j) \setminus B(x_{a,j}^A, \lambda_j^A R),$$

and the image of $\omega^A$ is connected to the image of $\omega^B$.

It is natural to ask the following:

Question 4.5. How the maps $u_j$ behaves on the degenerate neck regions $A_{r_j R, \delta}(x_{a,j}) = B(x_{a,j}, \delta) \setminus B(x_{a,j}, r_j R)$ when $j \to \infty$?

This question is related to the analysis of bubble maps of Case C. It has certain similarity with the blow-up analysis of harmonic maps from degenerating Riemann surfaces [62]. Set $T_j = \log(r_j R)$ and $T(\delta) = \log(\delta)$, then we have the following bubble-neck decomposition:

Proposition 4.6. Notations and assumptions as above.

1. "Asymptotic boundary conditions":

$$\lim_{j \to \infty} \omega(u_j, P_{T_j, T_j + L}) = \lim_{\delta \to 0} \omega(u_j, P_{T(\delta) - L, T(\delta)}) = 0, \quad \forall L \geq 1,$$

$$\lim_{j \to \infty} \text{osc}_{P_{T_j, T_j + L}} u_j = \lim_{\delta \to 0} \text{osc}_{P_{T(\delta) - L, T(\delta)}} u_j = 0,$$

where $P_{T_j, T_j}$ is the cylinder corresponding to $A_{T_j^1, T_j^2}(x_{a,j}) \subset M_j$,

$$\omega(u_j, P_{T_j^1, T_j^2}) \equiv \sup_{t \in [T_j^1, T_j^2 - 1]} \int_{A_{e_j t, e_j t + 1}(x_{a,j})} |\nabla^2 u_j|^2 + |\nabla g_j u_j|^4 dV_j.$$

2. "bubble domain and neck domain": after selection of a subsequence, which we still denote by $u_j$, the following two alternatives hold:

$$(2.1) \quad \lim_{\delta \to 0} \lim_{j \to \infty} \omega(u_j, P_{T_j, T(\delta)}) = 0,$$

$$(2.2) \quad \text{there exists some number } N_2 > 0 \text{ which is independent of } j, \text{ and } 2N_2 \text{ sequences of numbers } \{a_1^j\}, \{b_1^j\}, \ldots, \{a_{N_2}^j\}, \{b_{N_2}^j\}, \text{ such that } T_j \leq a_1^j \ll b_1^j \leq \cdots \leq a_{N_2}^j \ll b_{N_2}^j \leq T(\delta) \quad (a_\alpha^j \ll b_\alpha^j \text{ means } \lim_{j \to \infty} b_\alpha^j - a_\alpha^j = \infty)$$

and

$$|b_\alpha^j - a_\alpha^j| \ll |T_j|, \quad \text{i.e. } \lim_{j \to \infty} \frac{|b_\alpha^j - a_\alpha^j|}{|T_j|} = 0.$$

Denote

$$J_j^\alpha \equiv P_{a_\alpha^j, b_\alpha^j}, \quad \alpha = 1, \cdots, N_2,$$

$$I_j^0 = P_{T_j, a_1^j} \subset T_j^{N_2} = P_{b_{N_2}^j, T(\delta)}, \quad I_j^\alpha = P_{b_{\alpha}^j, a_{\alpha}^j + 1}, \quad \alpha = 1, \cdots, N_2 - 1.$$

Then

(i) $\alpha = 0, 1, \cdots, N_2, \lim_{\delta \to 0} \lim_{R \to \infty} \lim_{j \to \infty} \omega(u_j I_j^\alpha) = 0$. 
(ii) \( \forall \alpha = 1, \ldots, N_2 \), there is a bubble tree which consists of at most finitely many bubble maps, i.e., finite energy biharmonic maps from \( \mathbb{R}^4 / \Gamma \) to \( N \). Here, for simplicity of notations, we assume there is only one such bubble map, namely, there is a biharmonic map

\[
\omega_{C, \alpha} : \mathbb{R}^4 / \Gamma \rightarrow N,
\]

such that

\[
\lim_{\delta \rightarrow 0} \lim_{R \rightarrow \infty} \lim_{j \rightarrow \infty} E_{bi}(u_j, J_j^\alpha) = E_{bi}(\omega_{C, \alpha}),
\]

where \( J_j^\alpha \) is the region in \( M_j \) corresponding to \( J_j^\alpha \), and

\[
E_{bi}(u_j, J_j^\alpha) \equiv \int_{J_j^\alpha} \left( |\nabla_{g_j}^2 u_j|^2 + |\nabla_{g_j} u_j|^4 dV_{g_j} \right).
\]

(iii) Each bubble \( \omega_{C, \alpha} \) given above is a continuous map, and

\[
\lim_{R \rightarrow \infty} \text{osc}_{\mathbb{R}^4 / \Gamma \setminus B_R} \omega_{C, \alpha} = 0.
\]

Proof. (1) follows easily from our assumption. (2) can be shown in the same sprit as in [62].

For (i), it can be argued as in the proof for Proposition 3.1 of [62].

For (ii), we only need to verify that \( \omega_{C, \alpha} \) is map from \( \mathbb{R}^4 / \Gamma \) to \( N \). Let \( J_j^\alpha \) be the annulus

\[
B(x_{a,j}, e^{b_j^\alpha}) \setminus B(x_{a,j}, e^{\delta_2^\alpha}) \subset M_j \text{ corresponding to } J_j^\alpha.
\]

Proposition 2.6 says that the neck region between \( \omega^B \) and \( u_\infty \) in \( M_j \) looks like a potion of a flat cone \( \mathbb{R}^4 / \Gamma \) for large \( j \), moreover,

\[
r^2 |R_{g_j}| \leq C \max \left\{ \left( \frac{T_j}{r} \right)^{\varepsilon_5}, \left( \frac{T_j}{r_\infty} \right)^{\varepsilon_5} \right\}
\]

on \( J_j^\alpha \), where \( \varepsilon_5 \) and \( r_\infty \) are small positive constants. So by the relation \( a_j^\alpha \ll b_j^\alpha \), after scaling of the scale \( \lambda_j^{C, \alpha} \equiv e^{\frac{a_j^\alpha - \delta_2^\alpha}{2}} \), \( \left( J_j^\alpha, (\lambda_j^{C, \alpha})^{-2} g_j \right) \) converges to \( \mathbb{R}^4 / \Gamma \) as \( j \rightarrow \infty \), and

\[
\left( T_j, (\lambda_j^{C, \alpha})^{-2} g_j \right) \rightarrow N \text{ converges to a bubble } \omega_{C, \alpha}.
\]

For (iii), the first part can be argued as in Lemma 4.2 and the second part follows from the result of Lemma 4.7 which is stated and proved separately afterwards. \( \square \)

Lemma 4.7. The bubble maps \( \omega^B : M_a \rightarrow N \) and \( \omega_{C, \alpha} : \mathbb{R}^4 / \Gamma \rightarrow N \) above is uniformly continuous at the infinity in the sense that, for any \( \delta > 0 \), there is \( R > 0 \) such that

\[
\text{osc}_{M \setminus B_R} \omega < \delta,
\]

where \( M = M_a \) and \( \omega = \omega^B \), or \( M = \mathbb{R}^4 / \Gamma \) and \( \omega = \omega_{C, \alpha} \).

Since biharmonic maps are in general not conformally invariant in dimension 4, we can not apply the removable singularity theorem to see that \( \lim_{|x| \rightarrow \infty} \omega(x) \) is well defined, see e.g. [54, 33].

Proof. We shall only prove the lemma for \( \omega^B \). Since the problem is scaling invariant in dimension 4, we may assume that for some fixed small \( \varepsilon_1 > 0 \) to be determined by the arguments later,

\[
\int_{M_a \setminus B_1} |\nabla_{h_a}^2 \omega^B|^2 + |\nabla_{h_a} \omega^B|^4 dV_{h_a} < \varepsilon_1.
\]
Then the lemma is an immediate consequence of the following Lemma 4.8. □

Lemma 4.8. There is a constant $\varepsilon_1 > 0$ depending on $N$ such that, if $u : M_a \to N$ is a biharmonic map from a Ricci flat ALE manifold (orbifold) of order 4 to $N$ and satisfying

$$\int_{M_a \setminus B_1} |\nabla^2_{h_a} u|^2 + |\nabla_{h_a} u|^4 dV_{h_a} < \varepsilon_1,$$

then $u$ is uniformly continuous at the infinity in the sense that for any $\delta > 0$, there is $R > 0$ independent of $u$ such that

$$\text{osc}_{M_a \setminus B_R} u < \delta.$$

The proof of Lemma 4.8 will be given in Appendix C.

4.3. Blow-up at the singularity II: multiple ALE bubbles case. In this subsection, we shall discuss how to construct the whole bubble tree for the convergence of biharmonic maps in the case that there are several ALE bubble manifolds (orbifolds) emerging (at the same orbifold singularity) from the G-H convergence of Einstein manifolds.

From the analysis in the previous subsections, we know that three types of bubble maps will appear at the orbifold singularity, namely, $\omega^A : \mathbb{R}^4 \to N$, $\omega^B : M_{ALE} \to N$, and $\omega^C : \mathbb{R}^4/\Gamma \to N$, where $M_{ALE}$ represent the ALE bubble manifolds (orbifolds) in the tree $\text{Tr}_{ALE}$ (see Remark 2.5).

Bubble maps of type $\omega^B$ occur in the place where the Riemannian curvatures concentrate, and the number of bubble maps of kind $\omega^B$ is equal to the number of ALE bubble manifolds (orbifolds) in the tree $\text{Tr}_{ALE}$ (if one of them is defined on a ALE orbifold, then we can argue as in Lemma 4.2 to show that it is continuous at the orbifold singularities). Bubble maps of type $\omega^C$ appear in degenerating neck regions such as those discussed in Proposition 2.6. Bubble maps of type $\omega^A$ appear over the regions where there is no degeneration of the metrics at its blow-up scale (they are on bubble maps of type $\omega^B$). With these in mind, one can easily construct the whole bubble map tree by induction. Note that the whole process will be terminated in finite steps, since the number of nontrivial ALE bubble manifolds (orbifolds) is finite and the number of nontrivial bubble maps of type $\omega^A$ and $\omega^C$ must be finite by the energy gap Theorem 2.14) and the finite total energy assumption.

Question 4.9. Is there an energy gap phenomenon for biharmonic maps from the ALE spaces in the tree $\text{Tr}_{ALE}$?

Now we state the main theorem in this paper to end this section.

Theorem 4.10. Let $(M_j, g_j)$ be a sequence of Einstein manifolds as in Theorem 2.1 and let $u_j$ be a sequence of biharmonic maps from $(M_j, g_j)$ to $N$, satisfying

$$\int_{M_j} |\nabla_{g_j}^2 u_j|^2 + |\nabla_{g_j} u_j|^4 dV_{g_j} \leq \Lambda$$

for some $\Lambda > 0$. Without loss of generality, we assume that there is only one ALE bubble manifold and there is at most one bubble in each case in the above blow up analysis, then
(up to a subsequence) we have

\[ \int_{M_j} |\nabla^2 g_{j} u_j|^2 + |\nabla g_{j} u_j|^4 dV_{g_j} = \int_{M_\infty} |\nabla^2 g_{\infty} u_{\infty}|^2 + |\nabla g_{\infty} u_{\infty}|^4 dV_{g_{\infty}} + E_{bi}(\omega^b) + \sum_{a=A,B,C} E_{bi}(\omega^a), \]

where \( E_{bi}(\cdot) \) is the biharmonic energy of the corresponding bubble map. Moreover the images of all the bubbles and the image of \( u_\infty \) are all connected.

By the analysis discussed before, to prove the above main theorem, we need only to show that the biharmonic energy of the maps \( u_j \) over the neck regions \( I_j^\alpha \) is converging to 0 and the image of \( u_j \) over \( I_j^\alpha \) is converging to a point. More precisely, we will prove the following theorem for the extrinsic case in Section 6 (the intrinsic case will handled in Section 7).

**Theorem 4.11.** Let \( I_j^\alpha = A_{\epsilon_j}^{\alpha-\epsilon_j}(x_{a,j}) \) be the annulus region in \( M_j \) corresponding to \( I_j^\alpha \) in Proposition 4.6, then

\[ \lim_{\delta \to 0} \lim_{R \to \infty} \lim_{j \to \infty} \int_{I_j^\alpha} |\nabla^2 g_{j} u_j|^2 + |\nabla g_{j} u_j|^4 dV_{g_j} = 0 \]

and

\[ \lim_{\delta \to 0} \lim_{R \to \infty} \lim_{j \to \infty} \text{osc}_{I_j^\alpha} u_j = 0. \]

The proof of the above theorem is the most difficult part in this paper.

5. GEOMETRY OF THE DEGENERATING NECK REGIONS

To prove Theorem 4.11, we need to investigate the asymptotic behavior of biharmonic maps \( u_j \) over the neck regions \( I_j^\alpha \), which are sub-annuli of the degenerating neck regions \( A_{r_j R, \delta}(x_{a,j}) \subset M_j \). Now the main difficult issue is that we do not have satisfactory coordinates on those degenerating neck regions \( A_{r_j R, \delta}(x_{a,j}) \) at this stage. Recall that \( A_{\delta_2, \delta_1}(\cdot) \) represents \( B(\cdot, \delta_1) \setminus B(\cdot, \delta_2) \). In this section, we shall explore the refined geometric picture of the degenerating neck region \( A_{r_j R, \delta}(x_{a,j}) \) and construct two types of *good global coordinates* on this region: (\( x \)) and (\( y \)). In the coordinates (\( x \)), the metrics \( g_j \) are \( C^0 \) close to the flat metric. While, in the coordinates (\( y \)), the metrics \( g_j \) are \( C^4 \) close to the flat metric in some weighted function space.

**Remark 5.1.** The arguments in this section are valid in higher dimensions in the domain and hence the scheme developed in this paper can be applied to geometric systems defined over non-collapsed degenerating Einstein manifolds of higher dimensions.

Firstly, we recall the definition of weighted Hölder space. The weighted Hölder space \( C_{\eta}^{k,\alpha}(A_{r_1, r_2}) \) (\( 0 < \alpha < 1 \)) with weight \( \eta \) (a function of \( |x| = r \)) are the space of functions in \( C_{\eta}^{k,\alpha}(A_{r_1, r_2}(0)) \) with bounded weighted norm \( ||f||_{C_{\eta}^{k,\alpha}} \), which is defined as follows:

\[ ||f||_{C_{\eta}^{k,\alpha}} = \sum_{j=0}^{k} \sup_{A_{r_1, r_2}(0)} \eta^{-1} |x|^j |D^j f| + \sup_{x \neq y} \min \left( \eta^{-1} |x|^{k+\alpha}, \eta^{-1} |y|^{k+\alpha} \right) \frac{|D^k f(x) - D^k f(y)|}{|x - y|^{\alpha}}. \]
Analogous weighted norms for spaces of functions on $\mathbb{R}^n \setminus B_R$ were defined in [35] (page 252-253). By similar arguments as in the proof of Theorem 1 in [35], we have the following Schauder’s estimates for linear elliptic equations:

**Theorem 5.2.** Set

$$L(u) = \sum_{k,l} a_{kl}(x) \partial_k \partial_l u + \sum_k a_k(x) \partial_k u + a(x) u,$$

where $a_{k,l} \in C^{p,\alpha}_{r_1}(A_{r_1,r_2})$, $a_k \in C^{p,\alpha}_{r_1}(A_{r_1,r_2})$, $a \in C^{p,\alpha}_{r_1}(A_{r_1,r_2})$ with norms less than $K_1$. Furthermore, we assume

$$\sum_{k,l} a_{kl}(x) \xi_k \xi_l \geq K_2 \sum_k \xi_k^2 \quad (K_2 > 0).$$

Let $u$ be a solution of $L(u) = f$, then we have

$$||u||_{C^{\eta+2,\alpha}} \leq \hat{C} \hat{K}(||u||_{C^{\eta}} + ||f||_{C^{p,\alpha}}),$$

where $\hat{C} > 0$ is defined by $\eta(s) \leq \hat{C} \eta(r)$ for $1/2r < s < 3/2r$, and $\hat{K} > 0$ depends only on $K_1$, $K_2$, $p$, $i$, $\alpha$, and $n$.

We put the proof in Appendix B.

The main result in this section is the following:

**Theorem 5.3.** For $\delta_1 > 0$ small enough and for $R > 0$ and $j \geq 0$ sufficiently large, there exist two coordinates $(x)$ and $(y)$ on $\tilde{A}_{r_j,R,\delta_1}(x_{a,j}) \subset M_j$ such that the following properties hold: in coordinates $(x)$, there hold

$$||g_{j,kl}(x) - \delta_{kl}||_{C^{0}} < \eta_j(|x|),
|x(\cdot)| \equiv r(\cdot) = d_{g_{j,\cdot}}(\cdot, x_{a,j}) \quad \text{and} \quad \nabla \partial_r \partial_r = 0,$$

and in coordinates $(y)$, there hold

$$||g_{j,kl}(y) - \delta_{kl}||_{C^{4,\alpha}} < C,$$

where $\eta_j(t) = O(r^{-\frac{\varepsilon_5}{t\varepsilon_5}} + (r_jR_0)^{\varepsilon_5 t^{-\varepsilon_5}})$, $r_{\infty} > 0$, $\varepsilon_5 > 0$, $r_j > 0$ are the same as in the curvature estimate in Proposition 2.6, and $R_0 > 0$ is some fixed large number. Moreover,

$$||y - x||_{C^{4,\alpha}_{\{|y|\eta_j(|y|)}} < C,$

in particular, $|x - y| < \eta_j(|x|)|x|$.

