ENERGY IDENTITY FOR STATIONARY YANG MILLS

AARON NABER AND DANIELE VALTORTA

ABSTRACT. Given a principal bundle $P \to M$ over a Riemannian manifold with compact structure group G, let us consider a stationary Yang-Mills connection A with energy $\int_M |F_A|^2 \leq \Lambda$. If we consider a sequence of such connections A_i, then it is understood by [Tia00] that up to subsequence we can converge $A_i \to A$ to a singular limit connection such that the energy measures converge $|F_{A_i}|^2 dv_g \to |F_A|^2 dv_g + \nu$, where $\nu = e(x) d\lambda^{n-4}$ is the $n-4$ rectifiable defect measure. Our main result is to show, without additional assumptions, that for $n-4$ a.e. point the energy density $e(x)$ may be computed explicitly as the sum of the bubble energies arising from blow ups at x. Each of these bubbles may be realized as a Yang Mills connection over S^4 itself.

This energy quantization was proved in [Riv02] assuming a uniform L^1 hessian bound on the curvatures in the sequence. In fact, our second main theorem is to show this hessian bound holds automatically. Precisely, given a connection A as above we have the apriori estimate $\int_M |\nabla^2 F_A| < C(\Lambda, \dim G, M)$ for the curvature. It is important to note this result is proved in tandem with the energy quantization, and not before it. Indeed, we will in fact prove an effective version of the energy identity, and it is this effective version which will lead to both the L^1 hessian bound and the classical energy quantization results. In the course of the proof we will provide a quantitative version of the bubble tree decomposition which hold in all dimensions with effective estimates for a fixed stationary connections. To produce to strongest estimates in the paper we introduce an ϵ-gauge condition, which generalizes the usual Coulomb gauge and which will exist, with effective control, even over singular regions. On these ϵ-gauges we will provide a new superconvexity estimate which will be a key tool in analyzing higher dimensional annular regions.

CONTENTS

1. INTRODUCTION

This paper is focused on studying principle bundles $P \xrightarrow{G} M$ over Riemannian manifolds (M^n, g) with compact structure groups $G \subseteq SO(k)$. A Yang-Mills connection A on P is a critical point of the Yang-Mills L^2 curvature functional

$$\mathcal{F}[A] \equiv \int |F_A|^2. \tag{1.1}$$

Most of the results of this paper are local in nature, and therefore it will be sufficient to consider connections over some ball $B_2 \subseteq M$. For technical simplicity we will restrict ourselves to smooth critical points of (1.1), however all of our techniques generalize to stationary points in more general singular contexts (e.g. admissible connections in the sense of [TT04] or stationary connections in the sense of [RPb],[RPc]), and we will make comments on the necessary ingredients to make such generalizations, which are fairly straight forward.

It is often times the case one is interested in not just a fixed connection, but in sequences and limits of such connections. For instance, when studying moduli spaces [Don89],[DT98] of connections, or when considering contradiction arguments. This is the appropriate context to study the energy identity. Indeed, if A_i are a sequence of Yang-Mills connections satisfying the uniform energy bound $\int |F_A|^2 \leq \Lambda$, then after passing to a subsequence it is known by [Uhl82],[TT04],[RPa] that we can converge the A_i modulo gauge

$$A_i \to A,\tag{1.2}$$

to a connection A which is smooth away from a set of $n - 4$ measure zero. More than that, by [Tia00] we can limit the energy measures

$$|F_{A_i}|^2 dv_g \to |F_A|^2 dv_g + \nu = |F_A|^2 dv_g + e(x)\, d\lambda^{n-4}\big|_S\,,\tag{1.3}$$

where $\nu = e(x)\, d\lambda^{n-4}$ is the $n - 4$ rectifiable defect measure supported on $S = \mathrm{supp}[\nu]$.

The goal of this paper is then two fold. We wish to study better the regularity properties of stationary Yang-Mills connections, and we wish to understand better the defect measure $e(x)d\lambda^{n-4}$ which arises as the singular part of the limit of energy measures from a sequence of Yang-Mills connections.

To state accurately our results on the structure of $e(x)$ let us first define with some accuracy the notion of a bubble. Notationally, let us remark that if A is a connection on the pointed manifold (M, g, x), then we write $r^{-1}A$ to denote the induced connection on $(M, r^{-2}g, x)$. This has the effect of rescaling the ball $B_r(x) \to B_1(x)$ to unit size, and therefore *blows up A* at x at scale r. Given this we define the notion of bubbling:

Definition 1.1. We define the following:

(1) A bubble B is a smooth Yang Mills connection on $\mathbb{R}^n \times G$ which is invariant under translation with respect to some $n - 4$ subspace $\mathcal{L}_B \subseteq \mathbb{R}^n$. We define the energy of B to be $E[B] \equiv \int_{\mathcal{L}_B^\perp} |F_B|^2$.

(2) We say that B is a bubble at $x \in \mathrm{supp}\{\nu\}$ if there exists a sequence $x_i \to x$ and $r_i \to 0$ such that the blow ups converge $r_i^{-1}A_i \to B$. We denote by $\mathcal{B}[x]$ the collection of all bubbles at x.

Notice that if B is a bubble, then by restricting B to $\mathcal{L}_B^\perp$ and transforming conformally we may view B as a smooth Yang Mills connection on S^4. The bubbles at a point $x \in \mathrm{supp}[\nu]$ turn out to be related to energy density $e(x)$. It turns out that if $B_1, \ldots, B_k \in \mathcal{B}[x]$ are *distinct* bubbles at x then one can rather easily prove the inequality

$$e(x) \geq \sum_{B_j} E[B_j] = \sum_{B_j} \int_{\mathcal{L}_{B_j}^\perp} |F_{B_j}|^2\,.\tag{1.4}$$

In words, the energy density $e(x)$ is at least as large as the energy contribution of every bubble at x. It has been an open problem about whether this inequality is an equality. This was first considered and proved for four dimensional instantons. In higher dimensions, the first results were due to [Tia00], where for generalized instantons it was shown that for $n - 4$ a.e. point that $x \in \mathrm{supp}[\nu]$ we can indeed compute $e(x)$

explicitly by the energy identity

$$e(x) = \sum_{B_j \in \mathcal{B}[x]} \int_{\mathcal{L}_{B_j}^{\perp}} |F_{B_j}|^2 \,, \tag{1.5}$$

for a collection of distinct bubbles. There is a variety of work in the literature toward dropping the instanton assumption. In dimension four the energy identity (1.5) was proved in full generality in [Riv02]. In higher dimensions, the best result is by Riviere [Riv02], where it was shown that (1.5) holds if one additionally assumes a uniform L^1 bound on the hessian of the curvature. The idea of [Riv02] was in the spirit of [LR02], where a similar result was proved for harmonic maps, and exploited certain Lorentz space estimates. The first main result of this paper is to prove (1.5) in full generality, and in particular we drop the assumed L^1 hessian bound.

In fact, the second main result of this paper is to show that if A is a stationary Yang-Mills connection, then one does in fact automatically have the apriori L^1 hessian estimate

$$\int_M |\nabla^2 F_A| < C(\Lambda, k, M) \,, \tag{1.6}$$

where the dependence of C above on M is only on the C^2 geometry of M. It is worth noting that this result is proved in tandem with the energy identity of (1.5), and not before it. In fact, as the paper is arranged we shall prove the energy identity first. However, fundamentally both results will follow from an effective version of the energy identity, which will describe the breakup of a fixed stationary connection. This will be outlined in Section 1.3 and is described in rigor in Section 9.

1.1. **Main Result for Stationary Yang Mills Connections.**

Our regularity results are all local in nature, and therefore we will only ever consider connections on open balls in manifolds. Precisely, we consider in this subsection a Yang-Mills connection A living on a principle G-bundle $P \to B_2(p) \subseteq M$, with $G \subseteq SO(k)$ compact, which satisfies

$$|\sec| < K^2 \,,$$
$$\mathrm{inj}(p) > K^{-1} \,. \tag{1.7}$$

Our main result on the regularity of stationary Yang Mills connections is the following:

Theorem 1.2 (L^1 Hessian Estimate)**.** *Let A be a stationary Yang-Mills connection satisfying (1.7) and $\fint_{B_2} |F_A|^2 \leq \Lambda$. Then we have*

$$\fint_{B_1} |\nabla^2 F_A| < C(n, k, K, \Lambda) \,. \tag{1.8}$$

Remark 1.1. In fact one can easily see that the estimate does not depend on the lower injectivity radius bound by simply lifting to a local cover.

1.2. **Main Results for Weak Limits of Yang Mills.** Let us now discuss limits $A_i \to A$ of stationary Yang-Mills connections with uniformly bounded energy. In this case we consider the defect measure

$$|F_{A_i}|^2 dv_g \to |F_A|^2 dv_g + \nu, \tag{1.9}$$

where by [Tia00] we have that the defect measure $\nu = e(x)\lambda^{n-4}$ is $n-4$ rectifiable. Recall from Definition 1.1 the precise meaning of a bubble at a point x. Our main result is the following energy quantization, which tells us that we may compute the energy density $e(x)$ through the bubbles at x:

Theorem 1.3 (Energy Identity). *Let $A_i \to A$ be a limit of stationary Yang-Mills connections satisfying* (1.7) *and $\fint_{B_2} |F_{A_i}|^2 \leq \Lambda$, and let ν be the associated defect measure. Then for $n-4$ a.e. $x \in supp\{\nu\}$, there exists a finite collection of distinct bubbles $B_1, \ldots, B_k \in \mathcal{B}[x]$ such that*

$$e(x) = \sum_{B_j} E[B_j] = \sum_{B_j} \int_{\mathcal{L}_{B_j}^{\perp}} |F_{B_j}|^2 . \tag{1.10}$$

1.3. **Outline of Proofs and Techniques.** Let us now outline the paper and the techniques involved in the proofs of Theorem 1.2 and Theorem 1.3. Everything in this subsection is rough in nature, and is meant to convey intuition without being dragged down by the large number of necessary technical details needed for the rigorous statements.

To begin with some basics, by covering our ball $B_1(p)$ with balls of small radius we can always assume our bound $K << 1$ from (1.7) is very small. In particular, by writing in harmonic coordinates we can assume we are on a chart so that our ball is a Euclidean ball with metric g_{ij} which satisfies for all $\alpha < 1$

$$\|g_{ij} - \delta_{ij}\|_{C^{1,\alpha}} \leq C(n, \alpha)K . \tag{1.11}$$

The K in the above is not quite the same as the one in (1.7), however it is does tend to zero as the sectional curvature does. In fact, there is really little lost in just assuming we are working on Euclidean space itself, as the proof of the general case requires only some minor extra technical work of a non fundamental nature. We will do just that for the remainder of this outline.

The proof of the main Theorems will center around the two main decomposition theorems given as Theorem 9.1 and Theorem 11.1. The content of the quantitative annulus/bubble decomposition of Theorem 11.1 is to split a ball $B_1(p)$ into two primary types of pieces

$$B_1(p) \subseteq \bigcup_a \mathcal{A}_a \cup \bigcup_b \mathcal{B}_b , \tag{1.12}$$

where $\mathcal{B}_b \subseteq B_{r_b}(x_b)$ and $\mathcal{A}_a \subseteq B_{r_a}(x_a)$ are quantitative versions of bubble and annular regions, such that we have the quantitative content covering control

$$\sum r_a^{n-4} + \sum r_b^{n-4} \leq C . \tag{1.13}$$

This annulus/bubble decomposition will be the primary decomposition used toward the proof of the L^1 hessian estimate. For the energy identity we will rely on the quantitative bubble tree decomposition of Theorem

9.1, which we will describe after we have discussed the quantitative bubble and annular regions in more detail.

The quantitative bubble regions $\mathcal{B}_b \subseteq B_{r_b}(x_b)$ are relatively easy to analyze, however the quantitative annular regions $\mathcal{A}_a$ will require several new ideas and will take up the majority of our discussion. Let us begin with a brief discussion of our goals with both, and then we will turn our attention to the methods.

Recall that a bubble, see Definition 1.1, is a smooth solution to the Yang-Mills equation on $\mathbb{R}^n$ which is invariant under translation by some $n-4$ dimensional subspace $\mathcal{L} \subseteq \mathbb{R}^n$. Likewise, a quantitative bubble $\mathcal{B}_b$ should be a solution which is close to looking like a bubble in an appropriate sense. The actual definition, given in Section 4, requires a little work because one has to account for the possibility for bubbles inside of bubbles, however for our outline the main point to emphasize is that it follows directly from the definition that the quantitative bubbles $\mathcal{B}_b$ are regions which are nearly invariant by $\mathcal{L}$ and have uniformly bounded curvature:

$$r_b^{4-n} \int_{B_{r_b}} |F[\mathcal{L}]|^2 < \delta \,,$$

$$r_b^2 |F| \leq C \text{ on } \mathcal{B}_b \,. \tag{1.14}$$

Using elliptic estimates one can then obtain pointwise scale invariant hessian estimates $r_b^4 |\nabla^2 F_A| < C$ on the curvature, which in particular lead to the integral estimates

$$\int_{\mathcal{B}_b} |\nabla^2 F_A| \leq C r_b^{n-4} \,. \tag{1.15}$$

Since the above are straightforward we will take it all in a blackbox in this outline and refer to Section 4 for a more detailed description of the quantitative bubble regions.

To discuss quantitative annular regions let us begin with a review of δ-flat and δ-weakly flat balls. Namely, we say a ball $B_r(x)$ is δ-flat if one has the scale invariant curvature estimate $r^2 |F_A| < \delta$ in $B_r(x)$. This is essentially the strongest condition one might ask for on a ball, and in fact too strong for practical applications. A weaker condition is that of a δ-weakly flat ball. As in Definition 3.1 we say that $B_r(x)$ is δ-weakly flat if for some $n-4$ subspace $\mathcal{L} \subseteq B_r(x)$ we have the scale invariant estimate $d(x, \mathcal{L})^2 |F_A| < \delta$ on $A_{\delta r, r}(\mathcal{L}) \cap B_r(x)$. That is, $B_r(x)$ is δ-weakly flat if the curvature is small away from $\mathcal{L}$, however it is still quite possible to have large curvature concentration near $\mathcal{L}$ itself. This situation happens frequently, and is in fact typical when studying defect measures. One could attempt to cover $B_1(p)$ as in (1.12) by quantitative bubble regions and weakly flat balls, unfortunately such a covering cannot necessarily be built to satisfy the content estimate (1.13), which will be crucial.

A δ-annular region $\mathcal{A} \subseteq B_{2r}$ is a region which looks δ-weakly flat on many scales. Slightly more precisely, if $\mathcal{C}$ is a closed set of center points and $r_x : \mathcal{C} \to \mathbb{R}$ is a positive function then

$$\mathcal{A} \equiv B_{2r} \setminus \overline{B}_{r_x}(\mathcal{C}) \equiv B_{2r} \setminus \bigcup_{x \in \mathcal{C}} \overline{B}_{r_x}(x) \,. \tag{1.16}$$

There are variety of useful technical conditions given in Definition 5.1 in the definition of a δ-annular region, e.g. a Vitali condition on the balls, however the relevant assumptions to keep in mind is that for each center point $x \in \mathcal{C}$ and all $r_x < s < 2r$ we have that $B_s(x)$ is δ-weakly flat and that $\mathcal{C} \cap B_s(x)$ looks approximately like

the subspace $\mathcal{L}_x = x + \mathcal{L}$. Thus as claimed we have that annular regions are those which look weakly flat on a potentially arbitrary number of scales. We also have that $\mathcal{C}$ looks approximately like an $n - 4$ dimensional space, and it is convenient to define the packing measure $\mu = \sum r_x^{n-4}$ associated to it. In Theorem 5.3 we prove a structure theorem for annular regions, which from the analysis point of view is the most important in the paper. The main results of this structure theorem are the following:

$$cs^{n-4} < \mu(B_s(x)) < Cs^{n-4} \text{ for } r_x < s < 2r,$$

$$r^{4-n} \int_{\mathcal{A}} |F|^2 < \epsilon,$$

$$r^{4-n} \int_{\mathcal{A}} |\nabla^2 F| < \epsilon. \tag{1.17}$$

The first result is an Ahlfor's regularity type result on μ, which tells us that $\mathcal{C}$ approximates an $n - 4$ dimensional space in a strong sense. For the energy identity of Theorem 1.3 it is the second estimate above which plays the key role, while for the L^1 hessian estimate of Theorem 1.2 is it the third estimate which plays the important role. Let us first roughly see how to conclude the main theorems once (1.17) is known, and then the rest of the outline will focus on the proof of (1.17) itself.