**Remark 5.4.** Here we say coordinates on the degenerating neck region, we mean a map

$\varphi = \mathcal{L}^{-1} \circ \text{proj} : \mathbb{R}^4 \rightarrow A_{r_j,R,\delta_1}(x_{a,j}),$

where $\mathcal{L}$ is a diffeomorphism from $A_{r_j,R,\delta_1}(x_{a,j})$ to $\mathbb{R}^4 / \Gamma$, and $\text{proj}$ is the natural projection from $\mathbb{R}^4$ to $\mathbb{R}^4 / \Gamma$.

**Remark 5.5.** The above theorem says that, in the $(y)$ coordinates, the metric on the degenerating neck region $A_{r_j,R,\delta_1}(x_{a,j}) \subset M_j$ is close in $C^4$ topology (in fact $C^k$ for any $k \in \mathbb{N}$ by the argument given in the proof) to the flat metric in some weighted sense. It is easy to see that $\eta_j$ satisfies the requirement in Theorem 5.2. We remark that the notation $\eta_j$ often
appears in the rest of the paper, it means a function of order $\eta_j$ or it serves as a positive upper bound.

To prove Theorem 5.3, we shall apply some arguments in [4] with some modifications. The proof will be completed in two steps. Firstly, we outline the construction of $C^1$-coordinates on the degenerating neck regions.

**Lemma 5.6.** One can construct on $A_{r_j R, \delta_1} (x_{a,j}) \subset M_j$ coordinates $(x)$ for small enough $\delta_1 > 0$ and sufficiently large $R > 0$ and $j > 0$, such that

$$\|g_{j,kl}(x) − \delta_{kl}\|_{C^0} < \eta_j(|x|).$$

**Proof.** Recall that $(M_j, r_j^{-2}g_{j,x_{a,j}})$ converges in pointed G-H topology (in $C^\infty$ topology outside some compact set) to $(M_a, h_a, p_a)$ (see Proposition 2.4 or [1]). The fact that $(M_a, h_a)$ is a Ricci flat ALE manifold (orbifold) gives the existence of the following diffeomorphism (see Section 2 of [27] or Lemma (3.1) of [4]):

$$\Pi_R : A_{\overline{R}, 2R}(p_a) \subset M_a \rightarrow C(S^3/\Gamma).$$

Moreover, the limit metric

$$\lim_{R \rightarrow \infty} R^{-2} \Pi_R^* h_a$$

is the standard flat metric on the flat cone $C(S^3/\Gamma)$. By taking the universal covering of $A_{R, 2R}(p_a) \subset (M_a, h_a)$, we assume $\Gamma = e$. Let us temporarily ignore the fundamental group.

For large $R > 0$, we set $\bar{S}_R := \Pi_R^{-1}(\{1\} \times S^3)$. Hence

$$\|R \bar{A}_R + Id\| \leq a(R),$$

where $\bar{A}_R$ stands for the shape operator of $\bar{S}_R$, and $a(R) \rightarrow 0$, as $R \rightarrow \infty$.

Notice that $r_j^{-1}A_{r_j R, 2r_j R}(x_{a,j})$ converge smoothly to $A_{\overline{R}, 2R}(p_a) \subset M_a$, it is easy to see that $\partial B(x_{a,j}, r_j R)$ is smooth, and $(R)^{-1}\bar{S}_R$ and $(r_j R)^{-1}\partial B(x_{a,j}, r_j R)$ are closed in $C^\infty$ topology. Hence $\partial B(x_{a,j}, r_j R) \equiv S_{r_j R}$ is diffeomorphic to $S^3$, and

$$\lim_{R \rightarrow \infty} \lim_{j \rightarrow \infty} (r_j R)^{-1} S_{r_j R} = (S^3, g_{S^3})$$

in $C^k$ topology, where $g_{S^3}$ is the standard round sphere metric. And

$$\|((r_j R)A_{r_j R} + Id\| < a(j, R, \sigma) \rightarrow 0$$

by (36) and the arguments at the beginning of the proof. Denote the diffeomorphism by

$$\Upsilon_{r_j R} : \bar{\theta} \in S_{r_j R} \rightarrow \theta \in S^3.$$  

Motivated by the construction of coordinates around isolated orbifold singularity in Page 342 of [4], we consider a map

$$\Phi_{r_j R} : S^3 \times [r_j R, \delta_1] \rightarrow M_j$$

$$(\theta, t) \mapsto \exp\{(t - r_j R)\nu_{r_j R}(\bar{\theta})\} = \exp\{(t - r_j R)\nu_{r_j R}(\Upsilon_{r_j R}(\theta))\}$$

where $\nu_{r_j R}(\bar{\theta})$ is outward unit normal vector.

**Proposition 5.7.** $\Phi_{r_j R}$ is an into-diffeomorphism for some fixed constant $\delta_1 > 0$ independent of $j$ and $R$.  

The above proposition will be proved soon afterward, now we turn back to the construction of coordinates.

Taking a point \( \hat{\theta} \in S_{r_j R} \), we consider parallel orthonormal vector fields \( E_1(t), \ldots, E_4(t) \) along \( \sigma_{r_j R}(t) = \Phi_{r_j R}(\theta, t) \). Let \( X_{r_j R}(t) \) be a Jacobi field (see Section 2 of [4]) along \( \sigma_{r_j R}(t) \) with

\[
\begin{align*}
X_{r_j R}(r_j R) &= r_j R E_k(r_j R), \\
\nabla_{\partial t}X_{r_j R}^\perp(r_j R) &= -r_j R A_{r_j R}(E_k^\perp(r_j R)), \\
\nabla_{\partial t}(X_{r_j R} - X_{r_j R}^\perp)(r_j R) &= E_k(r_j R) - E_k^\perp(r_j R)
\end{align*}
\]

for some \( k = 1, \ldots, 4 \). Here \( A_{r_j R} \) stands for the shape operator of \( S_{r_j R} \), \( \nu_{r_j R} \perp X_{r_j R}^\perp \). Then \( X_{r_j R}^i(t) = (X_{r_j R}(t), E_i(t)) \) satisfies (see Section 5 of [49] or Section 3 of [4])

\[
X_{r_j R}^i(t) = X_{r_j R}^i(r_j R) + (\partial_t X_{r_j R}^i)(r_j R)(t - r_j R) + \int_{r_j R}^{t} \int_{r_j R}^{s} X_{r_j R}^i(u) K_i^i(u) \, duds,
\]

where \( K_i^i(u) = (\text{Riem}(E_i(u), (\partial_t \sigma_{r_j R})(u))(\partial_t \sigma_{r_j R})(u), E_i(u)) \).

Thanks to the curvature estimate

\[
r^2 |R_{g_j}| \leq C_7 \max \left\{ \frac{(r_j)^{\varepsilon_5}}{r}, \frac{(r_{\infty})^{\varepsilon_5}}{r_{\infty}} \right\}.
\]

in Proposition 2.6, we find by direct calculations using (37) and (39) that

\[
|X_{r_j R}^i(t)| < Ct, \\
(\partial_t X_{r_j R}^i)(r_j R)t = \delta_k^i(1 + a(j, R)) t,
\]

\[
|\int_{r_j R}^{t} \int_{r_j R}^{s} X_{r_j R}^i(u) K_i^i(u) \, duds| < C \max \left\{ \varepsilon_5 r_j R^{1-\varepsilon_5} + t R^{-\varepsilon_5} - \frac{t^2}{4} \varepsilon_5, \left[ \varepsilon_5 \cdot \frac{t^{1+\varepsilon_5}}{r_{\infty}^{\varepsilon_5}} \right] \right\}
\]

for \( t \in [r_j R, \delta_1] \). Noticing that

\[
a(j, R) \to 0, \text{ as } R, j \to \infty
\]

by (37), we know that the terms \( a(j, R)t \) and \( t R^{-\varepsilon_5} \) have to cancel each other out. Indeed

\[
(\partial_t X_{r_j R}^i)(N r_j R) = (\partial_t X_{r_j R}^i)(r_j R) + \int_{r_j R}^{N r_j R} X_{r_j R}^i(u) K_i^i(u) \, du
\]

\[
= (\partial_t X_{r_j R}^i)(r_j R) + O(R^{-\varepsilon_5} - (N R)^{-\varepsilon_5}),
\]

where \( N > 0 \) is some large number. Therefore we obtain the following key estimate

\[
|t^{-1} X_{r_j R}^i(t) - \delta_{ik}| \leq C(r_j R t^{-1}) + C(r_j R^{1-\varepsilon_5} t^{-1}) + C \max(r_{\infty}^{-\varepsilon_5} t^{\varepsilon_5}, (r_j^{\varepsilon_5} t^{-\varepsilon_5})
\]

Now fix \( R = R_0 \) in (41) for some sufficiently large \( R_0 \). Then for \( t \in [r_j R, \delta_1] \), we have

\[
|t^{-1} X_{r_j R_0}^i(t) - \delta_{ik}| \leq \eta_j(t),
\]

where \( \eta_j(t) = O(r_{\infty}^{-\varepsilon_5} t^{\varepsilon_5} + (r_j R_0)^{\varepsilon_5} t^{-\varepsilon_5}) \).
The above estimate means that in the coordinates
\[(42) \quad \Phi_{r, R_0}: S^3 \times [r, R, \delta_1] = \{x \in \mathbb{R}^4 : r \leq |x| \leq \delta_1 \} \to A_{r, R, \delta_1}(x_{a, j}) \subset M_j,\]
the metric \(g_j\) satisfies
\[|((\Phi_{r, R_0}^*) g_j)_{kl} - \delta_{kl}| \leq \eta_j(|x|).\]

We just outline it, for details one can see Step 2 in the proof of Theorem (1.1) in [4]. In fact from (41) we have that the shape operator \(A_t, t \in [r, R, \delta_1] \) of \(S_t = \partial B(x_{a, j}, t)\) satisfies
\[(43) \quad |t \bar{A}_t + Id| \leq \eta_j(t).\]

The standard Euclidean metric on \(S^3 \times [r, R, \delta_1]\) can be written as
\[dr^2 + r^2 d\theta^2.\]
So from (43) the metric \(g_j\) in \((x)\) coordinates takes the form
\[dr^2 + r^2 (1 + \eta_j(r))^2 d\theta^2.\]
Hence
\[|g_{j, kl}(x) - \delta_{kl}| \leq \eta_j(|x|).\]
\[\square\]

**Remark 5.8.** From the above construction, we know that \(|x(\cdot)| = r(\cdot) = d_{g_j}(\cdot, x_{a, j})\) and \(\nabla_{\partial_r} \partial_r = 0.\)

The final step for the proof of the Theorem 5.3 is to improve the regularity of the metric on the basis of Lemma 5.6.

**Lemma 5.9.** We can construct coordinates \((y)\) on \(A_{r, R, \delta_1}(x_{a, j})\) such that
\[\|g_{j, kl}(y) - \delta_{kl}\|_{C^4_{a, j}} < C.\]
This estimate says that if we introduce the coordinates \(\tilde{y} = \rho^{-1} y\) on \(A_{\rho, 2\rho}(x_{a, j}) \subset M_j,\) then
\[\|g_{j, kl}(\rho \tilde{y}) - \delta_{kl}\|_{C^4(B_2 \setminus B_1)} < \eta_j(\rho)\]
for all \(\rho \in [r, R, \delta_1].\)

**Proof.** As in the proof of Lemma 5.6, at very point \(\tilde{\theta} \in S, r, R_0,\) we consider parallel orthonormal vector fields \(E_1(t, \tilde{\theta}), \ldots, E_4(t, \tilde{\theta})\) along \(\sigma_{r, R_0}(t) = \Phi_{r, R_0}(\tilde{\theta}, t)\) which satisfy
\[r_j R_0 \nu_{r, R_0}(\tilde{\theta}) = x^i(r_j R_0, \tilde{\theta}) E_i(r_j R_0, \tilde{\theta}).\]
Here \((x) = (r, \theta)\) are the coordinates constructed in Lemma 5.6. Then
\[r \partial_r = x^i E_i(r)\]
along \(\sigma_{r, R_0},\) since both \(E_i(t)\) and \(\partial_r\) are parallel invariant and \(|x| = r.\)

Now we take geodesics \(\sigma_1, \ldots, \sigma_K (\sigma_k(t) = \Phi_{r, R_0}(\theta_k, t), \gamma_{r, R_0}(\theta_k) = \tilde{\theta}_k)\) such that
\[\{B(\sigma_k(t_l); \delta t_l)\}_{k=1, \ldots, K; l=1, \ldots}\]
cover \(A_{r, R, \delta_1}(x_{a, j}).\) We denote \(\{B(\sigma_k(t_l); \delta t_l)\}\) by \(B_{k, l}.\)
For all \( k \) and \( l \), we take harmonic coordinates associated to \( E_1(t, \tilde{\theta}_k), \cdots, E_4(t, \tilde{\theta}_k) \)
\[
\mathbf{H}_{k,l} : B_{k,l} \to \mathbb{R}^4
\]
as in [4]. Moreover we set \( \mathbf{H}_{k,l}(\sigma_k(t)) = (\theta_k, t) \). By means of the arguments in Step 3 in the proof of Theorem (1.1) in [4] with some minor modifications, we have
\[
||(H_{k,l}^* g_j)_{mn} - \delta_{mn}||_{C^{1,\alpha}_{(\eta_j)}(H_{k,l}(B_{k,l}))} \leq C,
\]
where \(|z|\) is the absolute value of \( \mathbf{H}_{k,l} \) coordinates functions. Then by elliptic estimates (see Schauder’s estimates for elliptic equations in Theorem 5.2), we know that in the above coordinates, the metric satisfies
\[
||(H_{k,l}^* g_j)_{mn} - \delta_{mn}||_{C^{2,\alpha}_{(\eta_j)}(H_{k,l}(B_{k,l}))} \leq C
\]
for all \( p \in \mathbb{N} \). Since Einstein metrics satisfy an elliptic equation in harmonic coordinates. In fact in harmonic coordinates, the equation for the metric takes the form [4]
\[
\triangle_g g_{kl} = -2 \text{Ric}_{kl} + Q_{kl}(g, \partial g).
\]
For the completeness of the argument, we outline the proof of (44). Let \( \omega_1 \) be the frame on \( B(\sigma_j R_0(t); 2\delta t) \) obtained by radial parallel transport of the dual frame of \( \{E_1, \cdots, E_4\} \) defined along \( \sigma_j R_0(t) \) as above. By estimates for harmonic coordinates in Fact (2.9) in [4], we have \( (\lambda < 1) \)
\[
|dH_t - \omega_1| \leq C \sup_{B(\sigma_j R_0(t); \lambda \delta t)} |Rm| \left( \lambda \delta t \right)^2 \text{ on } B(\sigma_j R_0(t); \lambda \delta t),
\]
and then by the curvature estimates in Proposition 2.6 (see (40)), we get
\[
|dH_t - \omega_1| \leq \eta_j(t) \text{ on } B(\sigma_j R_0(t); \lambda \delta t),
\]
where \( H_t \) stand for the harmonic coordinates associated to the frame \( \{E_1, \cdots, E_4\} \). Now we denote \( \Phi_{\sigma_j R_0}(\theta, t\xi) \) by \( \Phi_t(\theta, \xi) \). The last estimate shows that
\[
|d(t^{-1}H_t \circ \Phi_t) - t^{-1}\Phi_t^* \omega_1| \leq \eta_j(t), \text{ on } B((\theta, 1); \lambda \delta) \subset S^3 \times (1/2, 3/2).
\]
From the proof of Lemma 5.6, it is easy to see that
\[
|t^{-1}\Phi_t^* \omega_1(\theta, \xi) - (dz_1, \cdots, dz_4)| \leq \eta_j(t),
\]
where \( (dz_1, \cdots, dz_4) \) form a basis of cotangent bundle of \( \mathbb{R}^4 \). Therefore
\[
|d(t^{-1}H_t \circ \Phi_t) - (dz_1, \cdots, dz_4)| \leq \eta_j(t).
\]
Since \( (H_t \circ \Phi_t)(\theta, 1) = 0 \in \mathbb{R}^4 \), we get
\[
|(t^{-1}H_t \circ \Phi_t) - (z_1, \cdots, z_4)| \leq \eta_j(t).
\]
Thanks to the estimate
\[
|(\Phi_t^* g_j)_{kl} - \delta_{kl}| \leq \eta_j(t) \text{ on } B((\theta, 1); \lambda \delta) \subset S^3 \times (1/2, 3/2),
\]
induced from Lemma 5.6, and notice that
\[
(t^{-1}H_t \circ \Phi_t)^* (\Phi_t^* g_j) = (t^{-1}H_t)^* g_j,
\]
we can get
\[ |t^{-1}(H_t)g_j - \delta_{mn}| < \eta_j(t) \quad \text{on } t^{-1}H_t(B(\sigma_{r_j}R_0(t); 2\delta t)). \]

Next by elliptic estimates [19], we obtain
\[ ||t^{-1}(H_t)g_j - \delta_{mn}||_{C^{1,\alpha}} < \eta_j(t) \quad \text{on } t^{-1}H_t(B(\sigma_{r_j}R_0(t); 2\delta t)). \]

Therefore we get (44).