The proof of the L^1 hessian estimate of Theorem 1.2 is now nothing more than a combination of the covering (1.12) with the content estimate (1.13) and the scale invariant integral estimates (1.15), (1.17). Indeed:

$$\int_{B_1} |\nabla^2 F| \leq \sum_a \int_{\mathcal{A}_a} |\nabla^2 F| + \sum_b \int_{\mathcal{B}_b} |\nabla^2 F| \leq C \Big(\sum_a r_a^{n-4} + \sum_b r_b^{n-4} \Big) \leq C. \tag{1.18}$$

To discuss the energy identity of Theorem 1.3 we describe a quantitative version, which is given in Theorem 9.2. To accomplish this we need the quantitative bubble tree decomposition of Theorem 9.1, which is a refinement of the decomposition of (1.12) (though in fact we prove the following decomposition first). Under the assumption that $B_1(p)$ is weakly flat with respect to $\mathcal{L} = \mathcal{L}^{n-4} \subseteq B_1$ we will decompose the ball

$$B_1(p) \subseteq \bigcup_a \mathcal{A}_a \cup \bigcup_b \mathcal{B}_b \cup \bigcup_c B_{r_c}(x_c), \tag{1.19}$$

where in addition to the content estimate (1.13) for the δ-annular and δ-bubble regions we also have the small content estimate $\sum_c r_c^{n-4} < \epsilon$. The key difference between the two decompositions is that by assuming B_1 is weakly flat and throwing out this set of small content we can assume every annular region $\mathcal{A}_a$ and every bubble region $\mathcal{B}_b$ are with respect to the *same* $n - 4$ plane $\mathcal{L}$. This does not hold for the decomposition in (1.12). More than that, the quantitative energy identity of Theorem 9.2 gives us that for each $q \in \mathcal{L}$ such that the 4-plane $\mathcal{L}_q^{\perp} \equiv q + \mathcal{L}^{\perp}$ satisfies $\mathcal{L}_q^{\perp} \cap \bigcup_c B_{r_c} = \emptyset$, then if $\mathcal{B}_q \equiv \mathcal{L}_q^{\perp} \cap \bigcup_b \mathcal{B}_b$ are the bubble regions which intersect the slice $\mathcal{L}_q^{\perp}$ we have the following:

$$\#\{\mathcal{A}_a \cap \mathcal{L}_q^{\perp} \neq \emptyset\}, \ \#\{\mathcal{B}_b \cap \mathcal{L}_q^{\perp} \neq \emptyset\} \leq N = N(n, k, \Lambda),$$

$$\Big| \int_{B_1(q)} |F_A|^2 - \int_{\mathcal{B}_q} |F_A|^2 \Big| < \epsilon. \tag{1.20}$$

The second estimate above is primarily due to the first estimate in (1.17), so that most slices in an annular region have small energy, and that the constant N above is independent of δ, so that a typical slice only intersects a bounded number of small energy regions. The δ-independent bound on N takes a bit of work, and morally follows because the covering is built to satisfy the correct nontriviality assumptions. More specifically, each time an annular region is intersected by $\mathcal{L}_q^\perp$ there is a corresponding bubble region which also gets intersected, and further each such bubble region contains some definite amount of energy, so that this may happen only a uniformly bounded number of times. To conclude the classical energy identity we study the $n-4$ rectifiable defect measure ν of a limiting sequence and observe that a.e. there exists a tangent measure and it is a multiple of the Hausdorff measure on some $\mathcal{L}$. In particular, by blowing up at such points we get that neighborhoods are arbitrarily weakly flat, and so we may apply (1.20) with $\epsilon \to 0$. This will conclude the classical energy identity. See Section 10 for more on this.

What is then left in our outline is to describe the proof of the structure theorem estimates of (1.17) for annular regions. For $n = 4$ there are several known ways of doing such estimates, unfortunately all such previous methods break down in higher dimensions. To describe our methods, which take place over Sections 5, 6, and 7 and are related to the arguments of [JN], we begin with a discussion of gauges. The choice of a good gauge is often crucial in regularity issues, and in the context of Yang Mills the standard gauge condition is a Coulomb gauge. Unfortunately, the existence of a Coulomb gauge is typically only over balls which are geometrically very simple. For instance, if $B_r(x)$ is δ-flat as above, then it is easy to prove the existence of a Coulomb gauge on $B_r(x)$. However, if $B_r(x)$ is only δ-weakly flat, then a Coulomb gauge simply need not exist on the ball.

Instead, in Section 6 we describe the notion of a harmonic ϵ-gauge, which in Theorem 6.2 we show does exist on weakly flat balls. To describe it recall that our structure group $G \subseteq SO(k)$ so that we have an induced vector bundle $E \to B_1$. In Definition 6.1 we say that sections $V^1, \ldots, V^k \in \Gamma(B_r, E)$ form a harmonic ϵ-gauge on $B_r(x)$ if the following hold:

$$\Delta V^a = 0\,,$$

$$|V^a| \le 1 + \epsilon\,, \quad \fint_{B_r} |\langle V^a, V^b \rangle - \delta^{ab}| < \epsilon\,, \quad r^2 \fint_{B_r} |\nabla V|^2 < \epsilon^2\,. \tag{1.21}$$

We call the sections an ϵ-gauge because the condition $\fint_{B_r} |\langle V^a, V^b \rangle - \delta^{ab}| < \epsilon$ only guarantees that away from a set of small measure that the V^a form an ϵ-orthonormal basis. On a δ-flat ball it is not so hard to see that the V^a form an actual ϵ-orthonormal basis at every point, however if the ball is only δ-weakly flat ball then this need not hold.

Section 6 and in particular Section 7 are dedicated to proving much more powerful estimates on ϵ-gauges than those in (1.21). Before describing these, let us outline how the V^a may be used to control the curvature on annular regions. Indeed, using the definition of curvature one may compute at any point the bounds

$$|F(V)| \le 2|\nabla^2 V|\,, \quad |\nabla^2 F(V)| \le 2|\nabla^4 V| + 2|\nabla F||\nabla V| + |F||\nabla^2 V|\,. \tag{1.22}$$

Imagine now that we are at a point x such that $|\langle V^a, V^b \rangle - \delta^{ab}| < \epsilon$ and $r^2|F|, r^3|\nabla F| < 1$, then we have

$$|F|^2(x) \leq 8 \sum |\nabla^2 V^a|^2, \quad |\nabla^2 F|(x) \leq 8 \sum |\nabla^4 V^a| + r^{-3} |\nabla V| + r^{-2}|\nabla^2 V|. \tag{1.23}$$

In particular, if $\mathcal{A}$ is an annular region and $d(x, \mathcal{C})$ is the distance from a point in $\mathcal{A}$ to the ball centers which were subtracted away, imagine we could prove

$$\int_{\mathcal{A}} d(x, \mathcal{C})^{-3} |\nabla V| < \epsilon, \tag{1.24}$$

$$\int_{\mathcal{A}} d(x, \mathcal{C})^{-2} |\nabla^2 V|, \quad \int_{\mathcal{A}} |\nabla^2 V|^2 < \epsilon, \tag{1.25}$$

and that $|\langle V^a, V^b \rangle - \delta^{ab}| < \epsilon$ in $\mathcal{A}$, then we will have finished the annular structure estimates of Theorem 5.3 given in (1.17) and hence the proof of our main Theorems.

We will indeed prove these integral gradient estimates on V, though the pointwise ϵ-orthogonality on $\mathcal{A}$ is a little too much to hope for. However, we will show in Theorem 6.4 that the V^a do at least form a legitimate vector bundle gauge on all of $\mathcal{A}$, and in Theorem 7.1 we will see they even form an ϵ-orthonormal basis away from a set whose $n-4$ content is less than ϵ. If one is careful, this turns out to be good enough because one can iterate on this bad set in order to eventually get the curvature bounds on all of $\mathcal{A}$.

The integral estimates on V and the ϵ-orthogonality on $\mathcal{A}$ all follow from the first estimate of (1.24), and therefore our main goal is to show this. Indeed, the various other integral estimates on the hessian of V follow from (1.24) in combination with ϵ-regularity theorems in $\mathcal{A}$, while the ϵ-orthonormality follows from (1.24) combined with a telescoping argument. We refer the reader to Section 7 for more on this, and focus now on the proof of (1.24).

In the proof of (1.24) we begin by defining a smoothing of $d(x, \mathcal{C})$. Indeed, recall the packing measure μ associated to the annular region, and let us define the Green's function and associated Green's distance function:

$$-\Delta G_\mu = \mu,$$
$$b^{-2} = G_\mu. \tag{1.26}$$

Note that if μ were exactly the $n-4$ Hausdorff measure on a $n-4$ dimensional subspace then b would be proportional to the distance to that subspace. Using the Ahlfor's regularity on μ we prove in Lemma 7.4 that in $\mathcal{A}$ we at least have the uniform estimates

$$cd(x, \mathcal{C}) < b(x) < Cd(x, \mathcal{C}), \text{ and } c < |\nabla b| < C, \tag{1.27}$$

and thus b is a legitimate smoothing of the distance $d(x, \mathcal{C})$.

Now let ϕ be a reasonable cutoff function with $\phi \equiv 1$ on $\mathcal{A}$, and let us define the scale invariant quantity

$$S(r) \equiv r \cdot r^{-3} \int_{b=r} |\nabla V||\phi||\nabla b|. \tag{1.28}$$

Then the estimate (1.24) is equivalent to the Dini estimate

$$\int_0^\infty \frac{S(r)}{r} < \epsilon \,. \tag{1.29}$$

In order to prove this we will show in Proposition 7.6 the following superconvexity:

$$r\frac{d}{dr}\Big(d\frac{d}{dr}S\Big) \geq (1-\epsilon)^2 S(r) - e(r) \,, \tag{1.30}$$

where $e(r) \leq \epsilon\mu(\{x \in \mathcal{C} : cr_x < r < Cr_x\})$. Note that if there was no error, then this superconvexity tells us that $S(r)$ decays and grows polynomially in r, which is more than enough for (1.29). In fact, this is exactly what happens in the four dimensional case. In the general case, we can at do an ode estimate using (1.30) in Proposition 7.7 in order to conclude

$$\int_0^\infty \frac{S(r)}{r} \leq \int_0^\infty \frac{e(r)}{r} \leq \epsilon \int_0^\infty \frac{\mu(\{x \in \mathcal{C} : cr_x < r < Cr_x\})}{r} \leq \epsilon\mu(\mathcal{C}) < \epsilon \,, \tag{1.31}$$

which finishes the proof of the Dini estimate, and hence the Theorems themselves.

2. Preliminaries

2.1. Stationary Yang Mills and Monotonicity.
Given a principal bundle $P \to M$ over a Riemannian manifold we can consider the Yang Mills functional, which associates to a connection A the L^2 curvature

$$\mathcal{F}[A] = \int_M |F_A|^2 \,. \tag{2.1}$$

A stationary Yang-Mills connection is one which is a critical point of the above functional. Such a connection solves the Yang-Mills equations

$$d_A F_A = 0 \,,$$
$$d_A^* F_A = 0 \,. \tag{2.2}$$

The terminology *stationary* comes from the fact that such a connection also solves the stationary equation

$$\mathrm{div}\Big(|F_A|^2 g_{ij} - 4\langle F_i, F_j\rangle\Big) = 0 \,. \tag{2.3}$$

By pairing the above with the radial vector field ∇d_x^2 from a point $x \in M$ we see as in [Pri83] that the scale invariant curvature functional

$$\theta_r(x) \equiv r^{4-n} \int_{B_r(x)} |F_A|^2 \,, \tag{2.4}$$

is a monotone quantity with

$$\frac{d}{dr}\theta_r(x) = 4r^{4-n} \int_{S_r(x)} \big|F_A[\partial_r]\big|^2 \,. \tag{2.5}$$

It will often be useful in the constructions to focus on the following, which measures the energy a whole ball:

$$\bar{\theta}_r(x) \equiv \sup_{y \in B_r(x)} \theta_r(y). \tag{2.6}$$

2.2. **Symmetry of Connections.** We briefly review the notion of symmetry in this subsection. In this paper we will only be interested in top dimensional symmetry (i.e., $k = n - 4$ in what follows). However we give the general definition as it is instructive:

Definition 2.1 ((k, ϵ)-Symmetry of Connections). If A is a connection on $B_1(p)$ with $K < \epsilon$ from (1.11), then we say A is (k, ϵ)-symmetric if the following hold:

(0) For the radial vector field $\partial_r = \nabla d_p$ we have $\fint_{B_1} r^4 |F[\partial_r]|^2 < \epsilon$.

(k) There exists a k-dimensional subspace $\mathcal{L}^k$ such that $\fint_{B_1} |F[\mathcal{L}]|^2 < \epsilon$.

Remark 2.1. We say A is 0-symmetric if only condition (0) holds above.

Using the monotone quantity $\theta_r(x)$ the following is a nice exercise, see for instance [Yu] or [NV]:

Theorem 2.2. *Let A be a stationary Yang-Mills with $\fint_{B_2} |F_A|^2 \leq \Lambda$. Then for each $\epsilon, \tau > 0$ there exists $\delta(n, \epsilon, \tau, \Lambda) > 0$ such that if $K < \delta$ from (1.11) and there exists $x_0, \ldots, x_k \in B_1$ with*

(1) *if $\mathcal{L}^\ell = span\{x_1 - x_0, \ldots, x_\ell - x_0\}$ then $d(x_{\ell+1}, \mathcal{L}^\ell) > \tau$,*

(2) *$|\theta_3 - \theta_\delta|(x_i) < \delta$,*

then B_1 is (k, ϵ)-symmetric.

2.3. **Defect Measures.** The first technical result we review is from [Tia00] and relates the symmetry of the defect measure to the symmetry of the converging connections:

Theorem 2.3 (Symmetries of Defect Measures [Tia00]). *Let $A_i \to A$ with $|F_{A_i}|^2 dv_{g_i} \to |F_A|^2 dv_g + \nu$ be converging stationary Yang-Mills connections with $\fint_{B_4} |F_A|^2 \leq \Lambda$. The following hold:*

(1) *If $\int_{B_2} |F_{A_i}[\mathcal{L}^k]|^2 \to 0$ and $K_i \to 0$ from (1.11) then A and ν are both translation invariant by $\mathcal{L}^k$. If $k = n - 4$ then A is smooth and $\nu \equiv c \lambda_{\mathcal{L}}^{n-4}$ is a constant multiple of the Hausdorff measure on $\mathcal{L}$.*

(2) *If A_i are ($n - 4, i^{-1}$)-symmetric wrt $\mathcal{L}^{n-4}$ then $A \equiv 0$ and $\nu \equiv c \lambda_{\mathcal{L}}^{n-4}$ is a constant multiple of the Hausdorff measure on $\mathcal{L}$.*

The above can be viewed as the basis for the basic regularity result:

Theorem 2.4 ([Tia00]). *Let $A_i \to A$ with $|F_{A_i}|^2 dv_{g_i} \to |F_A|^2 dv_g + \nu$ be converging stationary Yang-Mills connections with $\fint_{B_4} |F_A|^2 \leq \Lambda$. Then $\nu = e(x) \lambda_S^{n-4}$ is $n - 4$ rectifiable with density $e(x) > \epsilon_{n,k}$.*

2.4. **ϵ-regularity for Stationary Yang-Mills.** One of the key tools in the study of any nonlinear equation are ϵ-regularity theorems. In this subsection we will discuss two such theorems. The first is the classical ϵ-regularity theorem, which tells us that balls with small energy must be smooth. Precisely, we have the following:

Theorem 2.5 ([Uhl82]). *There exists an $\epsilon_{n,k} > 0$ such that if A is a stationary Yang-Mills connection with $K < \epsilon_{n,k}$ from (1.11) and $\theta(0,2) = 2^{4-n} \fint_{B_2} |F_A|^2 < \epsilon_{n,k}$, then $\sup_{B_1} |F_A|^2 \leq C(n,k) \fint_{B_2} |F_A|^2 \leq 1$.*

For convenience, we also introduce the concept of regularity scale at a point.