Next by similar arguments as in Step 4 in the proof of Theorem (1.1) in [4], if the intersection \( B_{k_1,l_1} \cap B_{k_2,l_2} \) is not empty, then
\[
||H_{k_1,l_1} \circ H_{k_2,l_2} -Id||_{C^{p+1,\alpha}_{|z|\eta_j(|z|)}}(H_{k_2,l_2}(B_{k_1,l_1} \cap B_{k_2,l_2})) \leq C,
\]
where \(|z|\) is the absolute value of \( H_{k_2,l_2} \) coordinates functions. Indeed by the above construction, \( t^{-1}_l H_{k,l} \circ \Phi_{\theta} \) are close to the restriction of the standard coordinates functions on \( \mathbb{R}^4 \) (we denote it by \( H_\infty \)) for all \( k \) and \( l \), since every flat bundle on \( S^3 \) is trivial. If \( B_{k_1,l_1} \) intersects \( B_{k_2,l_2} \) (assume \( \frac{t_{l_1}}{t_{l_2}} = \beta^{-N} < 1 \) for some \( \beta > 1 \) and \( N \in \mathbb{N} \)), then by (46), we get
\[
|t^{-1}_{l_1}H_{k_1,l_1} \circ \Phi_{r_jR_0}(\theta, \xi t_{l_1}) - t^{-1}_{l_2}H_{k_2,l_2} \circ \Phi_{r_jR_0}(\theta, \xi t_{l_2})| \\
\leq \eta_j(t_{l_1}) + \beta^{N}t^{-1}_{l_2}H_{k_2,l_2} \circ \Phi_{r_jR_0}(\theta, \beta^{-N}\xi t_{l_2}) - H_\infty(\theta, \xi t_{l_1}) \\
\leq \eta_j(t_{l_1}) + \eta_j(t_{l_2}) \leq C \eta_j(t_{l_1}).
\]

Thus (47) can be derived by apriori estimates for harmonic functions (see Theorem 5.2). Since all components of \( H_{k_1,l_1} \circ H_{k_2,l_2} \) are harmonic functions with respect to \( g_j \). And Moreover by applying Theorem 5.2 again, it actually holds for all \( p \in \mathbb{N} \) that
\[
||H_{k_1,l_1} \circ H_{k_2,l_2} -Id||_{C^{p+1,\alpha}_{|z|\eta_j(|z|)}}(H_{k_2,l_2}(B_{k_1,l_1} \cap B_{k_2,l_2})) \leq C.
\]

Now we take a partition of unity \( \rho_{k,l} \) associated to the covering \( \{B_{k,l}\} \) so that if \( B_{k_1,l_1} \) intersects with \( B_{k_2,l_2} \), then
\[
||\rho_{k_1,l_1} \circ H_{k_2,l_2}^{-1}||_{C^{p+1,\alpha}_{|z|\eta_j(|z|)}}(H_{k_2,l_2}(B_{k_1,l_1} \cap B_{k_2,l_2})) \leq C.
\]

For such a construction, we refer to Page 325 of [4].

Finally, we define a smooth map \( \Psi : A_{r_jR_0}(x_{a,j}) \to \mathbb{R}^4 \) by
\[
\Psi(\cdot) = \sum_{k,l} \rho_{k,l}(\cdot)H_{k,l}(\cdot).
\]

Then we have
\[
||\Psi \circ H_{k_1,l_1}^{-1} - Id||_{C^{p+1,\alpha}_{|z|\eta_j(|z|)}}(H_{k_1,l_1}(B_{k_1,l_1})) \\
= ||\sum_{k,l} \rho_{k,l} \circ H_{k_1,l_1}^{-1} H_{k,l} \circ H_{k_1,l_1}^{-1} - Id||_{C^{p+1,\alpha}_{|z|\eta_j(|z|)}}(H_{k_1,l_1}(B_{k_1,l_1})) \leq C.
\]

Therefore \( \Psi \) defines coordinates \( (y) \) such that
\[
||g_{j,k}(y) - \delta_{kl}||_{C^{p,\alpha}_{|z|\eta_j(|z|)}} < C.
\]

The proof is finished.
Remark 5.10. By the above construction, $H_\infty$ in the proof is the restriction of the identity map, so
$$\|y - x\|_{C^1_0(|y|)} < C,$$
in particular $|x - y| < \eta_j(|x|)|x|$, where $|x(t)| = r = d_{g_j}(\cdot, x_{a,j})$. In fact, the map $\Phi_{g_jR_0}(\theta, t)$ define the coordinates $(x)$, $\Phi_{g_jR_0}(\theta, t\xi) = \Phi_t(\theta, \xi)$, $t \to \Psi_{k_i} \circ \Phi_{t} \to H_\infty$; the map $\Psi$ define the coordinate $(y)$, and $\Psi \circ H_{k_i}^{-1}$ is closed to $Id$.

Now we give the proof of Proposition 5.7. We follow the arguments in Step 1 of the proof for Theorem (1.1) in [4].

Proof. Let $J$ be a nontrivial Jacobi field along the geodesic $\sigma_{r,R}(t) = \Phi_{r,R}(\theta, t)$ with $J(r,R) \in T_{p_{S_{r,R}}} \mathbb{H}$ and $\frac{\partial}{\partial t} J(r,R) = A_{r,R}(J(r,R))$. The first key observation is that for large $R$, and for some small enough $\delta_1$, $J(t)$ never vanishes as long as $d_{g_j}(\sigma_{r,R}(t), S_{r,R}) = t - r_{r,R}$. Indeed, let $y(t)$ be the solution of the ordinary differential equation:
$$\begin{align*}
\frac{d^2}{dt^2} y(t) + C \max \left( \frac{r_j^{\varepsilon_5}}{t^{2+\varepsilon_5}}, \frac{t^{2+\varepsilon_5}}{r_j^{\varepsilon_5}} \right) y(t) &= 0, \quad \forall t \in [r,R, \delta_1), \\
y(r,R) = 1, \quad \frac{dy}{dt} (r,R) &= (r,R)^{-1} (1 - a(j,R)),
\end{align*}$$
where a constant $C$ is taken so that the sectional curvature of $M_j$ at $\sigma_{r,R}(t)$ is bounded from above by
$$C \max \left( \frac{r_j^{\varepsilon_5}}{t^{2+\varepsilon_5}}, \frac{t^{2+\varepsilon_5}}{r_j^{\varepsilon_5}} \right).$$
See the curvature estimate in (40) and the meaning of $a(j,R)$ in (37).

If $y(t)$ is positive on $[r,R, \delta_1]$, then $\frac{dy}{dt}(t) \leq (r,R)^{-1} (1 - a(j,R))$, since $\frac{dy}{dt}(t) \leq 0$. Thus
$$\begin{align*}
\frac{dy}{dt}(t) &= (r,R)^{-1} (1 - a(j,R)) - C \int_{r,R}^{t} \max \left( \frac{r_j^{\varepsilon_5}}{s^{2+\varepsilon_5}}, \frac{s^{2+\varepsilon_5}}{r_j^{\varepsilon_5}} \right) y(s) ds \\
&\leq (r,R)^{-1} (1 - a(j,R)) - C \int_{r,R}^{t} \max \left( \frac{r_j^{\varepsilon_5}}{s^{2+\varepsilon_5}}, \frac{s^{2+\varepsilon_5}}{r_j^{\varepsilon_5}} \right) \left( 1 + \frac{s}{r_j(R)} (1 - a(j,R)) \right) ds.
\end{align*}$$

Therefore direct computations show that if we choose $R$ large and $\delta_1$ small, $\frac{dy}{dt}(t)$ is positive on $[r,R, \delta_1]$, and hence so is $y(t)$. Then applying the comparison theorem on Jacobi fields to our situation, we have
$$y(t) \leq |J(t)|/|J(r,R)|, \quad \text{as long as } d_{g_j}(\sigma_{r,R}(t), S_{r,R}) = t - r_{r,R}.$$This shows that $S_{r,R}$ has no focal points as long as $d_{g_j}(\sigma_{r,R}(t), S_{r,R}) = t - r_{r,R}$.

Suppose $S_{r,R}$ has the cut locus $C_{r,R}^+$ outside $B(x_{a,j}, r_{r,R})$. Let $q$ be a point of $C_{r,R}^+$ which is closest to $S_{r,R}$, i.e., $d_{g_j}(q, S_{r,R}) = d_{g_j}(C_{r,R}^+, S_{r,R})$, and $\sigma_{r,R} : [r,R, \delta_1] \to M_j$ a geodesic such that $d_{g_j}(\sigma_{r,R}(t), S_{r,R}) = t - r_{r,R} < t_{0}$ on $[r,R, r_{r,R} + t_{0}]$, $t_{0} = d_{g_j}(q, S_{r,R})$ and $\sigma_{r,R}(r_{r,R} + t_{0}) = q$.

Then since $S_{r,R}$ has no focal points along $\sigma_{r,R}|_{r,R, r_{r,R} + t_{0}}$, it turns out from the standard theory of Riemannian geometry (Lemma 2 in P.226 of [20], Theorem (1.5) in Chap.4 of [37]) that $\sigma_{r,R}$ satisfies:
$$d_{g_j}(\sigma_{r,R}(t), S_{r,R}) = r_{r,R} + 2t_{0} - t, \quad \text{on } [r_{r,R} + t_{0}, r_{r,R} + 2t_{0}].$$
Now we show that \( C^+_{r_jR} \cap A_{r_jR, \delta_1}(x_{a,j}) \) must be empty for large \( R \) and small \( \delta_1 \). In fact if there are a divergence sequences \( \{R_k\} \) and a sequence \( \{\delta_k\} \) converging to 0 such that \( C^+_{r_jR_k} \cap A_{r_jR_k, \delta_k}(x_{a,j}) \) is not empty for any \( k \), then we can take \( t_k = \delta_k \) and

\[
\sigma_{r_jR_k} : [r_jR_k, r_jR_k + 2t_k] \to M_j
\]

with \( t_k = d_{g_j}(C^+_{r_jR_k}, S_{r_jR_k}) \) as above.

We claim that

\[
\lim_{k \to \infty} \lim_{j \to \infty} \frac{t_k}{r_jR_k} = \infty.
\]

Then the metric space

\[
\left\{ x \in M_j : 1/2 \leq (t_k + r_jR_k)^{-1}d_{g_j}(x; x_{a,j}) \leq 2 \right\}, (t_k + r_jR_k)^{-1}d_{g_j})
\]

converges to \( \{ x \in \mathbb{R}^4 : 1/2 \leq |x| \leq 2 \} \) in \( C^{1,\alpha} \)-topology, as \( j \to \infty \) firstly and then \( k \to \infty \) (note that the curvature estimate in Proposition 2.6 is enough to ensure this convergence by Gromov compactness theorem, see for example Theorem 2.2 in [1]). And the geodesic \( \sigma_{r_jR_k} \) converges to a line in \( \mathbb{R}^4 \) from \( x_0 \) with \( |x_0| = 1/2 \) to \( \hat{x}_0 \) with \( |\hat{x}_0| = 1/2 \) passing through a point \( x \) with \( |x| = 1 \). This is a contradiction. So the proof of Proposition 5.7 is completed.

Finally, we prove the above claim. Since \( (M_j, r_j^{-2}g_j, x_{a,j}) \) converges in pointed G-H topology to the Ricci flat ALE manifold \((M_a, h_a, p_a)\), and moreover \( (M_j \setminus B(x_{a,j}; r_jR_k), r_j^{-2}g_j, x_{a,j}) \) converges in \( C^0 \) topology to \( (M_a \setminus B(p_a, R_k), h_a) \), \( \sigma_{r_jR_k} \) converges to a geodesic segment \( \sigma_{R_k} : [R_k, R_k + 2t_k] \to M_a \) satisfying

\[
d_{h_a}(\sigma_{R_k}(t), S_{R_k}) = R_k + 2t_k - t, \quad \text{on } [R_k + t_k, R_k + 2t_k].
\]

By the arguments in subsection 2.2 of [28], we have \( t_k \geq CR_k^{1+\varepsilon} \) for some \( \varepsilon > 0 \), hence the claim is valid. \( \square \)

6. Neck analysis on degenerating neck regions

In this section, we shall analyze the behavior of biharmonic maps \( u_j \) defined on the degenerating neck regions \( I_j^\alpha \) and then prove Theorem 4.11. As in Section 5, for simplicity of notations, we shall ignore the effect of the fundamental group of \( A_{r_jR, \delta_1}(x_{a,j}) \) when we carry out the neck analysis. We remark again that the notation \( \eta_j \) often appears in the rest of the paper, it means a function of order \( \eta_j \) or it serves as a positive upper bound, it’s precise expression may differ line by line.

To begin with, we shall firstly recall the following three key results given in the previous sections. In Proposition 4.6, we denote the set \( I_j^\alpha \) to be the cylinder corresponding to \( \tilde{I}_j^\alpha = A_{e_j^{\alpha-2}, e_j^{\alpha+1}}(x_{a,j}) \subset M_j \).

\((\mathcal{R}_1)\) According to the bubble-neck decomposition in Proposition 4.6, \( \tilde{I}_j^\alpha \subset A_{r_jR, \delta_1}(x_{a,j}) \), and

\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{j \to \infty} \omega(u_j, I_j^\alpha) = 0,
\]
where
\[ \omega(u_{j}, I_{j}^\alpha) \equiv \sup_{t \in [v_{j}^\alpha, \alpha_{j}^{\alpha+1}-1]} \int_{A_{t, x_{a,j}}(x_{a,j})} |\nabla_{g_{j}}^{2} u_{j}|^{2} + |\nabla_{g_{j}} u_{j}|^{4}dV_{g_{j}}. \]

(\mathcal{R}_2) By Theorem \ref{5.3}, for \( \delta > 0 \) small enough and for \( R > 0 \) and \( j > 0 \) sufficiently large, there exist coordinates \( (x) \) and \( (y) \) on \( A_{r_{j}, R}(x_{a,j}) \subset M_{j} \) such that
\[ \|g_{j,kl}(x) - \delta_{kl}\|_{C^{0}} < \eta_{j}(|x|), \]
\[ |x(\cdot)| \equiv r(\cdot) = d_{g_{j}}(\cdot, x_{a,j}) \quad \text{and} \quad \nabla_{\partial_{y}} \partial_{r} = 0, \]
and
\[ \|g_{j,kl}(y) - \delta_{kl}\|_{C^{4,\alpha}_{\eta_{j}(|y|)}} < C, \]
where \( \eta_{j}(t) = O(r_{-\varepsilon_{5}} t^{\varepsilon_{5}} + (r_{j}R_{0})^{\varepsilon_{5}} t^{-\varepsilon_{5}}) \), \( r_{\infty}, \varepsilon_{5} \), \( r_{j} \) are the same as in the curvature estimate in Proposition \ref{2.6}, and \( R_{0} > 0 \) is some fixed large number. Moreover,
\[ \|y - x\|_{C^{1}_{\eta_{j}(|y|)}} < C, \]
in particular \( |y - x| < \eta_{j}(|x|)|x|. \)

(\mathcal{R}_3) The proof of the main Theorem \ref{1.2} can be reduced to be the proof of Theorem \ref{4.11} by the arguments in Section 4.

6.1. Decay estimates for tangential part energy. The main task in this subsection is to show that similar decay estimates as in Lemma \ref{3.3} for biharmonic maps on those degenerating neck regions still hold. We remark that the decay along degenerating neck regions may not be as fast as in the case of a fixed domain shown in Lemma \ref{3.3}, however, it is already good enough for our purpose as we will see soon.

Suppose \( b_{j}^{\alpha} = -l_{j,a}^{\alpha} L, d_{j}^{\alpha+1} = -l_{j,a}^{\alpha} L \) for some universal constant \( L > 0 \). Set
\[ A_{l} = A_{j, a^{l} - (l-1)L} = B(x_{a,j}, e^{-(l-1)L}) \setminus B(x_{a,j}, e^{-L}), \quad l^{\alpha}_{j,a} \leq l < l^{\alpha}_{j,b}. \]
By (\mathcal{R}_2) in the above, for some \( \tilde{L} \),
\[ A_{l} \subset \tilde{A}_{l-1} \cup \tilde{A}_{l} \cup \tilde{A}_{l+1}, \]
where \( \tilde{A}_{l} = \{ q \in M_{j}; e^{-lL} \leq |y(q)| \leq e^{-(l-1)L} \} \). So if we can prove the energy decay on \( \tilde{A}_{l} \), then the energy decay on \( A_{l} \) follows easily.

Set \( s = |y| \),
\[ \tilde{F}_{l}(v_{j}) = \int_{\tilde{A}_{l}} \frac{v_{j}^{2}}{|y|^{4}}dy, \]
for any smooth functions \( v_{j} \) on \( I_{j}^{\alpha} \subset M_{j} \), and
\[ (u_{j})^{*}(s) = \frac{1}{|\partial B_{s}|} \int_{\partial B_{s}} u_{j}d\sigma, \]
where \( (y) \) are the coordinates constructed in Theorem \ref{5.3}, \( dy \) is the volume form of the flat metric, and \( d\sigma \) is volume form of the round sphere of radius \( s \). Here we use the \( (y) \) coordinates instead of the \( (x) \) coordinates for the reason that the metric coefficients of \( g_{j} \) in \( (x) \) coordinates are not regular enough for our neck analysis carried out later.
Our aim is to prove the following decay estimates.