Definition 2.6. Given a smooth Yang-Mills connection A, we define the regularity scale $r_A(x)$ by

$$r_A(x) = \sup \left\{ s \geq 0 \;\; s.t. \;\; \forall y \in B_s(x) \;, \;\; |F_A(y)| \leq s^{-2} \right\}. \tag{2.7}$$

Note that this quantity is scale-invariant, and that an immediate consequence of the ϵ-regularity theorem is that $\theta(0,2r) \leq \epsilon_{n,k}$ implies $r_A(x) \geq r$.

Let us now discuss one further ϵ-regularity result which will play a role in our paper. The following tells us that instead of assuming the energy is small on a ball, we need only assume the energy is small in a sufficient number of directions. That is, if a ball is sufficiently symmetric in the sense of Definition 2.1 then automatically the energy is small and a smaller ball is smooth. Precisely:

Theorem 2.7. *Let A be a stationary Yang-Mills connection with $\fint_{B_4} |F_A|^2 \leq \Lambda$. Then for each $\epsilon > 0$ there exists $\delta(n, \Lambda, \epsilon) > 0$ such that if $K < \delta$ from (1.11) and B_2 is $(n-3, \delta)$ symmetric, then $\sup_{B_1} |F_A| \leq \epsilon$.*

Proof. We will only sketch the proof, as we refer to [CN13] for a verbatim argument in the nonlinear harmonic maps context. So indeed, assume for some $\epsilon > 0$ such a $\delta(n, \Lambda, \epsilon) > 0$ does not exist, so that we can find a sequence of connections A_i for which B_2 is $(n-3, \delta_i)$-symmetric, but the curvature is not uniformly bounded. We may pass to a subsequence to limit $A_i \to A$ with corresponding defect measure $|F_{A_i}|^2 dv_{g_i} \to |F_A|^2 dv_{\mathbb{R}^n} + \nu$. However, by theorem 2.3 we then have that both A and ν are now $n-3$ symmetric, which is to say invariant under translations of some $n-3$ subspace in $\mathbb{R}^n$. However, by theorem 2.4 we also have that ν is $n-4$ rectifiable. Combining these two points tells us that $\nu = 0$ must be trivial. Therefore A is smooth and reduces to a Yang-Mills connection on some $\mathbb{R}^3$, and is therefore itself a flat connection. Thus we have that $|F_{A_i}|^2 dv_{g_i} \to 0$. In particular, for far enough down the sequence we may apply the classical ϵ-regularity of Theorem 2.5 to get a contradiction, and thus prove the Theorem.   □

2.5. **Yang Mills in Four Dimensions.** In this subsection we recall a couple basic points of four dimensional solutions to the Yang Mills equations. The results of this subsection are well understood, even if they are packaged in a form which is not completely standard. In addition to the classical methods for proving such results, the methods of Sections 6 and 9 may be used to give distinct proofs which have the advantage of generalizing to higher dimensions (as we shall see).

We begin by discussing four dimensional annular regions. The notion of an annular region will be a central point to this paper, especially in higher dimensions where the analysis has been lacking to study such regions. In dimension four they are quite well understood, and so we briefly review them here. The

classical techniques used to prove the results of this section do not pass to higher dimensions, where the singularities are not isolated. However, it is still helpful for intuition to review the isolated singularity case, and additionally we will explicitly use these results in the study of finite regions in higher dimensions. The main result is the following:

Theorem 2.8 (Annular Regions in $n = 4$). *Let A be a stationary Yang-Mills connection on a four dimensional space with $r_+^{4-n} \int_{B_{2r_+}} |F_B|^2 \leq \Lambda$, and let $\mathcal{A} \equiv A_{r_-,r_+}(p)$. There exists $\epsilon(k) > 0$ such that if*

(1) $K < \epsilon$ from (1.11).
(2) $d_x^2 |F_A|(x) < \epsilon$ for $x \in A_{r_-,r_+}$ with $d_x \equiv d(p,x)$.

then there exists $\alpha(k) > 0$ and $C(k) > 0$ such that if $\delta \equiv \sup_{A_{r_-,2r_-}} |F_A| + \sup_{A_{r_+,2r_+}} |F_A|$ then we have the improved estimate for $x \in A_{r_-,r_+}$:

$$d_x^2 |F_A|(x) < C\left(\left(\frac{d_x}{r_+}\right)^\alpha + \left(\frac{r_-}{d_x}\right)^\alpha\right) \cdot \delta. \tag{2.8}$$

In particular, we have that $\int_{A_{r_-,r_+}} |F_A|^2 < C(k)\delta^2$ and $\int_{A_{r_-,r_+}} |\nabla^2 F_A| < C(k)\delta$

The classical method for proving results like the above is through a *three annulus* type lemma. The techniques in Section 7 also give a (more involved) proof of the above result, though has the advantage of working in higher dimensions.

We now turn our attention to more global information about four dimensional solutions. The content of the next result is to see that the energy of a bubble can concentrate on at most a finite number of regions.

Theorem 2.9. *Let B be a stationary Yang-Mills connection on $\mathbb{R}^4 \times G$ with finite energy $\int_{\mathbb{R}^4} |F_B|^2 \leq \Lambda$. We have $C(k,\Lambda), N(k,\Lambda)$ such that $\exists$ disjoint balls $\{B_{r_i}(c_i)\}_1^N \subseteq \mathbb{R}^4$ with $\int_{B_{r_i}} |F_B|^2 > \epsilon(k)$ such that for every $\eta > 0$ if $R \geq R(k,\Lambda,\eta)$ then*

$$\left| \int_{\mathbb{R}^4} |F_B|^2 - \int_{\bigcup B_{Rr_i}(c_i)} |F_B|^2 \right| < \eta. \tag{2.9}$$

Remark 2.2. There is a local version as well if A is a connection on $B_{2r}(0^4)$ with $\int_{B_{2r}} |F_A|^2 \leq \Lambda$. Then one has the estimate $\left| \int_{B_r} |F_B|^2 - \int_{B_r \cap \bigcup B_{Rr_i}(c_i)} |F_B|^2 \right| < \eta$.

In particular, if we consider a sequence of bubbles then the resulting limit may split into at most N independent limiting bubbles. Classically one could prove this by contradiction using a bubble tree argument and Theorem 2.8. The techniques of Section 11 can be used to give an effective proof, even in higher dimensions.

3. Weakly Flat Balls

In Section 2.4 it was discussed that the structure of a Yang-Mills connection A which is $(n-3, \delta)$-symmetric is quite trivial, namely A is smooth and ϵ-flat. In this subsection we want to study the structure of connections which are $(n-4, \delta)$-symmetric. In this case, A need not be trivially smooth, however the structure of A is still quite simple and will form the basic building block for the top stratum of the defect

measure. To understand this better let us discuss the notion of a weakly flat ball:

Definition 3.1 (δ-weakly flat). If $K < \delta$ from (1.11), then we say a connection A is δ-weakly flat on $B_1(p)$ with respect to $\mathcal{L}^{n-4}$ if

(1) For each $y \in \mathcal{L}_p \cap B_1$ and $\delta \leq r \leq 1$ we have $r^4 \fint_{B_r(y)} |F[\mathcal{L}]|^2 < \delta$.

(2) For $y \in B_1(p) \setminus B_\delta(\mathcal{L}_p)$ we have the estimate $|F_A|(y) \leq \delta \, d(y, \mathcal{L}_p)^{-2}$.

Remark 3.1. As with all definitions in this paper we will apply this in a scale invariant sense to any ball. That is, we call $B_r(p)$ a δ-weakly flat ball if after rescaling $r^{-1}B_r(p) \to B_1(\tilde{p})$ we have that the above holds.

Therefore a weakly flat ball does not have uniformly small curvature, but the curvature is quite small away from a neighborhood of a $n-4$ plane $\mathcal{L}$. The notion of a weakly flat ball will come into play at several stages, in particular in our defining of annular regions (one of our two main building blocks) in Section 5. Our first result, which follows quite easily from Theorem 2.7, tells us that balls which are very $n-4$ symmetric must be weakly flat:

Theorem 3.2 (Existence of Weakly Flat Balls). *Let A be a stationary Yang-Mills connection satisfying* (1.11) *with $\fint_{B_2} |F_A|^2 \leq \Lambda$ and let $\epsilon > 0$. Then there exists $\delta(n, k, K, \Lambda, \epsilon) > 0$ such that if B_2 is $(n-4, \delta)$-symmetric, then either:*

(1) $\sup_{B_{3/2}} |F| \leq \epsilon$, *or*

(2) B_1 *is ϵ-weakly flat and* $\fint_{B_1} |F|^2 > \epsilon_{n,k}$.

Remark 3.2. The constant $\epsilon_{n,k} > 0$ is from the ϵ-regularity of Theorem 2.5.

Let us now consider the following self-improvement theorem for weakly flat balls, which will be used for local estimates in the study of both quantitative bubble regions and annular regions. In short, it gives us a local pinching estimate which tells us that if $B_1(p)$ looks δ-weakly flat on many scales, then B_1 improves and it actually $10^{-2}\delta$-weakly flat.

Theorem 3.3 (Curvature Pinching of Weakly Flat Balls). *Let A be a stationary Yang-Mills connection with $0 < \kappa < 1$ and $\delta > 0$. There exists $c(k, \kappa) > 0$ and $\delta'(n, k, \Lambda, \kappa, \delta) > 0$ such that if $K < \delta'$ from* (1.11) *with $\fint_{B_{c^{-1}}} |F_A|^2 \leq \Lambda$, $\fint_{B_{c^{-1}}} |F_A[\mathcal{L}]|^2 < \delta'$ and such that $B_r(p)$ is δ-weakly flat ball wrt $\mathcal{L}$ for $c \leq r \leq c^{-1}$, then we have that $B_1(p)$ is $\kappa\delta$-weakly flat.*

If $n = 4$ the above can be viewed as a rewriting of Theorem 2.8. However, in higher dimensions the above result is morally much weaker because the bound on $\fint_{B_2} |F_A[\mathcal{L}]|^2$ depends on δ. In particular, iterating this result to obtain further improvements on more scales requires apriori better bounds in the $\mathcal{L}$-directions. Improving this estimate to a true generalization of Theorem 2.8 which is scale independent is accomplished in the structure theorem on annular regions in Section 5. Regardless, the above local version will be useful at several stages as a more coarse estimate.

We end with a key covering tool, which for us will be the dichotomy that either a ball is weakly flat, or away from a set of small $n-4$ content we must have the energy strictly drops. Precisely, our result is the following:

Theorem 3.4 (Weakly Flat Covering). *Let A be a stationary Yang-Mills connection with $K < \delta$ from* (1.11)*, $\fint_{B_{16}} |F_A|^2 \le \Lambda$. Then there exists $\eta(n, k, K, \Lambda, \delta) > 0$ such that either*

(1) *$B_4(p)$ is δ-weakly flat, or*

(2) *We can cover $B_1(p) \subseteq \bigcup_c B_{r_c}(x_c) \cup \bigcup_d B_{r_d}(x_d)$ such that*

    (a) *$\sum_c r_c^{n-4} < \delta$,*

    (b) *$\overline{\theta}_{r_d}(x_d) \le \overline{\theta}_1(p) - \eta$ with $\sum_d r_d^{n-4} < C(n, k, K, \Lambda, \delta)$.*

3.1. **Proof of Theorem 3.2.**

The proof is really just an application of theorem 2.3 with a contradiction argument. Thus let us assume for some $\epsilon > 0$ no such $\delta > 0$ exists. Then there exits a sequence A_i of connections which are $(n-4, \delta_i)$-symmetric with $\int_{B_2} |F_{A_i}|^2 \le \Lambda$. After passing to a subsequence we have

$$A_i \to A,$$
$$|F_{A_i}|^2 dv_{g_i} \to |F_A|^2 dv_g + \nu. \tag{3.1}$$

Using theorem 2.3 we have that $A \equiv 0$ with $\nu = c\lambda_{\mathcal{L}}^{n-4}$ a constant multiple of the $n-4$ Hausdorff measure on $\mathcal{L}$. In particular, for far enough into the sequence B_1 is ϵ-weakly flat.

To finish the proof we have two options, either $\nu(B_1) \ge \epsilon_{n,k}$ or not. In the first case we have (2), thus let us assume $\nu(B_1) < \epsilon_{n,k}$. In this case we have by Theorem 2.5 that A_i are uniformly smooth on $B_{1/2}$ sufficiently far in the sequence, and in particular we actually have $\nu \equiv 0$. Thus we have that $A_i \to 0$ smoothly on $B_{3/2}$, and for far enough into the sequence we have $|F| < \epsilon$ on the ball, which shows case (1) is satisfied. $\square$

3.2. **Proof of Theorem 3.3.**

For $\delta, \kappa > 0$ fixed let us assume no such $\delta'(n, k, \Lambda, \delta, \kappa) > 0$ exists. Thus we have a sequence of connections A_i with $\int_{B_{c^{-1}}} |F_{A_i}[\mathcal{L}]|^2 \to 0$ such that for $c \le r \le c^{-1}$ we have that $B_r(p)$ is δ-weakly flat. We will choose $c = c(n, \kappa) > 0$ before the end of the proof. Passing to a subsequence we can use theorem 2.3 to conclude that $A_i \to A$ and $|F_{A_i}|^2 dv_{g_i} \to |F_A|^2 dv_g + \nu$, where A defines a smooth $n-4$ symmetric Yang Mills connection on $B_{c^{-1}}(p)$ and $\nu = c\lambda_{\mathcal{L}}^{n-4}$. In particular, we can now apply theorem 2.8 to conclude for $c\delta \le d(x, \mathcal{L}) \le 1$ the estimate

$$d(x, \mathcal{L})^2 |F_A|(x) \le C(k)c(k, \kappa)^\alpha \left(\left(\frac{d_x}{1}\right)^\alpha + \left(\frac{\delta}{d_x}\right)^\alpha\right)\delta. \tag{3.2}$$

For $c(k, \kappa)$ sufficiently small we see that $B_1(p)$ is $\frac{1}{2}\kappa\delta$-weakly flat. In particular, for far enough in the sequence we have that $B_1(p_i)$ is $\kappa\delta$-weakly flat, which proves the Theorem. $\square$

3.3. **Proof of Theorem 3.4.** Let us define the set

$$E_\eta \equiv \{x \in B_1 : |\theta_{10} - \theta_\eta| < \eta\}. \tag{3.3}$$

Picking $\tau = \delta^2$ and $c(n) \equiv 10^{-6n}$ and using Theorem 2.2, if $\eta < \eta(n, k, \Lambda, \delta')$ and $Vol(B_{2\tau}E_\eta) > c(n)\delta\tau^4$ then we know B_8 is $(n - 4, \delta')$-symmetric, and hence by Theorem 3.2 that B_4 is δ-weakly flat so that (1) is satisfied. Therefore, we may assume this is not the case and choose a covering

$$B_{2\tau}(E_\eta) \subseteq \bigcup_c B_{r_c}(x_c), \tag{3.4}$$

where $r_c \equiv \tau$ and $\sum r_c^{n-4} < \delta$. Let us define $r_d \equiv \eta$, and note now that for every point $x \in B_1 \setminus \bigcup_c B_{r_c}(x_c)$ we have that $\overline{\theta}_{r_d}(x) \leq \overline{\theta}_1(p) - \eta$. By picking a Vitali covering of $B_1 \setminus \bigcup_c B_{r_c}(x_c) \subseteq \bigcup B_{r_d}(x_d)$ of such balls we have completed the construction. □