**Theorem 6.1.** Let \( v_j = u_j - (u_j)^* \), and \( s = |y| = e^t \), then when \( j \) and \( R \) are large enough and \( \delta \) is small enough, for \( I_{j,a}^\alpha - 1 \leq l < I_{j,b}^\alpha + 1 \), we have that for some \( \theta > 0 \),

\[
F_l \leq C\varepsilon^2 \left( e^{-\theta(l-I_{j,a}^\alpha)}L + e^{-\theta(I_{j,b}^\alpha-l)\hat{L}} \right),
\]

and

\[
\int_{(-L,-(l-1)) \times S^3} \left( |\triangle g S^3 u_j|^2 + |\nabla g S^3 u_j|^2 + |\partial_t^\alpha \nabla g S^3 u_j|^2 + |\partial_t^\alpha \triangle g S^3 u_j|^2 \right) dtd\theta \\
\leq C\varepsilon^2 \left( e^{-\theta(l-I_{j,a}^\alpha)}L + e^{-\theta(I_{j,b}^\alpha-l)\hat{L}} \right).
\]

The fact (\( \mathfrak{R}_2 \)) indicates that the \((y)\) coordinates and the \((x)\) coordinates in Theorem 5.3 are close to each other in \( C^1 \) norm, so by the following simple fact

\[
\int_{A_j} \frac{1}{r^4} \left( |\triangle g S^3 u_j|^2 + |\nabla g S^3 u_j|^4 + \frac{1}{r^2} |\partial_r g \nabla g S^3 u_j|^2 + \frac{1}{r^4} |\partial_t \triangle g S^3 u_j|^2 \right) dx \\
\leq C \int_{A_{l-1} \cup A_{l} \cup A_{l+1}} \frac{1}{s^4} \left( |\triangle g S^3 u_j|^2 + |\nabla g S^3 u_j|^4 + \frac{1}{s^2} |\partial_s g \nabla g S^3 u_j|^2 + \frac{1}{s^4} |\partial_t \triangle g S^3 u_j|^2 \right) dy,
\]

we have

**Theorem 6.2.** Set \( r = |x| \), then for \( r \in \left[ e^{k_0^\alpha}, e^{s_0^\alpha+1} \right] \),

\[
\int_{A_{j,r,e^L r}} \frac{1}{r^4} \left( |\triangle g S^3 u_j|^2 + |\nabla g S^3 u_j|^4 + \frac{1}{r^2} |\partial_r g \nabla g S^3 u_j|^2 + \frac{1}{r^4} |\partial_t \triangle g S^3 u_j|^2 \right) dx \\
\leq C\varepsilon^2 \left[ \left( \frac{r}{e^{s_0^\alpha+1}} \right)^\theta + \left( \frac{e^{k_0^\alpha}}{r} \right)^\theta \right],
\]

where \( A_{j,r,e^L r} \) := \( \{ r \leq |x| \leq e^L r \} \subset M_j \).

The proof of the above theorem is parallel to the one for Lemma 3.3, we just outline it.

Notice that the fact (\( \mathfrak{R}_1 \)) in the above ensure that

\[
\int_{\tilde{A}_j} |\nabla g_{j,l}^2 u_j|^2 + |\nabla g_{j,l} u_j|^4 dV_{g_j} < \varepsilon^2
\]

for any given \( \varepsilon > 0 \), as long as \( j \) is large enough. And from the fact (\( \mathfrak{R}_2 \)), we know that for any \( \eta_0 > 0 \), there exists \( j_0 > 0 \), \( R_0 > 0 \) and \( \delta_0 > 0 \) such that for all \( \rho \in \left[ e^{k_0^\alpha}, e^{s_0^\alpha+1} \right] \), the metric \( g_j \) satisfies

(48) \[\| g_{j,k,l}(\rho y) - \delta_{k,l} \|_{C^0(B_2 \setminus B_1)} < \eta_0 \]

when \( j > j_0 \), \( R > R_0 \) and \( \delta < \delta_0 \), where \( \tilde{y} = \rho^{-1}y \). So that we can apply the \( \varepsilon \)-regularity Theorem 2.9 on the region \( \tilde{A}_l \). In the rest of paper, we shall mainly focus on the region \( I_{j}^\alpha \), we may not point this out when we have used the \( \varepsilon \)-regularity theorem.

Rewrite the biharmonic map equation as in the proof of Lemma 3.3 (see (19)) to put the affect of the metric into the tension term (or error term), we get the following key lemma:
Lemma 6.3. $v_j(y) = u_j - (u_j)^*$ is an $\eta_0$-approximate biharmonic function (w.r.t. the flat metric) in the sense of Definition 3.1, where $\eta_0 > 0$ is the constant in Theorem 3.4. That is
\[
\Delta^2 v_j(s, \theta) = a_1 \nabla \Delta v_j + a_2 \nabla v_j + a_3 v_j + a_4 v_j
\]
\[
+ \frac{1}{|\partial B_s|} \int_{\partial B_s} b_1 \nabla \Delta v_j + b_2 \nabla v_j + b_3 v_j + b_4 v_j d\sigma + h_j(y),
\]
where the coefficients satisfy the conditions in Definition 3.1. Moreover the corresponding $h_j$ satisfies
\[
|||y|^{4(1-1/p)} h_j||^{2}_{L^{p}(\tilde{\mathcal{A}})} \leq \epsilon^2 j^2 (|y|).
\]

Finally, argue as in Corollary 3.5, we can prove the decay estimates in Theorem 6.1.

Remark 6.4. From Theorem 5.3, $\eta_j(t) = O(r^{-\varepsilon_5} t^{\varepsilon_5} + (r_j R_0)^{\varepsilon_5} t^{-\varepsilon_5})$, so we can take $\vartheta = \min(1, 2\varepsilon_5)$ in Theorem 6.1 by the arguments in Corollary 3.5, where $\varepsilon_5 > 0$ is the constant in the curvature estimate in Proposition 2.6.

6.2. A new Pohožaev type identity. In this subsection, we shall develop new Pohožaev type arguments that can be applied to degenerating neck regions. The advantage of these new arguments is that they depend much less on the coordinate functions than those for a fixed domain in Subsection 3.2. The arguments here are much more complicated, due to the fact that we are not able not do analysis on the whole geodesic ball $B(x_{a,j}, r)$ and we need to take care of the degeneration of the metrics.

In the sequel, for simplicity of notation, we shall write $u$ for $u_j$ if there is no confusion. Set $A_{j,r, cr} \equiv B(x_{a,j}, cr) \setminus B(x_{a,j}, r)$. Let $(x)$ be the coordinates in Theorem 5.3. Then $|x| = r = d_{g_j}(\cdot, x_{a,j})$ and $\nabla_{\partial_r} \partial_r = 0$. Multiplying the extrinsic biharmonic map equation by $r \partial_r u$, by using integration by parts as in Subsection 3.2, we get that
\[
0 = \int_{A_{j,r, cr}} (r \partial_r u)(\Delta^2 g_j u) dV_{g_j}
\]
\[
= \int_{\partial A_{j,r, cr}} (r \partial_r u)(\partial_r \Delta g_j u) d\sigma_{g_j} - \int_{\partial A_{j,r, cr}} \partial_r (r \partial_r u) \Delta g_j u d\sigma_{g_j}
\]
\[
+ \int_{A_{j,r, cr}} \Delta g_j (r \partial_r u) \Delta g_j u dV_{g_j}.
\]

Direct calculations show that
\[
\int_{\partial A_{j,r, cr}} r \frac{\Delta g_j u^2}{2} d\sigma_{g_j} = \int_{A_{j,r, cr}} \text{div}_{g_j} \left( \frac{\Delta g_j u^2}{2} r \partial_r \right) dV_{g_j}
\]
\[
= \int_{A_{j,r, cr}} r \partial_r (\Delta g_j u) \Delta g_j u dV_{g_j} + \int_{A_{j,r, cr}} \text{div}_{g_j} (r \partial_r) \frac{\Delta g_j u^2}{2} dV_{g_j}.
\]

Combining (49) and (50) gives that
\[
\int_{\partial A_{j,r, cr}} (r \partial_r u)(\partial_r \Delta g_j u) + r \frac{\Delta g_j u^2}{2} - \partial_r (r \partial_r u) \Delta g_j u d\sigma_{g_j}
\]
\[
= \int_{A_{j,r, cr}} \text{div}_{g_j} \left( \frac{\Delta g_j u^2}{2} r \partial_r \right) dV_{g_j} - \int_{A_{j,r, cr}} \Delta g_j (r \partial_r u) \Delta g_j u dV_{g_j}.
\]
Set
\[ \int_{A_{j,r,cr}} \text{div}_{g_j} \left( \frac{\Delta_{g_j} u}{2} r \partial_r \right) \, dV_{g_j} - \int_{A_{j,r,cr}} \Delta_{g_j} (r \partial_r u) \Delta_{g_j} u \, dV_{g_j} = P(g_j, r), \]
then it follows that
\[ \int_{\partial A_{j,r,cr}} (r \partial_r u)(\partial_r \Delta_{g_j} u) + r \frac{|\Delta_{g_j} u|^2}{2} - \partial_r (r \partial_r u) \Delta_{g_j} u \, d\sigma_{g_j} = P(g_j, r). \]
Rewrite the term in the left hand side of (51) as in Subsection 3.2
\[ \square \equiv (r \partial_r u)(\partial_r \Delta_{g_j} u) + r \frac{|\Delta_{g_j} u|^2}{2} - \partial_r (r \partial_r u) \Delta_{g_j} u, \]
the identity (51) is equivalent to
\[ \int_{\partial A_{j,r,cr}} r(\square_1 + \square_2) \, d\sigma_{g_j} = P(g_j, r), \]
Here the forms of \( \square_1 \) and \( \square_2 \) are expressed in the same way as in (21) and (22) (replacing the metric \( g \) with \( g_j \)).

**Lemma 6.5.** Under the assumptions of (K_1) and (K_2), the error term satisfies
\[ |P(g_j, r)| \leq \int_{A_{j,r,cr}} \eta_j(r) \frac{\nabla_{g_j} u}{r} |\Delta_{g_j} u| + \eta_j(r) |\Delta_{g_j} u|^2 \, dV_{g_j}. \]

**Proof.** By (50),
\[ P(g_j, r) = \int_{A_{j,r,cr}} r \partial_r (\Delta_{g_j} u) \Delta_{g_j} u + \text{div}_{g_j} (r \partial_r) \frac{|\Delta_{g_j} u|^2}{2} - \Delta_{g_j} (r \partial_r u) \Delta_{g_j} u \, dV_{g_j}. \]
The idea is to show that
\[ \Delta_{g_j} (r \partial_r u) \Delta_{g_j} u - r \partial_r (\Delta_{g_j} u) \Delta_{g_j} u \]
is almost equal to \( \text{div}_{g_j} (r \partial_r) \frac{|\Delta_{g_j} u|^2}{2} \), which is almost equal to \( 2 |\Delta_{g_j} u|^2 \) by applying the following key lemma.

Set \( \partial B_r(x_{a,j}) \) be the geodesic sphere \( \partial B_r(x_{a,j}) \) with the metric \( g_r \equiv \frac{g_j}{r} \), where \( g_r \) is the metric of \( \partial B_r(x_{a,j}) \) induced from \((M_j, g_j)\). Let \( \tilde{m} \) be the mean curvature, \( \nabla \) be the gradient operator and \( \tilde{\triangle} \) be the associated Laplace operator of \( \partial B_r(x_{a,j}) \). So
\[ g_j = dr^2 + g_r = r^2 g_r \]
and
\[ \Delta_{g_j} u = \partial_r^2 u + \frac{\tilde{m}}{r} \partial_r u + \frac{\tilde{\triangle} u}{r^2}. \]
To complete the proof of Lemma 6.5, we need the following lemma, which will be proved in Appendix A.

**Lemma 6.6.** Let \( \tilde{\nabla}, \tilde{\triangle} \) and \( \tilde{m} \) be as above, then on \( B_{2\rho}(x_{a,j}) \setminus B_{\rho/2}(x_{a,j}) \), the followings hold:
1) \[
\triangle_{g_j} r = (A(r, \theta))^{-1} \frac{\partial A(r, \theta)}{\partial r} = \frac{3 + \eta_j(r)}{r},
\]
\[
\text{div}_{g_j}(r \partial_r) - 4 = \eta_j,
\]
where \(A(r, \theta) drd\theta\) is the volume form;

2) \[
r \partial_r \triangle_{g_j} r = \frac{3 + \eta_j(r)}{r};
\]

3) for any smooth function \(f\), on \(\tilde{\partial}B_{\rho}(x_{a,j})\),
\[
|\nabla_{S^3} f - \tilde{\nabla} f| = C \eta_j |\nabla_{S^3} f|,
\]
\[
|\Delta_{S^3} f - \tilde{\Delta} f| = C \eta_j (|\Delta_{S^3} f| + |\nabla_{S^3} f|),
\]
where the norm is taken with respect to the standard metric on \(S^3\) and it is equivalent to the norm defined by \(\frac{\eta_j}{\rho}\), and
\[
|\partial_r \Delta_{S^3} f - \partial_r \tilde{\Delta} f| = C \eta_j (|\Delta_{S^3} f| + |\nabla_{S^3} f|);
\]

4) on \(\tilde{\partial}B_{\rho}(x_{a,j})\),
\[
|m - 3| = \eta_j;
\]

5) on \(\tilde{\partial}B_{\rho}(x_{a,j})\),
\[
|\partial_r \tilde{m}| = \frac{\eta_j}{\rho}.
\]

By direct computations, we have
\[
\triangle_{g_j} (r \partial_r u) \triangle_{g_j} u = \left[(\triangle_{g_j} r) \partial_r u + 2 \partial^2_r u + r \Delta_{g_j} \partial_r u\right] \triangle_{g_j} u.
\]
By (54), we have
\[
r \triangle_{g_j} \partial_r u = r \left( \partial^2_r + \frac{\tilde{m}}{r} \partial_r + \frac{\tilde{\Delta}}{r^2}\right) \partial_r u
\]
\[
= r \left( \partial^3_r + \frac{\tilde{m}}{r} \partial^2_r + \frac{\tilde{\Delta}}{r^2} \partial_r\right) u,
\]
and
\[
r \partial_r \triangle_{g_j} u = r \partial_r \left( \partial^3_r + \frac{\tilde{m}}{r} \partial^2_r + \frac{1}{r^2} \partial_r \tilde{\Delta}\right) u + r \partial_r \left( \frac{\tilde{m}}{r}\right) \partial_r u - \frac{2}{r^2} \tilde{\Delta} u.
\]
Hence

\[ r \triangle g_j \partial_r u - r \partial_r \triangle g_j u \]
\[ = r \left( \frac{\Delta}{r^2} \partial_r - \frac{1}{r^2} \partial_r \tilde{\Delta} \right) u - r \partial_r \left( \frac{\tilde{m}}{r} \right) \partial_r u + \frac{2}{r^2} \tilde{\Delta} u. \]

So that from (55), we get

\[ \triangle g_j (r \partial_r u) \triangle g_j u \]
\[ = r \partial_r \triangle g_j u \Delta g_j u + \left[ (\triangle g_j r) \partial_r u + 2 \partial_r^2 u - r \partial_r \left( \frac{\tilde{m}}{r} \right) \partial_r u + \frac{2}{r^2} \tilde{\Delta} u \right] \Delta g_j u \]
\[ + r \left( \frac{\Delta}{r^2} \partial_r - \frac{1}{r^2} \partial_r \tilde{\Delta} \right) u \Delta g_j u. \]  

(56)

On \( A_{j;1/2r,2r} \) with the scaled metric \( \tilde{g}_j = \frac{g_j}{r^2} \) (we denote it by \( \tilde{A}_{j;1/2r,2r} \)), \( \tilde{\partial}_r = r \partial_r \) has length one. At any point \( q \in \partial B_r(x_{a,j}) \), we can choose \( \tilde{\partial}_i, i = 1, 2, 3 \), such that \( (\tilde{\partial}_r, \tilde{\partial}_1, \tilde{\partial}_2, \tilde{\partial}_3) \) is a normal frame at \( q \in \tilde{A}_{j;1/2r,2r} \). From Proposition 2.6, on \( A_{j;1/2r,2r} \) with the metric \( g_j \),

\[ r^2 |Rm(g_j)| \leq \eta_j(r), \]

so on \( \tilde{A}_{j;1/2r,2r} \)

\[ |Rm(\tilde{g}_j)| \leq \eta_j(r). \]

Therefore, by using the Ricci identity, we have

\[ \left| \left[ \frac{\Delta}{r^2} \partial_r - \frac{1}{r^2} \partial_r \tilde{\Delta} \right] u \right| \leq \eta_j(r) |\nabla u|. \]

Hence,

(57)

\[ \left| r \left[ \frac{\Delta}{r^2} \partial_r - \frac{1}{r^2} \partial_r \tilde{\Delta} \right] u \right| \leq \frac{\eta_j(r)}{r} |\nabla u|. \]

With the help of Lemma 6.6, from (57) and (56), we can obtain that

(58)

\[ |\triangle g_j (r \partial_r u) \triangle g_j u - r \partial_r \triangle g_j u \Delta g_j u - 2 |\triangle g_j u|^2| \]
\[ \leq \frac{\eta_j(r)}{r} |\nabla g_j u||\Delta g_j u|. \]

Indeed by Lemma 6.6,

\[ \left[ (\triangle g_j r) \partial_r u + 2 \partial_r^2 u - r \partial_r \left( \frac{\tilde{m}}{r} \right) \partial_r u + \frac{2}{r^2} \tilde{\Delta} u \right] = 2 \triangle g_j u + \frac{\eta_j(r)}{r} \partial_r u. \]

Again by using 1) of Lemma 6.6, (58) and (54), we have

\[ |P(g_j, r)| \leq | \int_{A_{j;1/2r,2r}} \frac{\eta_j(r)}{r} |\nabla g_j u||\Delta g_j u| + \eta_j(r)|\Delta g_j u|^2 dV_g |. \]

The proof of Lemma 6.5 is completed. \( \square \)
By Proposition 4.6 and Lemma 6.6, we can easily derive by using the same arguments as in Subsection 3.2 the following error estimate:

**Lemma 6.7.** On \( A_j \subset I^a_j \), for any \( \varepsilon > 0 \), when \( j \) and \( R > 0 \) are large enough and \( \delta > 0 \) is sufficiently small, we have

\[
|P(g_j, r)| = \varepsilon^2 \eta_j(r).
\]

**Proof.** By applying (i) of Proposition 4.6 (see also (R_1) in the beginning of this Section) which ensure that the energy of \( u_j \) on \( A_{j,cr} \) is less than \( \varepsilon^2 \), and the Hölder inequalities,

\[
\left| \int_{A_{j,cr}} \eta_j(r) |\nabla g_j u||\triangle g_j u| + \eta_j(r)|\triangle g_j u|^2 dV_{g_j} \right|
\]

is bounded by \( \varepsilon^2 \eta_j(r) \). See also the proof of Lemma 3.8. \( \square \)

Now we are going to analyse the relation between the tangential part energy and the radial part energy.

Using integration by part as in (26), we get

\[
\int_{A_{j,cr}} (\partial_r u)(\partial_r^3 u)dV_{g_j} = \int_{\partial(A_{j,cr})} (\partial_r u)(\partial_r^2 u) d\sigma_{g_j} - \int_{A_{j,cr}} (\partial_r^2 u)^2 + \left( \frac{3}{r} + \frac{\eta_j(r)}{r} \right) (\partial_r u)(\partial_r^2 u)dV_{g_j}.
\]

We have used \( \triangle g_j r = \frac{3}{r} + \frac{\eta_j(r)}{r} \) in the last formula, see 1) of Lemma 6.6. Recall that

\[
\square_2 = (\partial_r u)(\partial_r^3 u) + \frac{\tilde{m}}{r} - \frac{1}{r^2} (\partial_r u)(\partial_r^2 u) + \left( \frac{\partial_r \tilde{m}}{r} - \frac{2\tilde{m}}{r^2} \right) (\partial_r u)^2 + \frac{1}{r^2}(\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3}(\partial_r u)(\tilde{\Delta} u).
\]

We get

\[
\int_{A_{j,cr}} \square_2 dV_{g_j} = \int_{\partial(A_{j,cr})} (\partial_r u)(\partial_r^2 u) d\sigma_{g_j} - \int_{A_{j,cr}} (\partial_r^2 u)^2 + \left( \frac{3}{r} + \frac{\eta_j(r)}{r} \right) (\partial_r u)(\partial_r^2 u)dV_{g_j} + \int_{A_{j,cr}} \frac{\tilde{m}}{r} (\partial_r u)(\partial_r^2 u) + \left( \frac{\partial_r \tilde{m}}{r} - \frac{2\tilde{m}}{r^2} \right) (\partial_r u)^2 dV_{g_j} + \int_{A_{j,cr}} \frac{1}{r^2}(\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3}(\partial_r u)(\tilde{\Delta} u)dV_{g_j}.
\]
Next, by direct calculations using 4) and 5) of Lemma 6.6, we have

\[
\int_{A_{j,\overline{r},\tau}} \Box_2 dV_{g_j} \\
= \int_{\partial(A_{j,\overline{r},\tau})} (\partial_r u)(\partial^2_{\tau^2} u) d\sigma_{g_j} + \int_{A_{j,\overline{r},\tau}} \left( O\left( \frac{\eta_j}{r^2} \right) - \frac{6 + \eta_j}{r^2} \right) (\partial_r u)^2 - (\partial^2_{\tau^2} u)^2 dV_{g_j} \\
+ \int_{A_{j,\overline{r},\tau}} \frac{2 + \eta_j}{r} (\partial_r u)(\partial^2_{\tau^2} u) - \left( \frac{3}{r^2} + \frac{\eta_j}{r} \right) (\partial_r u)(\partial^2_{\tau^2} u) dV_{g_j} \\
+ \int_{A_{j,\overline{r},\tau}} \frac{1}{r^2} (\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3} (\partial_r u)(\tilde{\Delta} u) dV_{g_j}.
\]

Hence

\[
\int_{A_{j,\overline{r},\tau}} \Box_2 dV_{g_j} \\
= \int_{\partial(A_{j,\overline{r},\tau})} (\partial_r u)(\partial^2_{\tau^2} u) d\sigma_{g_j} + \int_{A_{j,\overline{r},\tau}} \left( O\left( \frac{\eta_j}{r^2} \right) - \frac{5}{r^2} \right) (\partial_r u)^2 - (\partial^2_{\tau^2} u)^2 dV_{g_j} \\
+ \int_{A_{j,\overline{r},\tau}} \left( - \frac{1}{2r} + \frac{\eta_j}{r} \right) \partial_r (\partial_r u)^2 - \frac{1}{r^2} (\partial_r u)^2 dV_{g_j} \\
+ \int_{A_{j,\overline{r},\tau}} \frac{1}{r^2} (\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3} (\partial_r u)(\tilde{\Delta} u) dV_{g_j}.
\]

By the use of the following formula

\[
(60) \quad \int_{\partial(A_{j,\overline{r},\tau})} \frac{1}{2r} (\partial_r u)^2 d\sigma_{g_j} = \int_{A_{j,\overline{r},\tau}} \frac{1}{2r} \partial_r (\partial_r u)^2 + \frac{1}{r^2} (\partial_r u)^2 + O\left( \frac{\eta_j}{r^2} \right) (\partial_r u)^2 dV_{g_j},
\]

we have

\[
\int_{A_{j,\overline{r},\tau}} \Box_2 dV_{g_j} \\
= \int_{\partial(A_{j,\overline{r},\tau})} (\partial_r u)(\partial^2_{\tau^2} u) d\sigma_{g_j} + \int_{A_{j,\overline{r},\tau}} \left( O\left( \frac{\eta_j}{r^2} \right) - \frac{5}{r^2} \right) (\partial_r u)^2 - (\partial^2_{\tau^2} u)^2 dV_{g_j} \\
- \int_{\partial(A_{j,\overline{r},\tau})} \frac{1}{2r} (\partial_r u)^2 d\sigma_{g_j} + \int_{A_{j,\overline{r},\tau}} \frac{\eta_j}{r} \partial_r (\partial_r u)^2 dV_{g_j} \\
+ \int_{A_{j,\overline{r},\tau}} \frac{1}{r^2} (\partial_r u)(\partial_r \tilde{\Delta} u) - \frac{\tilde{m}}{r^3} (\partial_r u)(\tilde{\Delta} u) dV_{g_j}.
\]

The formula (60) can be deduced as in (27), except that here we use \( \Delta_{g_j} r = \frac{3}{r} + \frac{\eta_j}{r} \).
Putting the boundary terms together, we have

\[
\int_{A_{j,r,cr}} \Box_2 dV_{g_j} = \int_{\partial(A_{j,r,cr})} \left( \nabla_r u (\nabla_r^2 u) - \frac{1}{2r} (\nabla_r u)^2 \right) d\sigma_{g_j}
\]

\[
+ \int_{A_{j,r,cr}} \left( O(\frac{\eta_j}{r^2}) - \frac{5}{r^2} \right) (\nabla_r u)^2 - (\nabla_r^2 u)^2 + \frac{\eta_j}{r} \partial_r (\nabla_r u)^2 dV_{g_j}
\]

\[
+ \int_{A_{j,r,cr}} \frac{1}{r^2} (\nabla_r u)(\nabla_r \triangle u) - \frac{\tilde{m}}{r^3} (\nabla_r u)(\tilde{\triangle} u) dV_{g_j}.
\]

Recall that

\[
\Box_1 = -\frac{1}{2} (\nabla_r^2 u)^2 + (\frac{\tilde{m}^2}{2r^2}) (\nabla_r u)^2 + \frac{1}{2r^4} (\tilde{\triangle} u)^2 + (\frac{\tilde{m}}{r^3}) (\nabla_r u)(\tilde{\triangle} u).
\]

Set

\[
\mathcal{R}(u_j, r) = \frac{3}{2} (\nabla_r^2 u_j)^2 + \left( \frac{10 - \tilde{m}^2}{2r^2} + O(\frac{\eta_j}{r^2}) \right) (\nabla_r u_j)^2
\]

and

\[
\mathcal{T}(u_j, r) = \frac{1}{2r^4} (\tilde{\triangle} u_j)^2 + \frac{1}{r^2} (\nabla_r u_j)(\nabla_r \tilde{\triangle} u_j).
\]

Note that by 4) of Lemma 6.6,

\[
\mathcal{R}(u_j, r) = \frac{3}{2} (\nabla_r^2 u_j)^2 + \left( \frac{1}{2r^2} + O(\frac{\eta_j}{r^2}) \right) (\nabla_r u_j)^2.
\]

Hence we have

\[
\int_{A_{j,r,cr}} \Box_2 + \Box_1 dV_{g_j}
\]

\[
= \int_{A_{j,r,cr}} \mathcal{T}(u_j, r) - \mathcal{R}(u_j, r) + \frac{\eta_j}{r} \partial_r (\nabla_r u_j)^2 dV_{g_j}
\]

\[
+ \int_{\partial(A_{j,r,cr})} (\nabla_r u_j)(\nabla_r^2 u_j) - \frac{1}{2r} (\nabla_r u_j)^2 d\sigma_{g_j}.
\]

From (52) we obtain that

\[
e \sqrt{A_{j,r,cr}} dV_{g_j} - \int_{A_{j,r,cr}} (\Box_1 + \Box_2) dV_{g_j} = \int_{A_{j,r,cr}} P(g_j, s) dV_{g_j} = \int_{\partial(A_{j,r,cr})} (\nabla_r u_j)(\nabla_r^2 u_j) - \frac{1}{2r} (\nabla_r u_j)^2 d\sigma_{g_j}.
\]

By similar argument as in the proof of Lemma 6.7, we get

\[
\int_{A_{j,r,cr}} (\nabla_r u_j)(\nabla_r^2 u_j) dV_{g_j} \leq C \varepsilon^2 r,
\]

it follows that

\[
\int_{A_{j,r,cr}} \frac{\eta_j}{r} (\nabla_r u_j)^2 dV_{g_j} \leq \varepsilon^2 \eta_j(r).
\]
Therefore, by combining (63), (62), (59) (i.e., \( P(g_j, r) = \varepsilon^2 \eta_j(r) \)) and the last inequality together, we have

\[
\begin{align*}
&c \int_{A_{j,cr}} \mathcal{R}(u_j, r) dV_{g_j} - \int_{A_{j,cr}} \mathcal{R}(u_j, r) dV_{g_j} \\
&= c \int_{A_{j,cr}} \mathcal{T}(u_j, r) dV_{g_j} - \int_{A_{j,cr}} \mathcal{T}(u_j, r) dV_{g_j} \\
&\quad + c \int_{\partial(A_{j,cr})} ((\partial_{r}u_j)(\partial^2_{r}u_j)) - \frac{1}{2r} (\partial_{r}u_j)^2 d\sigma_{g_j} \\
&\quad - \int_{\partial(A_{j,cr})} ((\partial_{r}u_j)(\partial^2_{r}u_j)) - \frac{1}{2r} (\partial_{r}u_j)^2 d\sigma_{g_j} + \varepsilon^2 \eta_j(r).
\end{align*}
\]

(64)

Set \( r_2(s) = e^{t_0 + s} \) and \( r_1(s) = e^{t_0 - s} \), then similarly to (32), we get the following formula:

\[
\begin{align*}
&c \int_{A_{j,cr}} \mathcal{R}(u_j, r) dV_{g_j} - \int_{A_{j,cr}} \mathcal{R}(u_j, r) dV_{g_j} \\
&= c \int_{A_{j,cr}} \mathcal{T}(u_j, r) dV_{g_j} - \int_{A_{j,cr}} \mathcal{T}(u_j, r) dV_{g_j} \\
&\quad + c \int_{\partial(A_{j,cr})} ((\partial_{r}u_j)(\partial^2_{r}u_j)) - \frac{1}{2r} (\partial_{r}u_j)^2 d\sigma_{g_j} \\
&\quad - \int_{\partial(A_{j,cr})} ((\partial_{r}u_j)(\partial^2_{r}u_j)) - \frac{1}{2r} (\partial_{r}u_j)^2 d\sigma_{g_j} + \int_{r_1}^{r_2} \varepsilon^2 \frac{\eta_j(r)}{r} dr.
\end{align*}
\]

(65)

Since

\[ \eta_j(t) = O(r^{-\varepsilon_5} t^{\varepsilon_5} + (r_j R_0)^{\varepsilon_5} t^{-\varepsilon_5}), \]

the last term in the right hand side of (65) can be controlled by

\[ \varepsilon^2 \max(\eta_j(r_1), \eta_j(r_2)) \leq C \varepsilon^2 (e^{\varepsilon_5 (t_0 - \log r_\infty)} + e^{\varepsilon_5 (\log(r_j R_0) - t_0)}) e^{\varepsilon_5 s}. \]

6.3. The energy identity and no neck property. Now we are in a position to prove the following:

**Theorem 6.8.** Let \( I_j^\circ \) be the annulus region in \( M_j \) corresponding to \( I_j^\circ \) in Proposition 4.6, then (for the extrinsic case)

\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{j \to \infty} \int_{I_j^\circ} \left| \nabla_{g_j}^2 u_j \right|^2 + \left| \nabla_{g_j} u_j \right|^4 dV_{g_j} = 0
\]

and

\[
\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{j \to \infty} \text{osc}_{I_j^\circ} u_j = 0.
\]
Proof. Recall $\bar{I}_j^\alpha = A_{b(j,\alpha),a(j,\alpha)}(x_{a,j})$, where $e^{b_\alpha} \equiv b(j,\alpha)$, $a(j,\alpha) \equiv e^{a_\alpha+1}$. If we set $c = e^L$, then from the decay estimates for the tangential part energy

$$\int_{A_{j,r,\tau}} \frac{1}{r^4} |(\triangle u_j + |\nabla S u_j|^4) + \frac{1}{r^2} |\partial_r \nabla S u_j|^2 + \frac{1}{r^2} |\partial_r \triangle S u_j|^2| dx$$

$$\leq C \varepsilon^2 \left[ \left( \frac{r}{a(j,\alpha)} \right)^\vartheta + \left( \frac{b(j,\alpha)}{r} \right)^\vartheta \right],$$

in Theorem 6.2, we get by the Cauchy-Schwarz inequality and 3) of Lemma 6.6 that for $r \in [b(j,\alpha), a(j,\alpha)] \subset [r_j R, \delta]$,}

$$\int_{A_{j,r,\tau}} |\mathcal{T}(u_j, r)| dV_{g_j} \leq \tau \int_{A_{j,r,\tau}} \frac{1}{r^2} (\partial_r u_j)^2 dV_{g_j}$$

$$+ C(\tau) \int_{A_{j,r,\tau}} \frac{1}{r^4} (|\tilde{\Delta} u_j|^2 + |\tilde{\nabla} u_j|^4) + \frac{1}{r^2} |\partial_r \tilde{\nabla} u_j|^2 + \frac{1}{r^2} |\partial_r \tilde{\Delta} u_j|^2 dV_{g_j}$$

$$\leq C(\tau) \varepsilon^2 \left[ \left( \frac{r}{a(j,\alpha)} \right)^\vartheta + \left( \frac{b(j,\alpha)}{r} \right)^\vartheta \right] + \tau \int_{A_{j,r,\tau}} \frac{1}{r^2} (\partial_r u_j)^2 dV_{g_j},$$

where $\tau > 0$ is a small number, and $C(\tau)$ is a bounded number depending on $\tau$. Due to Remark 6.4, we can take $\vartheta = 2\varepsilon_5 > 0$ (w.l.o.g. we assume that $2\varepsilon_5 \leq 1$). It is easy to see that the term $\tau \frac{1}{r^2} (\partial_r u_j)^2$ in (66) can be absorbed into $\mathcal{R}(u_j, r)$ if we choose $\tau > 0$ to be small enough.