4. δ-BUBBLE REGIONS

In this section we study a quantitative form of a bubble region. These will play a role in both our Quantitative Bubble Tree decomposition in Section 9 and the Annulus/Bubble decomposition of Section 11. Recall from Definition 1.1 the notion of a bubble B. We begin with a definition:

Definition 4.1 (δ-Bubble Region). Given that $K < \delta$ from (1.11) with $\mathbf{b} = \{b_j\}$ a discrete set and $r : \mathbf{b} \to \mathbb{R}$ such that $r_j \leq \delta$, then $\mathcal{B} = B_{\delta^{-1}}(p) \setminus \overline{B}_{r_j}(\mathbf{b})$ is a δ-bubble region with respect to $\mathcal{L} = \mathcal{L}_B^{n-4}$ if

(b1) If $\mathbf{b}^\perp \equiv \mathbf{b} \cap \mathcal{L}_{\mathbf{p}}^\perp$ then for any $b_i \in \mathbf{b}$ we have $b_i \in \mathcal{L}_{b_j^\perp}$ with $r_i = r_j^\perp$ for some $b_j^\perp \in \mathbf{b}^\perp$.

(b2) $B_{\delta^{-1}}(p)$ and $\{B_{r_j}(b_j)\}$ are δ-weakly flat.

(b3) $r_A > \overline{r}(\Lambda, \delta)$ on $\mathcal{B}$.

(b4) $\{B_{r_j/10}(b_j)\}$ are disjoint with $r_j^{4-n} \int_{B_{r_j}(b_j)} |F|^2 > \epsilon(n, k)$ for each $b \in \mathbf{b}$.

Remark 4.1. The constant $\overline{r}(\Lambda, \delta) > 0$ is fixed according to Theorem 4.2. Indeed, with a little work one could estimate $\overline{r}$ explicitly as a polynomial of δ, but this would require many pages of tolling with little added value.

Remark 4.2. In (b4) we can take $\epsilon(n, k) = \frac{1}{2}\epsilon_{n,k}$ where $\epsilon_{n,k}$ is from the ϵ-regularity of Theorem 2.5.

One should view a bubble region in the following manner. In dimension 4, consider a δ-weakly flat ball $B_\delta(p)$. Ideally, a perfect bubble would be a smooth nontrivial YM connection such that for all $x \in B_{\delta^{-1}}(p)$ we have $r_A \geq r(\Lambda, \delta)$. This implies that in some sense the curvature is not concentrating on too small scales on this bubble, and in particular the ϵ-regularity theorem ensures that a sequence of such bubbles would automatically converge smoothly with estimates to another YM connection. However, it may happen that this is not the case. Thus, if we have some concentration of energy on scales smaller than $r(\Lambda, \delta)$, we do not want these pieces to be part of our bubble, and so we cut them out by our balls $B_{r_j}(b_j)$. By definition, each one of this pieces will carry some definite amount of scale-invariant curvature, since otherwise we need not have cut them out in the first place.

In higher dimensions, the situation is very similar. The only difference is that, instead of cutting out n-dimensional balls, we cut out tubes around $n - 4$ dimensional planes.

Indeed, one should think of a bubble as being

$$\mathcal{B} = B_{\delta^{-1}}(p) \setminus \bigcup_{b_i \in \mathbf{b}^\perp} \overline{B}_{r_i}(b_i + \mathcal{L}). \tag{4.1}$$

For future convenience, instead of writing a bubble as subtracting off a tube around the planes $\mathcal{L}_{\mathbf{b}^\perp}$ from $B_{\delta^{-1}}$, we will instead cover those planes by a Vitali collection of balls centered on them and subtract this collection off. There is no fundamental difference except it is more convenient for technical reasons later. Condition (b1) is a rephrasing of this idea.

We will prove two primary results in this section. The first is an existence theorem, which will both tell us when bubble regions exist and fix for us our constant $\bar{r}$ in our definition. The criteria for existence will help us in our construction of bubble regions in our quantitative bubble tree decomposition. The second purpose of this section is a structure theorem, the results of which are mostly a straight forward consequence of the definition combined with the knowledge of the behavior of four dimensional solutions given in Section 2.5 . This is as opposed to the corresponding structure theorem on δ-annular regions which will be introduced in the next section, which will be quite challenging.

Let us begin with our main existence theorem. The result will build for us nontrivial bubble regions in the sense that the energy of any removed ball $B_{r_i}(b_i)$ will drop by some strict amount. This will be an important aspect of future constructions:

Theorem 4.2 (Existence of Bubble Regions). *Let A be a stationary Yang-Mills connection, and assume $\int_{B_{2\delta^{-1}}} |F[\mathcal{L}]|^2 < \delta'$, $K < \delta'$ from (1.11) and that $B_{\delta^{-1}}(p)$ is δ-weakly flat wrt $\mathcal{L}$. If $\delta' < \delta'(n, k, \Lambda, \delta)$, then for $\bar{r}(\Lambda, \delta) > 0$ in Definition 4.1 there exists a δ-bubble region $\mathcal{B} = B_{\delta^{-1}}(p) \setminus \overline{B}_{r_x}(\mathbf{b})$ with best subspace $\mathcal{L}$. If further we assume $B_r(b)$ is not a δ-weakly flat ball for some $\delta^3 < r < \delta^{-1}$ and some $b \in \mathbf{b}$, then we can build the bubble so that $|\bar{\theta}_2 - \bar{\theta}_{r_b}|(b) > \epsilon(n, k)$ for all $b \in \mathbf{b}$.*

Remark 4.3. $|\bar{\theta}_1 - \bar{\theta}_{r_b}|(b) > \epsilon(n, k)$ is a nontriviality condition, which roughly says that if the bubble is not really an annular region in disguise, then there is a definite energy drop blow the bubble.

Our primary structure theorem for bubble regions is now the following.

Theorem 4.3 (Structure of Bubble Regions). *Let A be a stationary Yang-Mills connection on a δ-bubble region $\mathcal{B} = B_{\delta^{-1}} \setminus \overline{B}_{r_b}(\mathbf{b})$ with $\int_{B_{2\delta^{-1}}} |F_A|^2 \leq \Lambda$ and $\int_{B_{2\delta^{-1}}} |F[\mathcal{L}]|^2 \leq \delta'$. For each $\epsilon > 0$ if $\delta < \delta(n, k, \Lambda, \epsilon)$ and $\delta' < \delta'(n, k, \Lambda, \delta)$ then:*

(1) If $b_i \in \mathbf{b}$ with $r_i \leq r \leq \delta^{-1}$ then $\sum_{r_j \in B_r(b_i)} r_j^{n-4} < N(n, k, \Lambda) r^{n-4}$.

(2) For $q \in \mathcal{L} \cap B_1$ let $b_{q,j} = \mathcal{L}_q^\perp \cap \mathcal{L}_{b_j^\perp}$, then $\left| \theta_{\delta^{-1}}(q) - \omega_{n-4} \int_{\mathcal{B}_q} |F_A|^2 - \sum \theta_{r_j}(b_{q,j}) \right| < \epsilon$.

(3) For $q \in \mathcal{L} \cap B_1 \ \exists \ c_i \in \mathcal{B}_q \cap B_1$ and $s_i > 0$ with $\#\{c_i\} \leq N(n, k, \Lambda)$ and $s_i^{4-n} \int_{B_{s_i}} |F|^2 > \epsilon(k)$ such that if $R \geq R(n, k, \Lambda, \epsilon)$ then $\left| \int_{\mathcal{B}_q \cap \bigcup B_{Rs_i}(c_i)} |F_A|^2 - \int_{\mathcal{B}_q} |F_A|^2 \right| < \epsilon$.

(4) We have the estimate $\delta^{n-4} \int_{\mathcal{B}} |\nabla^2 F_A| < C(n, k, \Lambda, \delta)$.