Recall that $A_l = A_{j,e^{-1}L,e^{-(l-1)}L}$, $b_j^\alpha = -l_{j,b}^\alpha L$, $a_j^{\alpha+1} = -l_{j,a}^\alpha L$ and

$$\eta_j = O \left( \left( \frac{r}{\delta_1} \right)^{\xi_5} + \left( \frac{r_j R}{r} \right)^{\xi_5} \right).$$

Set

$$\mathcal{I}_j^\alpha = (\bar{I}_j^\alpha \cup A_{j,a-1}) \setminus A_{j,b+1}.$$  

Due to (i) of Proposition 4.6 (see also $A_1$ in the beginning of this section), we can assume that

$$\int_{A_{j,a-1} \cup A_{j,b+1}} |\nabla_{g_j}^2 u_j|^2 + |\nabla_{g_j} u_j|^4 dV_{g_j} \leq \varepsilon^2 < \varepsilon_0.$$  

Recall that

$$\mathcal{R}(u_j, r) = \frac{3}{2} (\partial_r^2 u_j)^2 + \left( \frac{1}{2r^2} + O(\frac{\eta_j}{r^2}) \right) (\partial_r u_j)^2,$$
and notice that the boundary terms can be cancelled out by each other, from (64) we get
\[(c - 1) \int_{I_j^0} \frac{3}{2} (\partial_r^2 u_j)^2 + \left( \frac{1}{2r^2} + O \left( \frac{\eta_j}{r^2} \right) \right) (\partial_r u_j)^2 dV_j \leq c \int_{I_j^0,\delta(B)} \mathcal{R}(u_j, r) dV_j - \int_{I_j^0} \mathcal{R}(u_j, r) dV_j + C\varepsilon^2 \]
\[\leq (c + 1) \sum_{l_a}^{l_a+1} \int_{A_l} |T(u_j, r)| dV_j + C\varepsilon^2 \sum_{l_a}^{l_a+1} \eta_j (e^{-L}) \]
\[+ c \int_{\partial(I_j^0) \cup \partial(I_j^0)^\star} |(\partial_r u_j)(\partial_r^2 u_j)| + \frac{1}{2r} (\partial_r u_j)^2 d\sigma_j + C\varepsilon^2. \]

And then by (66), we have
\[\int_{I_j^0} \frac{3}{2} (\partial_r^2 u_j)^2 + \left( \frac{1}{2r^2} \right) (\partial_r u_j)^2 dV_j \leq C(c) C(\tau) \sum_{l_a}^{l_a+1} \varepsilon^2 \left[ \left( \frac{e^{-L}}{a(j, \alpha)} \right)^\delta + \left( \frac{b(j, \alpha)}{e^{-L}} \right)^\delta \right] + \int_{I_j^0} C(c) \left( \frac{\tau}{r^2} \right) (\partial_r u_j)^2 dV_j \]
\[+ C(c) \int_{\partial(I_j^0) \cup \partial(I_j^0)^\star} |(\partial_r u_j)(\partial_r^2 u_j)| + \frac{1}{2r} (\partial_r u_j)^2 d\sigma_j + C(c)\varepsilon^2 + C(c)\varepsilon^2. \]

In the above estimate, \(C(c) > 0\) is some constant dependent on \(c\). Next, by taking \(\tau\) small properly, we have
\[\int_{I_j^0} \frac{3}{2} (\partial_r^2 u_j)^2 + \frac{1}{2r^2} (\partial_r u_j)^2 dV_j \leq C \int_{\partial(I_j^0) \cup \partial(I_j^0)^\star} |(\partial_r u_j)(\partial_r^2 u_j)| + \frac{1}{2r} (\partial_r u_j)^2 d\sigma_j + C\varepsilon^2. \]

Here we have used the notation \(\hat{\partial}(B_R \setminus B_r) = \partial B_R \cup \partial B_r\).

Next by using the \(\varepsilon\)-regularity Theorem 2.9 and Sobolev embedding theorem, we get
\[\int_{\partial(I_j^0) \cup \partial(I_j^0)^\star} |(\partial_r u_j)(\partial_r^2 u_j)| + \frac{1}{2r} (\partial_r u_j)^2 d\sigma_j \leq C\varepsilon^2. \]

Therefore, we have that for any \(\varepsilon > 0\), there are constants \(R_0 > 0\), \(\delta_0 > 0\) and \(j_0 > 0\) depending only on \(\varepsilon\), such that, when \(j > j_0\), \(R > R_0\) and \(\delta < \delta_0\), if we take some \(c \geq 2\) in the above arguments, then
\[(67) \quad \int_{I_j^0} (\partial_r^2 u_j)^2 + \frac{1}{r^2} (\partial_r u_j)^2 dV_j \leq C\varepsilon^2. \]

Combining the last inequality with (66), the energy identity can be proved.
Recall that \( r_2(s) = e^{t_0 + s}, \) \( r_1(s) = e^{t_0 - s}, \) and

\[
\mathcal{R}(u_j, r) = \frac{3}{2}(\partial_r^2 u_j)^2 + \left( \frac{1 + \eta_j(r)}{2r^2} \right)(\partial_r u_j)^2.
\]

To prove the no neck property, we define

\[
F(s) = \int_{A_{j,r_1,r_2}} \mathcal{R}(u_j, r) dV_g,
\]

and

\[
G(s) = \int_{A_{j,e^{t_0},e^{t_0}}} \mathcal{R}(u_j, r) dV_g.
\]

It is easy to see that \( F \) and \( G \) are non-decreasing function of \( s \), and they are uniformly bounded by \( C \varepsilon^2 \) (see (67)). We remark that \( b_j^1 < t_0 < a_j^{a+1} < 0, s \in [0, \min(t_0 - b_j^1, a_j^{a+1} - t_0)] \).

First, we rewrite (65) to be

\[
\int_{\partial(A_{j,r_1,r_2})} ((\partial_r u_j)(\partial_r^2 u_j)) - \frac{1}{2r}(\partial_r u_j)^2 d\sigma_j = \int_{A_{j,r_1,r_2}} \mathcal{R}(u_j, r) dV_g
\]

\[
+ c \int_{\partial(A_{j,r_1,r_2})} ((\partial_r u_j)(\partial_r^2 u_j)) - \frac{1}{2r}(\partial_r u_j)^2 d\sigma_j
\]

\[
- c \int_{A_{j,r_1,r_2}} \mathcal{R}(u_j, r) dV_g \int_{r_1}^{r_2} \varepsilon^2 \frac{\eta_j(r)}{r} dr.
\]

Then by noticing that \( \frac{1}{\sqrt{3}} < \frac{2}{3} \),

\[
\partial_s F(s) = r_1(s) \int_{\partial B_{r_1}(x_{a,j})} \mathcal{R}(u_j, r) dV_g + r_2(s) \int_{\partial B_{r_2}(x_{a,j})} \mathcal{R}(u_j, r) dV_g,
\]

and

\[
\left| \int_{\partial(B_{r_2} \setminus B_{r_1})} (\partial_r u)(\partial_r^2 u) d\sigma \right|
\]

\[
\leq \frac{1}{\sqrt{3}} r_2 \int_{\partial B_{r_2}} \frac{3}{2}(\partial_r^2 u)^2 + \left( \frac{1}{2r_2^2} \right)(\partial_r u)^2 d\sigma
\]

\[
+ \frac{1}{\sqrt{3}} r_1 \int_{\partial B_{r_1}} \frac{3}{2}(\partial_r^2 u)^2 + \left( \frac{1}{2r_1^2} \right)(\partial_r u)^2 d\sigma,
\]

we get from (68) and (66) that (set \( c = e^L \))

\[
- \frac{2}{3} \partial_s F(s) - F(s) \leq e^L \left( \frac{2}{3} \partial_s G(s) - G(s) \right)
\]

\[
+ C \varepsilon^2 (e^{2(t_0 - \log r_\infty)} + e^{2\log(r_j R_0 - t_0)}) e^{\varepsilon s} + C \varepsilon^2 (e^{\theta(t_0 - a_j^{a+1})} + e^{\theta(b_j^1 - t_0)}) e^{\theta s}.
\]
We remark here that the term involving $\tau$ in (66) and the error terms involving $\eta_j(r)$ in $G(s)$ and $F(s)$ do not have essential effects to the validity of the last inequality, as long as we set $\tau$ properly small.

Recall that $0 < \vartheta = 2\varepsilon_5 \leq 1$. Multiplying $e^{-\frac{3}{2}s}$ to both sides of the above inequality, and noticing that $[b_j', a_j^{a+1}] \subset [\log(r_jR), \log \delta]$, we have

$$e^t \partial_s(e^{-\frac{3}{2}s}G)(s) + \partial_s(e^{-\frac{3}{2}s}F)(s) \geq -3e^{-\frac{3}{2}s}F(s) - C\varepsilon^2(\varepsilon_5(t_0-a_j^{a+1}) + \varepsilon_5(b_j'-t_0))e^{(\varepsilon_5-\frac{3}{2})s}.$$  

Integrating the above inequality (69) from $s = L$ to $s = 2L$, we have

$$e^L(e^{-3L}G(2L) - e^{-\frac{3}{2}L}G(L)) + (e^{-3L}F(2L) - e^{-\frac{3}{2}L}F(L)) \geq -3 \int_L^{2L} e^{-\frac{3}{2}s}F(s)ds - C\varepsilon^2(\varepsilon_5(t_0-a_j^{a+1}) + \varepsilon_5(b_j'-t_0)).$$

Notice that on $[L, 2L]$, $F(s) \leq F(2L)$, we have from (70) that

$$e^L G(L) + F(L) \leq e^{-\frac{3}{2}L}(F(2L) + e^L G(2L)) + 3e^L \int_L^{2L} e^{-\frac{3}{2}s}F(s)ds + C\varepsilon^2(\varepsilon_5(t_0-a_j^{a+1}) + \varepsilon_5(b_j'-t_0)) \leq e^{-\frac{3}{2}L}(F(2L) + e^L G(2L)) + 3LF(2L) + C\varepsilon^2(\varepsilon_5(t_0-a_j^{a+1}) + \varepsilon_5(b_j'-t_0)).$$

So

$$e^L G(L) \leq e^{-\frac{3}{2}L}G(2L) + (3L + 1)F(2L) + C\varepsilon^2(\varepsilon_5(t_0-a_j^{a+1}) + \varepsilon_5(b_j'-t_0)).$$

Thanks to the simple fact that $G(2L), F(2L) \leq G(3L)$, if we choose $L > 0$ such that

$$\max\{(3L + 1)e^{-\frac{3}{2}L}, e^{-\frac{3}{2}L}\} \leq \frac{1}{2}e^{-\frac{L}{2}},$$

then

$$G(L) \leq e^{-\frac{L}{2}}G(3L) + Ce^{(\varepsilon_5-1)L}\varepsilon^2(\varepsilon_5(t_0-a_j^{a+1}) + \varepsilon_5(b_j'-t_0)).$$

By applying an iteration argument, we get

$$G(L) \leq Ce^2(\varepsilon_5(t_0-a_j^{a+1}) + e_5(b_j'-t_0)) + Ce^2(\varepsilon_5(t_0-a_j^{a+1}) + e_5(b_j'-t_0),$$

since $G(s)$ are uniformly bounded by $Ce^2$.

Finally, similarly as in Subsection 3.2, the above radial energy decay together with the decay of the tangential part energy in (66) is enough to ensure the validity of the no neck property. The proof of Theorem 6.8 is now completed.

7. INTRINSIC BIHARMONIC MAPS

In this section, we shall show that the energy identity and no neck property as stated in Theorem 1.1 and Theorem 1.2 hold also for intrinsic biharmonic maps.

Recall that the equations of extrinsic and intrinsic biharmonic maps differ only by the following term (see Section 7 of [34]):

$$I(u) \equiv P(u)(B(u)(\nabla u, \nabla u))\nabla u B(u)(\nabla u, \nabla u) + 2B(u)(\nabla u, \nabla u)B(u)(\nabla u, \nabla P(u)).$$
That is to say, the intrinsic biharmonic map equations take the form
\begin{equation}
\Delta^2_g u = \Delta_g u(B(u)(\nabla_g u, \nabla_g u)) + 2\nabla_g \cdot \langle \Delta_g u, \nabla_g (P(u)) \rangle - \langle \Delta_g (P(u)), \Delta_g u \rangle + I(u).
\end{equation}

Here $B$ is the second fundamental form of $N \subset \mathbb{R}^K$ and $P(u)$ is the projection to the tangent space $T_{u}N$.

In view of the equation (71), it is easy to see that the three circle type method developed for deriving the decay estimate of the tangential part energy for extrinsic biharmonic maps in both the non-degenerating case (see Subsection 3.1) and the degenerating case (see Subsection 6.1) can be applied to intrinsic biharmonic maps by slightly modifying the arguments. Therefore, to handle the case of intrinsic biharmonic maps, it is sufficient to develop the necessary Pohožaev type arguments.

7.1. Non-degenerating case. We deal with the non-degenerating case. As in Subsection 3.2, we multiply the term $I(u)$ by $x^k e_k u$, notice that $x^k e_k u = r \partial_r u$ is a tangent vector, then
\begin{align*}
\partial_r u \cdot I(u) &= B(u)(\nabla u, \nabla u) \nabla_r B(u)(\nabla u, \nabla u) + 2B(u)(\nabla u, \nabla u) B(u)(\nabla u, \nabla (r \partial_r u)) \\
&= r \partial_r \left( \frac{|B(u)(\nabla u, \nabla u)|^2}{2} \right) + 2|B(u)(\nabla u, \nabla u)|^2 \\
&= \frac{1}{r^3} \partial_r \left[ \frac{r^4}{2} |B(u)(\nabla u, \nabla u)|^2 \right].
\end{align*}

By applying similar argument as is done for (26), we get
\begin{align*}
\int_{B(r)} \frac{1}{r^3} \frac{\partial_r}{\partial r} \left( \frac{r^4}{2} f \right) dV_g \\
&= \int_{\partial B(r)} \frac{\partial_r (\sqrt{g})}{r^3 \sqrt{g}} \left( \frac{r^4}{2} f \right) dV_g + \int_{\partial B(r)} \frac{1}{r^3} \left( \frac{r^4}{2} f \right) d\sigma_g.
\end{align*}

Notice that
\[ \frac{\partial_r \sqrt{g}}{\sqrt{g}} = \Delta_g r - \frac{3}{r} = O(r), \]
we have
\[ \int_{B(r)} \frac{1}{r^3} \partial_r \left[ \frac{r^4}{2} |B(u)(\nabla u, \nabla u)|^2 \right] dV_g = \int_{\partial B(r)} \frac{r}{2} |B(u)(\nabla u, \nabla u)|^2 d\sigma_g + O(r^2), \]

since
\[ \int_{B(r)} |B(u)(\nabla u, \nabla u)|^2 dV_g \leq \int_{B(r)} C|\nabla u|^4 dV_g \leq C. \]

It follows that
\[ \int_{B(r)} x^k e_k u \cdot I(u) dV_g = \int_{\partial B(r)} \frac{r}{2} |B(u)(\nabla u, \nabla u)|^2 d\sigma_g + O(r^2). \]
Thus, if we choose \( \varepsilon > \frac{1}{r^2} \), the energy identity and no neck property. On the other hand, by the decay estimate for the tangential part energy (using three circle type method), we know that the term \( \frac{1}{r^2} B(u)(\nabla u, \nabla u) \) has no essential effect for the proof of the energy identity and no neck property. On the other hand, by \( \varepsilon \)-regularity Theorem 2.9 and Sobolev embedding theorem, we have

\[
|B(u)(\partial_r u, \partial_r u)|^2 \leq C \varepsilon^2 \frac{1}{r^2} (\partial_r u)^2.
\]

Thus, if we choose \( \varepsilon > 0 \) small enough, this term can be absorbed into the left hand of (72).

The rest of the proof can be completed by applying the same arguments as in the case of extrinsic biharmonic maps.

7.2. Degenerating case. We denote the term \( I(u_j) \) by \( I_j(u) \). Here, for simplicity, we denote \( \nabla_{g_j} \) by \( \nabla_j \), and write \( u \) for \( u_j \). As in Subsection 6.2, we multiply the term \( I_j(u) \) by \( r \partial_r u \),
then we obtain that on $A_{j;r,cr}$,

$$
\begin{align*}
& r \partial_r u \cdot I_j(u) \\
& = B(u)(\nabla_j u, \nabla_j u) \nabla_{r \partial_r u} B(u)(\nabla_j u, \nabla_j u) + 2 B(u)(\nabla_j u, \nabla_j u) B(u)(\nabla_j u, \nabla_j (r \partial_r u)) \\
& = r \partial_r \left( \frac{|B(u)(\nabla_j u, \nabla_j u)|^2}{2} \right) + 2 |B(u)(\nabla_j u, \nabla_j u)|^2 \\
& = \frac{1}{r^3} \partial_r \left[ \frac{r^4}{2} |B(u)(\nabla_j u, \nabla_j u)|^2 \right].
\end{align*}
$$

Recall that

$$
\frac{\partial_r \sqrt{g_j}}{\sqrt{g_j}} = \Delta g_j r - \frac{3}{r} = \frac{\eta_j(r)}{r},
$$

where $\sqrt{g_j}$ is the square root of determinant of the metric $g_j$ in the $(x)$ coordinates in Theorem 5.3. Then again by using integration by parts as before, we have

$$
\int_{A_{j;r,cr}} \frac{1}{r^3} \partial_r \left[ \frac{r^4}{2} |B(u)(\nabla_j u, \nabla_j u)|^2 \right] dV_{g_j} = \int_{\partial A_{j;r,cr}} r \frac{r}{2} |B(u)(\nabla_j u, \nabla_j u)|^2 d\sigma_{g_j} + \varepsilon^2 \eta_j(r).
$$

In the above, we have used (i) of Proposition 4.6 which ensures that the energy of $u_j$ on $A_{j;r,cr}$ is less than $\varepsilon^2$.

Then by applying similar arguments as in the non-degenerating case, we have

$$
\int_{\partial A_{j;r,cr}} r (\Box_1 + \Box_2) - \frac{r}{2} |B(u)(\nabla_j u, \nabla_j u)|^2 d\sigma_{g_j} = P_1(g_j, r),
$$

where $P_1(g_j, r) = P(g_j, r) + \varepsilon^2 \eta_j(r)$. The three terms $\Box_1, \Box_2$ and $P(g_j, r)$ (recall that $P(g_j, r) \leq \varepsilon^2 \eta_j(r)$) are defined as before. So, we have

$$
\begin{align*}
c \int_{A_{j;r,cr}^{2r}} (\Box_1 + \Box_2) - \frac{1}{2} |B(u)(\nabla_j u, \nabla_j u)|^2 dV_{g_j} \\
- \int_{A_{j;r,cr}} (\Box_1 + \Box_2) - \frac{1}{2} |B(u)(\nabla_j u, \nabla_j u)|^2 dV_{g_j} \\
= \int_r^{cr} \frac{1}{s} P_1(g_j, s) ds.
\end{align*}
$$

Finally, by applying the same arguments as in Subsection 6.2 for extrinsic biharmonic maps, it is easy to see that the additional term $|B(u)(\nabla_j u, \nabla_j u)|^2$ can be easily handled and has no essential influence to the rest of the proof.

**Remark 7.1.** In fact all results derived in this paper which are valid for extrinsic biharmonic maps are also true for intrinsic biharmonic maps.
APPENDIX A. PROOF OF LEMMA 3.7 AND LEMMA 6.6

We prove Lemma 6.6 in detail, and outline the proof for Lemma 3.7. Recall that
\[(73)\] 
\[\|g_{j,kl}(x) - \delta_{kl}\|_{C^0} < \eta_j(|x|),\]
\[(74)\] 
\[\|g_{j,kl}(\rho \bar{y}) - \delta_{kl}\|_{C^4(B_2 \setminus B_1)} < \eta_j(\rho),\]
where \(\bar{y} = \rho^{-1}y, (x)\) and \((y)\) are the two coordinates given in Theorem 5.3.

\textbf{Proof.} Firstly, we prove 1) of Lemma 3.7. Since \(e_k\) is parallel invariant along the radial geodesics, if we write \(e_k = f^{kl} \partial_{x^l}\), then \(f^{kl}, l = 1, \cdots, 4\), satisfy the following equations:
\[\frac{df^{kl}}{dr} + \Gamma^l_{mn} \frac{dx^m}{dr} f^{kn} = 0, f^{kl}(0) = \delta_{kl}.\]
In normal coordinates on \(B_r\),
\[|g_{kl} - \delta_{kl}| < Cr^2,\]
so \(|\Gamma^l_{mn}| < Cr\). Therefore
\[e_k = \partial_{x^k} + h^{kl} \partial_{x^l}, h^{kl} = O(r^2).\]
Hence 1) of Lemma 3.7 follows. And 2) of Lemma 3.7 follows from direct calculations in local coordinates.

To prove 3) of Lemma 3.7, recall that \(\partial B_r\) is the sphere \(\partial B_r\) with metric \(g_r \equiv \frac{g}{r^2}\), where \(g_r\) is the metric of \(\partial B_r\) induced from \((M, g)\). By the fact that the standard metric \(g_{S^3}\) on \(S^3\) and the metric \(g_r\) satisfy
\[\|g_{S^3,kl} - \bar{g}_{r,kl}\|_{C^4} \leq Cr^2,\]
we get by direct calculations the estimates in 3).

Next by standard theory in Riemannian geometry (see for example [8]), for a fix metric \(g\), in the geodesic normal coordinates \(\bar{m} = r \Delta_r = 3 + O(r^2)\), so 4), 5) and 6) of Lemma 3.7 follow immediately.

Now we prove 1), 2), 4), 5) of Lemma 6.6. As in the proof of Lemma 5.6, at a point \(\hat{\theta} \in S_{r_0, r_0} \equiv \partial B_r(x_{a,j}, r_0),\) we consider nontrivial Jacobi field \(J\) along the geodesic \(\sigma_{r_0}(t) = \Phi_{r_0}(\theta, t),\) with initial conditions \(J(r_0) \in T_{\hat{\theta}} S_{r_0, r_0}, \frac{\partial J(r_0)}{\partial \theta} = A_{r_0}(J(r_0))\). Let \(A(t, \theta)\) be the matrix solution of the differential equation in \(T_{\hat{\theta}} S_{r_0, r_0},\)
\[\frac{\partial^2}{\partial t^2} A(t, \theta) + R(t, \theta) A(t, \theta) = 0,\]
satisfying \(A(r_0, \theta) = r_0 A_{r_0}(u_1, u_2, u_3), \frac{\partial A(r_0, \theta)}{\partial \theta} = A_{r_0}(A(r_0, \theta)).\) Here \(\{u_1, u_2, u_3\}\) form an orthonormal basis for \(T_{\hat{\theta}} S_{r_0, r_0},\) and \(R(t, \theta)\) is the composition of the curvature operator at \(\sigma_{r_0}(t)\) with the parallel translation along the geodesic.

As before due to the curvature estimate in Proposition 2.6, we have
\[|A(t, \theta) - t Id| = O(t \eta_j(t)), \quad |\frac{\partial}{\partial t} A(t, \theta) - Id| = \eta_j(t), \quad \text{and} \quad \left|\frac{\partial^2}{\partial t^2} A(t, \theta)\right| = \frac{\eta_j(t)}{t}.\]
It is easy to see that
\[(J_1(r), J_2(r), J_3(r)) = A(t, \theta)(u_1(r), u_2(r), u_3(r)),\]
where $u_l(r), l = 1, 2, 3$ are the parallel invariant vector fields along the geodesics. So

$$\det A(t, \theta) = \frac{J_1(r) \wedge J_2(r) \wedge J_3(r)}{u_1(r) \wedge u_2(r) \wedge u_3(r)}.$$

By properties of Jacobi fields, we know that along $\sigma_{r, R_0}(t), A(t, \theta) = v(\theta) \det A(t, \theta)$ for some constant $v(\theta)$. Here $A(t, \theta)$ is the square root of the determinant of the metric in the polar coordinates $(t, \theta)$. Or equivalently $d\text{vol}_{g_j} = A(t, \theta)drd\theta$. Therefore

$$\triangle_{g_j} r = m(r)$$

$$= \sum_{l=1}^{3} (\nabla_J l \partial_r, J_l(r)) = \sum_{l=1}^{3} (\nabla_{\partial_r} J_l(r), J_l(r)) = \text{tr}(\partial_r AA^{-1})$$

$$= \partial_r (\log(\det A(r, \theta))) = (A(r, \theta))^{-1} \partial A(r, \theta)$$

$$= \frac{3 + \eta_j(r)}{r}$$

and

$$r \partial_r \triangle_{g_j} r = -\frac{3 + \eta_j(r)}{r},$$

where $m(r)$ is the mean curvature of the geodesic sphere $\partial B(x_{a,j}, r)$ with the induce metric from $(M_j, g_j)$. See [59] and Lemma 2.3 in [63] for similar computations. By now we have proved 1), 2) of Lemma 6.6. Next by $\tilde{m} = rm(r) = r\triangle_{g_j} r$, we obtain the desired estimates in 4) and 5) of Lemma 6.6.

Finally, we prove 3) of Lemma 6.6. Notice that (73) and (74) indicate that the metric $g_j$ is close to the flat metric, and that $\triangle_{g_j} u, \partial_r u$ and $\tilde{\Delta} u$ are independent on coordinate functions. It follows from direct computations that

$$|\triangle_{g_j} u - \triangle_{g_{\text{euc}}} u| \leq \eta_j(r)|\triangle_{g_{\text{euc}}} u| + \frac{\eta_j(r)}{r} |\nabla_{g_{\text{euc}}} u|_{g_{\text{euc}}},$$

where $g_{\text{euc}}$ is the standard flat metric. So from 1), 2), 4), 5) of Lemma 6.6 and the following two formulas

$$\triangle_{g_{\text{euc}}} u = \partial^2_r u + \frac{3}{r} \partial_r u + \frac{\triangle_{g_3} u}{r^2}$$

and

$$\triangle_{g_j} u = \partial^2_r u + \frac{\tilde{m}}{r} \partial_r u + \frac{\tilde{\Delta} u}{r^2},$$

we can get the desired estimates. □
APPENDIX B. PROOF OF THEOREM 5.2

Firstly, we recall the classical Schauder’s interior estimates for linear elliptic equations ([16], see also [35]). Let \( B \) be an open ball in \( \mathbb{R}^n \). Let \( d_x = d(x, \partial B) \) and \( d_{xy} = \min(d_x, d_y) \). For functions \( f \) in \( C^p(B) \) define

\[
||f||_{C^{p,\alpha}_p} = \sum_{j=0}^p \sup_B d_x^{-\sigma} |D^j f| + \sup_{x \neq y} d_{xy}^{p+\alpha-\sigma} |D^p f(x) - D^p f(y)| \frac{1}{|x - y|^\alpha}.
\]

Theorem B.1. Set

\[
L(u) = \sum_{k,l} a_{k,l}(x) \partial_k \partial_l u + \sum_k a_k(x) \partial_k u + a(x) u,
\]

where \( a_{k,l} \in C^{0,\alpha}_0(A_{r_1, r_2}), a_k \in C^{p,\alpha}_{-1}(A_{r_1, r_2}), a \in C^{p,\alpha}_{-2}(A_{r_1, r_2}) \) with norms less than \( K_1 > 0 \), and

\[
\sum_{k,l} a_{k,l}(x) \xi_k \xi_l \geq K_2 \sum_k \xi_k^2 \quad (K_2 > 0).
\]

Let \( u \) be a solution of \( L(u) = f \), then

\[
||u||_{C^{p+2,\alpha}_0} \leq K(||u||_{C^{0,\alpha}_0} + ||f||_{C^{p,\alpha}_2}),
\]

where \( K > 0 \) depends only on \( K_1, K_2, p, i, \alpha, \) and \( n \).

We remark here that \( K \) in the above does not depend on the radius of the ball.

Now we list some properties of the \( C^{p,\alpha}_\sigma(A_{r_1, r_2}) \) norm (see also Lemma 1 in [35]):

Lemma B.2. Let \( B_x \) be the open ball \( \{y; |x - y| < |x|/2 \} \) in \( A_{r_1, r_2} \), and \( \check{C} > 0 \) be as in Theorem 5.2, then the two norms \( C^{p,\alpha}_\sigma(A_{r_1, r_2}) \) and \( C^{p,\alpha}_\sigma(B_x) \) have the following relations:

\[
\eta^{-1}(|x|)||f||_{C^{p,\alpha}_\sigma(B_x)} \leq 9\check{C}||f||_{C^{p,\alpha}_\sigma(A_{|\frac{1}{2}| |x|})} \quad (\sigma \leq 0)
\]

and

\[
4^{-(p+\alpha-\sigma)}||f||_{C^{p,\alpha}_\sigma(A_{r_1, r_2})} \leq \sup_{B_x \subset A_{r_1, r_2}} \eta^{-1}(|x|)||f||_{C^{p,\alpha}_\sigma(B_x)} \quad (\sigma \leq 0).
\]

Proof. For \( j \leq p \) and \( \sigma \leq 0 \), let \( y \) and \( z \) be points in \( B_x \), and \( d_y = d(y, \partial B_x) \),

\[
d_{y}^{-\sigma+j} |D^j f(y)| \leq \check{C} \eta(|x|) \eta^{-1}(|y|) |y|^{-\sigma+j} |D^j f(y)|.
\]

If \( |y - z| < \max(1/4|y|, 1/4|z|) \) \( (y \neq z) \), then

\[
d_{y_z}^{-\sigma+p+\alpha} |D^p f(y) - D^p f(z)| \frac{1}{|y - z|^\alpha}
\]

\[
\leq \check{C} \eta(|x|) \min(\eta^{-1}(y)|y|^{-\sigma+p+\alpha}, \eta^{-1}(z)|z|^{-\sigma+p+\alpha}) |D^p f(y) - D^p f(z)| \frac{1}{|y - z|^\alpha}
\]

\[
\leq \check{C} \eta(|x|) ||u||_{C^{p,\alpha}_\sigma(A_{|\frac{1}{2}| |x|})}.
\]
If $|y - z| \geq \max(1/4|y|, 1/4|z|)$, then
\[
d_{yz}^{-\sigma + p + \alpha} \frac{|D^p f(y) - D^p f(z)|}{|y - z|^\alpha}
\leq 4^\alpha \hat{C} \eta(|x|)(\eta^{-1}(y)|y|^{-\sigma + p}|D^p f(y)| + \eta^{-1}(z)|z|^{-\sigma + p}|D^p f(z)|)
\leq 8\hat{C} \eta(|x|)||u||_{C^p,\alpha}^{(A_{|x| \leq 3|x|})}.
\]
Therefore, the first inequality follows directly from the definition of the $C^p,\alpha(B_x)$ norm. The second inequality is a consequence of the following two inequalities:
\[
\eta^{-1}(|x|)||f||_{C^p,\alpha(B_x)} \geq \sum_{j=0}^p \eta^{-1}(|x|) d_x^{j-\sigma} |D^j f|
\geq 2^{\sigma - p} \sum_{j=0}^p \eta^{-1}|x|^{j-\sigma} |D^j f|,
\]
\[
\eta^{-1}(|x|)||f||_{C^p,\alpha(B_x)} \geq \eta^{-1}(|x|)||f||_{C^p,\alpha(B_x)} + 4^{\sigma - p - \alpha} \eta^{-1}|x|^{-\sigma + \alpha} \sup_{\{z:|x-z|<|x|/4\}} \frac{|D^p f(x) - D^p f(z)|}{|x - z|^\alpha}.
\]

Finally, we give the proof of Theorem 5.2.

Proof. By applying the first inequality in Lemma B.2 with $\eta \equiv 1$, it follows that
\[
||a_{k,l}||_{C^p,\alpha(B_x)}, ||a_k||_{C^{p,1,\alpha}(B_x)}, ||a||_{C^{p,2,\alpha}(B_x)} \leq 9K_1.
\]
From Theorem B.1, we have
\[
||u||_{C^{p+2,\alpha}(B_x)} \leq K(||u||_{C^p,\alpha}(B_x) + ||f||_{C^{p,\alpha}(B_x)}).
\]
Using the first inequality in Lemma B.2 again, we get
\[
\eta^{-1}(|x||u||_{C^{p+2,\alpha}(B_x)} \leq 9\hat{C}K(||u||_{C^p,\alpha}(A_{r_1, r_2}) + ||f||_{C^{p,\alpha}(A_{r_1, r_2})}),
\]
for all $B_x \subset A_{r_1, r_2}$.

Hence the desired estimate is an immediate consequence of the above inequality and the second inequality in Lemma B.2. \qed

**Appendix C. Proof of Lemma 4.8**

In this section, we shall prove Lemma 4.8 by modifying the arguments used in the proof of the case that the domain is $\mathbb{R}^4$ [34]. We remark that by similar arguments we can prove the removable singularity Theorem 2.12. For the definition of ALE manifolds and the regularity of the metrics of Ricci flat ALE manifolds, we refer to e.g. [4, 11]. In the following, we shall only prove the case of extrinsic biharmonic maps and it is easy to see that the case of
intrinsic biharmonic maps can be shown by slightly modifying the arguments as in Section 7, see also Section 7 of [34].

**Proof.** Since \((M_a, h_a)\) is an Asymptotically Locally Euclidean (ALE) Ricci flat manifold (orbifold) of order 4 (see Theorem 2.1), by definition, there exists a compact subset \(K\), a finite subgroup \(\Gamma \subset O(n)\) and a \(C^\infty\) diffeomorphism \(\Phi : M_a \setminus K \to (\mathbb{R}^4 \setminus B(0;R))/\Gamma\) such that \(\varphi = \Phi^{-1} \circ \text{proj}\) satisfies (where \(\text{proj}\) is the natural projection of \(\mathbb{R}^4\) to \(\mathbb{R}^4/\Gamma\))

\[
\varphi^* h_a (z) = \delta_{ij} + O(|z|^{-4}), \quad \partial_k (\varphi^* h_a)_{ij} (z) = O(|z|^{-5}),
\]

for \(z, w \in \mathbb{R}^n \setminus B(0;R)\), where \(\partial_k\) denotes \(\partial/\partial z_k\). Since \(M_a\) is Ricci flat, by elliptic estimates we obtain that

\[
\partial^\alpha (\varphi^* h_a)_{ij} (z) = O(|z|^{-(4+|\alpha|)})
\]

for any \(\alpha\) such that \(|\alpha| > 0\).

It is easy to see that we only need to prove the lemma for biharmonic maps on the domain \((\mathbb{R}^4 \setminus B(0,R), \varphi^* h_a)\). For simplicity, we write \(g\) for \(\varphi^* h_a\). By a scaling argument, we can assume that \(R_a = 1\). Set

\[
\mathbb{R}^4 \setminus B_1 = \bigcup_{l=0}^\infty A_l,
\]

where \(A_l = B_{e^{Ll}} \setminus B_{e^{(l-1)L}}\). Let \(\rho = |z| \in [e^{(l-1)L}, e^{Ll}]\), then we have

\[
\|g_{ij}(\rho y) - \delta_{ij}\| < \rho^{-4}
\]

and

\[
\partial^\alpha (g_{ij}(\rho y) = O(|\rho|^{-(4+|\alpha|)}),
\]

where \(z = \rho y\). Namely, the metric after scaling to \(B_2 \setminus B_1\) is close to the flat metric in \(C^k\) norm for any \(k \in \mathbb{N}\).

Since

\[
\int_{\mathbb{R}^4 \setminus B_1} |\nabla^2 g u|^2 + |\nabla g u|^4 dV_g < \epsilon_1,
\]

we can assume that for any \(\epsilon > 0\), there exists \(l_0\) such that

\[
\int_{A_l} |\nabla^2 g u|^2 + |\nabla g u|^4 < \epsilon^2 < \epsilon_0
\]

holds for \(l > l_0\).

By direct computations as in the proof for Lemma 3.3 (with the use of (75) and (76)), the biharmonic map equation (15) (for \(u(\rho y)\)) in local coordinates on \(B_2 \setminus B_1\) is equivalent to

\[
\Delta^2 u = \alpha_1(u) \nabla \Delta u \nabla u + \alpha_2(u) \nabla^2 u \nabla^2 u + \alpha_3(u) \nabla^2 u \nabla u + \alpha_4(u) \nabla u \nabla u + \tilde{\tau}(u).
\]

Here \(\tilde{\tau}(u)\) satisfies

\[
\|\tilde{\tau}(u)\|_{LP(B_2 \setminus B_1)} \leq C \epsilon \eta
\]
for any $p > 1$, where $\eta = O(\rho^{-4})$. After scaling back to the scale of $A_\ell$, the equation (15) for $u(z)$ in local coordinates becomes

$$\nabla^2 u = \alpha_1(u) \nabla \nabla u + \alpha_2(u) \nabla^2 \nabla^2 u + \alpha_3(u) \nabla^2 \nabla u + \alpha_4(u) \nabla^3 \nabla u + \tau(u).$$

In the above $\tau(u(z)) = \rho^{-4} \tilde{\tau}(u(\rho y))$.

As in the proof of Lemma 3.3, $v = u - u^*(r)$ ($|z| = r$) is an approximate biharmonic function in the sense of Definition 3.1, and

$$(78) \quad F_l(v) \leq C \varepsilon^2 e^{-lL}$$

for all $l > l_0$. In the above

$$u^*(r) = \frac{1}{|\partial B_r|} \int_{\partial B_r} u(r, \theta) d\sigma.$$

Then by (78), Lemma 3.2 and the Sobolev embedding theorem, we have for any $p + q \leq 3$, $q \geq 1$ and $t \in [0, \infty)$,

$$(79) \quad |\partial_t \nabla^{q+1} u|^2(t, \theta) \leq C \varepsilon e^{-t/2}.$$

The next step is to control the radial part energy by a Pohožaev type argument. Recall that $u$ is an approximate biharmonic map if and only if $\nabla^2 u - \tau(u)$ is normal to the tangent space $T_u N$. On the other hand, $\partial_t u$ is a tangent vector at $u(\cdot) \in N$. Therefore,

$$\int_{S^3} (\nabla^2 u - \tau(u)) \partial_t u d\theta = 0$$

for all $t$, where $d\theta$ is the volume element of $S^3$. Then by the equation

$$\nabla^2 u = e^{-4t} ((\partial_t + \nabla_{S^3})^2 - 4 \partial_t^2) u$$

and some computations as in Section 5 of [34], we can obtain

$$(80) \quad \partial_t G(t) = \int_{S^3} e^{4t} \tau(u) \partial_t u d\theta,$$

where

$$G(t) = \int_{S^3} 2 \partial_t u \partial_t^2 u - |\partial_t^2 u|^2 + |\Delta_{S^3} u|^2 - 2|\partial_t \nabla_{S^3} u|^2 - 4|\partial_t u|^2 d\theta.$$  

Now as in Section 5 of [34], we can prove that

$$G(t) = \int_{S^3} 2 \partial_t u \partial_t^2 u - |\partial_t^2 u|^2 + |\Delta_{S^3} u|^2 - 2|\partial_t \nabla_{S^3} u|^2 - 4|\partial_t u|^2 d\theta$$

is zero when $t \to \infty$. Indeed by the $\varepsilon$-regularity theorem (Theorem 2.9) and the Sobolev embedding theorem, we have for $1 \leq k \leq 3$,

$$\max_{\mathbb{R}^4 \setminus B_\rho} \rho^k |\nabla^k u| = o(1)$$

as $\rho \to \infty$.  

Therefore we have
\[
\partial_t \left( \int_{S^3} \partial_t u \partial_t^2 u \, d\theta \right) - \int_{S^3} 3/2 |\partial_t^2 u|^2 + 2 |\partial_t u|^2 \, d\theta
\]

(81)
\[
= \Theta(t) + \int_t^\infty \int_{S^3} e^{4t} \tau(u) \partial_t u \, d\theta \, ds,
\]
where
\[
\Theta(t) = \int_{S^3} -1/2 |\Delta_{S^3} u|^2 + |\partial_1 \nabla_{S^3} u|^2 \, d\theta.
\]
By using (79), \(|\Theta| \leq C\varepsilon^2 e^{-t}.

For some fixed \(t_0 \in [1, \infty)\), set
\[
F(t) = \int_{t_0-t}^{t_0+t} \int_{S^3} 3/2 |\partial_t^2 u|^2 + 2 |\partial_t u|^2 \, d\theta \, dt.
\]
\(F\) is defined for \(0 \leq t \leq t_0\). Integrating (81) from \(t_0 - t\) to \(t_0 + t\), we obtain
\[
F(t) \leq \frac{1}{2 \sqrt{3}} \left( \int_{\{t_0-t\} \times S^3} + \int_{\{t_0+t\} \times S^3} \right) 3/2 |\partial_t^2 u|^2 + 2 |\partial_t u|^2 \, d\theta
\]
\[
+ \int_{t_0-t}^{t_0+t} |\Theta(s)| \, ds + \int_{t_0-t}^{t_0+t} \int_s^\infty \int_{S^3} e^{4t} |\tau(u) \partial_t u| \, d\theta \, ds
\]
By direct computations (transforming back to \(z\)-coordinates), we have
\[
\int_{t_0-t}^{t_0+t} \int_s^\infty \int_{S^3} e^{4s} |\tau(u) \partial_t u| \, d\theta \, ds \, d\tilde{s}
\]
\[
\leq \int_{e^{t_0-t}}^{e^{t_0+t}} |\tau(u)| \|\nabla u\| \, dz \, dr
\]
\[
\leq C \int_{e^{t_0-t}}^{e^{t_0+t}} |\tau(u)| \|\nabla u\|_{L^4/3(\mathbb{R}^4 \setminus B_{r^*})} \|\nabla u\|_{L^4(\mathbb{R}^4 \setminus B_{r^*})} \, ds
\]
\[
\leq C \int_{e^{t_0-t}}^{e^{t_0+t}} |\tau(u)| \|\nabla u\|_{L^4/3(\mathbb{R}^4 \setminus B_{r^*})} \, ds.
\]
In the above inequality, we have used the assumption that \(\varepsilon_1 > 0\) is small.

Now we claim the following

Claim B:
\[
|\tau(u)| \|\nabla u\|_{L^4/3(\mathbb{R}^4 \setminus B_{r^*})} \leq C \varepsilon e^{-s}.
\]

By using the above estimate, we have
\[
F(t) \leq 1/2 \partial_t F + C \varepsilon^2 e^{t-t_0}.
\]
Multiplying \(-e^{-2t}\) to both sides of the above inequality and integrating from \(t = 1\) to \(t = t_0 - 1\), we get
\[
F(1) \leq e^{-2(t_0-1)} F(t_0 - 1) + C \varepsilon^2 e^{-t_0-1}.
\]
Thanks to the above inequality and (79), we can obtain the desired decay estimate for the biharmonic energy. Next, we can argue exactly as in Section 5 of [34] to show a decay estimate

$$\max_{|z|=e^{t_0}} |z| |\nabla u| \leq C \varepsilon |z|^{-1/2},$$

by which we can complete the proof of the lemma.

Finally, we prove Claim B.

Proof. Set $s = (l_s - 1)L$. Notice that $\tau = \rho^{-4}\tilde{\tau}$. Since $\eta(\rho) = C \rho^{-4}$ in (77), we have

$$\int_{\mathbb{R}^4 \setminus B_{s}} |\tau(u)|^{4/3}dz \leq \sum_{l=l_s}^{\infty} \int_{A_l} |\tau(u)|^{4/3}dz \leq \sum_{l=l_s}^{\infty} C e^{lL \times (4 - 4 \times 4/3)} \int_{B_1 \setminus B_{e^{-L}}} |\tilde{\tau}(u)|^{4/3}dy \leq \sum_{l=l_s}^{\infty} C \varepsilon e^{lL \times (4 - 4 \times 4/3 - 4 \times 4/3)} \leq \sum_{l=l_s}^{\infty} C \varepsilon e^{(-20/3)L} < C \varepsilon e^{-s}.$$
[12] Youmin Chen and Hao Yin. Uniqueness of tangent cone for biharmonic map with isolated singularity. Pacific Journal of Mathematics, 299(2):401–426, 2019.

[13] Wei Yue Ding and Gang Tian. Energy identity for a class of approximate harmonic maps from surfaces. Comm. Anal. Geom., 3(3-4):543–554, 1995.

[14] S. K. Donaldson. An application of gauge theory to four-dimensional topology. J. Differential Geom., 18(2):279–315, 1983.

[15] S. K. Donaldson and P. B. Kronheimer. The geometry of four-manifolds. Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1990.

[16] Avron Douglis and Louis Nirenberg. Interior estimates for elliptic systems of partial differential equations. Communications on Pure and Applied Mathematics, 8(4):503–538, 1955.

[17] Daniel S. Freed and Karen K. Uhlenbeck. Instantons and four-manifolds. Second edition. Mathematical Sciences Research Institute Publications, 1. Springer-Verlag, New York, 1991.

[18] Andreas Gastel. The extrinsic polyharmonic map heat flow in the critical dimension. Advances in Geometry, 6(4):501–521, 2006.

[19] David Gilbarg and Neil S. Trudinger. Elliptic partial differential equations of second order. Springer, 2003.

[20] Detlef Gromoll, Wilhelm Klingenberg, and Wolfgang Meyer. Riemannsche geometrie im grossen, volume 55. Springer-Verlag, 2006.

[21] Mikhael Gromov. Groups of polynomial growth and expanding maps. Inst. Hautes Etudes Sci. Publ. Math., 53:53–73, 1981.

[22] Mikhael Gromov. Pseudo holomorphic curves in symplectic manifolds. Inventiones Mathematicae, 82(2):307–347, 1985.

[23] Karsten Grove and Peter Petersen. Manifolds near the boundary of existence. Journal of Differential Geometry, 33(1991):379–394, 1991.

[24] John Happel and Howard Brenner. Low Reynolds number hydrodynamics with special applications to particulate media. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1965.

[25] Peter Hornung and Roger Moser. Energy identity for intrinsically biharmonic maps in four dimensions. Analysis & PDE, 5(1):61–80, 2012.

[26] Juergen Jost. Two-dimensional geometric variational problems. Pure and Applied Mathematics (New York). A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1991.

[27] Atsushi Kasue. A compactification of a manifold with asymptotically nonnegative curvature. In Annales scientifiques de l'Ecole normale supérieure, volume 21, pages 593–622. Elsevier, 1988.

[28] Atsushi Kasue. A convergence theorem for riemannian manifolds and some applications. Nagoya Mathematical Journal, 114:21–51, 1989.

[29] Tobias Lamm and Tristan Rivière. Conservation laws for fourth order systems in four dimensions. Communications in Partial Differential Equations, 33(2):245–262, 2008.

[30] Paul Laurain and Tristan Rivière. Energy quantization for biharmonic maps. Advances in Calculus of Variations, 6(2):191–216, 2013.

[31] John M. Lee and Thomas H. Parker. The Yamabe problem. Bulletin of the American Mathematical Society, 17(1):37–91, 1987.

[32] Fanghua Lin and Changyou Wang. Energy identity of harmonic map flows from surfaces at finite singular time. Calc. Var. Partial Differential Equations, 6(4):369–380, 1998.

[33] Lei Liu and Hao Yin. On the finite time blow-up of biharmonic map flow in dimension four. Journal of Elliptic and Parabolic Equations, 1(2):363–385, 2015.

[34] Lei Liu and Hao Yin. Neck analysis for biharmonic maps. Mathematische Zeitschrift, 283(3-4):807–834, 2016.

[35] Norman Meyers. An expansion about infinity for solutions of linear elliptic equations. J. Math. Mech., 12:247–264, 1963.

[36] Roger Moser. Remarks on the regularity of biharmonic maps in four dimensions. Communications on Pure and Applied Mathematics, 59(3):317–329, 2006.
[37] H. Nakagawa. Taiiki no riemann kikagaku (riemannian geometry in the large). Kaigai Shuppan Boeki: Tokyo, 1977.

[38] Hiraku Nakajima. Hausdorff convergence of einstein 4-manifolds. J. Fac. Sci. Univ. Tokyo, 35:411–424, 1988.

[39] Hiraku Nakajima. A convergence theorem for einstein metrics and the ale spaces. Translations of the American Mathematical Society-Series 2, 160:79–94, 1994.

[40] Thomas H. Parker. Bubble tree convergence for harmonic maps. Journal of Differential Geometry, 44(3):595–633, 1996.

[41] Thomas H. Parker and Jon G. Wolfson. Pseudo-holomorphic maps and bubble trees. J. Geom. Anal., 3(1):63–98, 1993.

[42] Jie Qing and Gang Tian. Bubbling of the heat flows for harmonic maps from surfaces. Communications on Pure and Applied Mathematics, 50(4):295–310, 1997.

[43] Tristan Rivière. Interpolation spaces and energy quantization for Yang-Mills fields. Comm. Anal. Geom., 10(4):683–708, 2002.

[44] Jonathan Sacks and Karen Uhlenbeck. The existence of minimal immersions of 2-spheres. Ann. of Math, 113(1):1–24, 1981.

[45] Christoph Scheven. Dimension reduction for the singular set of biharmonic maps. Adv. Calc. Var., 1(1):53–91, 2008.

[46] Richard Schoen and Shing-Tung Yau. Lectures on differential geometry, volume 2. International press Cambridge, MA, 1994.

[47] Leon Simon. Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems. Ann. of Math. (2), 118(3):525–571, 1983.

[48] Leon Simon. Theorems on regularity and singularity of energy minimizing maps. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1996. Based on lecture notes by Norbert Hungerbühler.

[49] P. D. Smith and Deane Yang. Removing point singularities of riemannian manifolds. Transactions of the American Mathematical Society, 333(1):203–219, 1992.

[50] Michael Struwe. Partial regularity for biharmonic maps, revisited. Calc. Var. Partial Differential Equations, 33(2):249–262, 2008.

[51] Paweł Strzelecki. On biharmonic maps and their generalizations. Calculus of Variations and Partial Differential Equations, 18(4):401–432, 2003.

[52] Gang Tian. On calabi’s conjecture for complex surfaces with positive first chern class. Invent. Math., 101(1):101–172, 1990.

[53] Changyou Wang. Biharmonic maps from R^4 into a Riemannian manifold. Math. Z., 247(1):65–87, 2004.

[54] Changyou Wang. Remarks on biharmonic maps into spheres. Calc. Var. Partial Differential Equations, 21(3):221–242, 2004.

[55] Changyou Wang. Stationary biharmonic maps from R^m into a Riemannian manifold. Comm. Pure Appl. Math., 57(4):419–444, 2004.

[56] Changyou Wang. Heat flow of biharmonic maps in dimensions four and its application. Pure and Applied Mathematics Quarterly, 3(2):595–613, 2007.

[57] Changyou Wang and Shenzhou Zheng. Energy identity for a class of approximate biharmonic maps into sphere in dimension four. Discrete and Continuous Dynamical Systems, 33(2):861–878, 2012.

[58] Changyou Wang and Shenzhou Zheng. Energy identity of approximate biharmonic maps to Riemannian manifolds and its application. Journal of Functional Analysis, 263(4):960–987, 2012.

[59] Guofang Wei and William Wylie. Comparison geometry for the bakry-emery ricci tensor. Journal of Differential Geometry, 83(2):377–406, 2009.

[60] J. G. Wolfson. Gromov’s compactness of pseudoholomorphic curves in symplectic geometry. Journal of Differential Geometry, 28(1988):383–405, 1988.

[61] Rugang Ye. Gromov’s compactness theorem for pseudo holomorphic curves. Trans. Amer. Math. Soc., 342(2):671–694, 1994.

[62] Miaomiao Zhu. Harmonic maps from degenerating riemann surfaces. Mathematische Zeitschrift, 264(1):63, 2010.
[63] Shunhui Zhu. The comparison geometry of ricci curvature. *In Comparison geometry, volume 30 of Math. Sci. Res. Inst. Pub*, pages 221–262, 1997.

SCHOOL OF MATHEMATICAL SCIENCES, SHANGHAI JIAO TONG UNIVERSITY, 800 DONGCHUAN ROAD, SHANGHAI, 200240, P. R. CHINA
Email address: chyoumin19@sjtu.edu.cn

SCHOOL OF MATHEMATICAL SCIENCES, SHANGHAI JIAO TONG UNIVERSITY, 800 DONGCHUAN ROAD, SHANGHAI, 200240, P. R. CHINA
Email address: mizhu@sjtu.edu.cn