Remark 4.4. In fact, it will be a consequence of Theorem 1.2 that we have the better estimate $\delta^{n-4} \int_{\mathcal{B}} |\nabla^2 F_A| < C(n, k, \Lambda)$, without the δ-dependence. However, to prove Theorem 1.2 it will be enough to first prove this weaker result.

Remark 4.5. (2) is the energy identity for bubbles and tells us that we can compute the energy at a point in terms of the energy of a slice.

Remark 4.6. (3) is a concentration condition and tells us that the energy of the bubble can separate into at most N clumps, independent of δ.

4.1. **Existence of Bubble Regions.** In this subsection we prove Theorem 4.2. Thus let us fix $\delta > 0$ and assume $\fint_{B_{2\delta^{-1}}} |F[\mathcal{L}]|^2 < \delta'$, which will be fixed later. For $0 < \delta'' < \delta$ fixed observe that $\forall\ x \in B_{\delta^{-1}}$ there exists a radius $r_x > 0$ with $\bar{r}(\delta'', \Lambda) < r_x < \delta^5$ such that $|\theta_{2r_x} - \theta_{\delta'' r_x}|(x) < \delta''$. Indeed, to see this consider the sequence of radii $s_a \equiv (\delta'')^a \delta^5$ and note for all $N \in \mathbb{N}$ that

$$\sum_{1}^{N} |\theta_{s_a} - \theta_{s_{a+1}}|(x) \le \Lambda. \tag{4.2}$$

Therefore for $N = \Lambda(\delta'')^{-1} + 1$ we see that for one of the radii $s_0, \ldots s_N$ we must have $|\theta_{s_a} - \theta_{s_{a+1}}|(x) < \delta''$ as claimed, otherwise by monotonocity of θ we contradict the above sum. Given this we can choose $\delta' < \delta'(\Lambda, \delta'')$ so that we can also be assured $r_x^{4-n} \fint_{B_{2r_x}(x)} |F[\mathcal{L}]|^2 < \delta''$, so that $B_{2r_x}(x)$ is $(n-4, \delta'')$-symmetric. We now pick $\delta'' \le \delta''(n, k, \Lambda, \delta)$ so that Theorem 3.2 holds with δ.

To finish the construction of $\mathcal{B}$ let us consider for $x \in \mathcal{L}_p^{\perp} \cap B_1$ the covering $\{B_{r_x}(x)\}$ of $\mathcal{L}_p^{\perp} \cap B_1$. Let $\{B_{r_j^{\perp}}(x_j^{\perp})\}$ be a Vitali subcovering with $x_j^{\perp} \in \mathcal{L}^{\perp} \cap B_1$ such that $r_j^{\perp} = r_{x_j}$ and with $\{B_{r_j^{\perp}/10}(x_j^{\perp})\}$ disjoint. By translating the centers balls over $\mathcal{L}$ and picking another Vitali subcovering we can extend this to a covering $\{B_{r_j'}(x_j')\}$ of B_1 such that for each x_j' there exists $x_i^{\perp}$ with $x_j' \in \mathcal{L}_{x_i^{\perp}}$ and $r_j' = r_i^{\perp}$. Now by using the uniform lower bound on the radii we can apply Theorem 3.2 to each ball $B_{r_j^{\perp}}(x_j^{\perp})$ in order to conclude that either $r_A(x_j^{\perp}) > \frac{1}{2} r_j^{\perp} \ge \bar{r}(\Lambda, \delta)$ on $B_{r_j^{\perp}}(x_j^{\perp})$ or that $B_{r_j^{\perp}}(x_j^{\perp})$ is $\delta/2$-weakly flat with $\theta_{r_j^{\perp}}(x_j^{\perp}) > \epsilon_{n,k}$. Let us now define $\mathbf{b}^{\perp} = \{\mathbf{b_j^{\perp}}\} \subseteq \{\mathbf{x_j'}\}$ to be the collection of balls for which this second condition holds, and let $\mathbf{b} \subseteq \{x_j'\}$ be the subset for which $b_j \in \mathcal{L}_{b_i^{\perp}}$ for some $b_i^{\perp} \in \mathbf{b}^{\perp}$. We then define $\mathcal{B} \equiv B_{\delta^{-1}} \setminus \bigcup B_{r_j}(b_j)$ as expected. Note that on $\mathcal{B}$ we have the regularity estimate $r_A \ge \frac{1}{2} \min r_j' \ge \bar{r}(\Lambda, \delta)$. Therefore with $\delta' < \delta'(n, k, \Lambda, \delta)$ we have that $\mathcal{B}$ is indeed a δ-bubble region.

Let us now focus on proving $|\theta_2 - \theta_{r_j}|(b_j) > \epsilon(n, k)$ under the assumption that for some $b \in \mathbf{b}$ and $\delta^4 < r < 1$ we have that $B_r(b)$ is not δ-weakly flat, at least if we choose $\delta' < \delta'(n, k, \Lambda, \delta)$ sufficiently small. So assume this is not the case for any δ', then we can find a sequence of connections A_i with bubble regions $\mathcal{B}_i \equiv B_1 \setminus B_{r_{i,x}}(\mathbf{b_i})$ as above with $\fint_{B_1} |F[\mathcal{L}_i]|^2 \to 0$. After passing to a subsequence we can limit $A_i \to A$ with $|F_{A_i}|^2 dv_{g_i} \to |F_A|^2 dv_g + \nu$ and $\mathcal{B}_i \to \mathcal{B} = B_1 \setminus B_{r_x}(\mathbf{b})$, and by using Theorem 2.3 we have that A defines a smooth Yang Mills connection on $\mathcal{L}^{\perp}$ and that ν is invariant under translation by $\mathcal{L}$. By our contradicting assumption we have that

$$\left| (|F_A|^2 dv_g + \nu)[B_2(b)] - (|F_A|^2 dv_g + \nu)[B_{r_b}(b)] \right| \le \epsilon(n, k). \tag{4.3}$$

By choosing $\epsilon(n, k)$ sufficiently small, from the ϵ-regularity of Theorem 2.5, we then have that $\nu \equiv 0$ on $A_{r_b,2}(b)$ with $d_x^2 |F_A| < \epsilon'(n, k)$ on $A_{r_b,2}(b)$, where $d_x = d(x, b)$ and $\epsilon'(n, k)$ is from Theorem 2.8. Indeed, since $B_{\delta^{-1}}(p)$ is δ-weakly flat we even have that $d_x^2 |F_A| < \epsilon'(n, k)$ on $A_{r_b,\delta^{-1}}(b)$. Now let us apply Theorem 6.2, and using that $B_{\delta^{-1}}(b)$ and $B_{r_b}(b)$ are δ-weakly flat, we then get the improved estimate for $x \in A_{r_b,\delta^{-1}}$:

$$d_x^2 |F_A|(x) < C(k)\delta\left((\delta\, d_x)^\alpha + \left(\frac{r_b}{d_x}\right)^\alpha\right), \tag{4.4}$$

Now recall our assumption is that for some $\delta^3 < r < \delta^{-1}$ we have that $B_r(b)$ is not δ-weakly flat. In particular, this implies in our case that for some $x \in A_{\delta^4,10}(b)$ we must have that $d_x^2 |F_A|(x) > \delta$. However, for $\delta < \delta(k)$ and $r_b < \delta^5$ as constructed we see from the above estimate that this is not possible, and thus we have found our desired contradiction and proved the Theorem. □

4.2. Structure of Bubble Regions.

In this subsection we give a proof of Theorem 4.3. The proof of the content estimate (1) follows immediately from $(b2)$ and $(b4)$ in the definition of a δ-bubble region. Indeed, using the lower bound $r_j^{n-4} \int_{B_{r_j}(x_j)} |F_A|^2 > \epsilon(n, k)$ together with the fact that $B_{r_j}(x_j)$ are δ-weakly flat we immediately have lower bounds on the slightly smaller balls

$$r_j^{n-4} \int_{B_{r_j/10}(x_j)} |F_A|^2 > \frac{1}{2}\epsilon(n, k). \tag{4.5}$$

Combining this with the disjoint property of $\{B_{r_j/10}(x_j)\}$ we have

$$\sum r_j^{n-4} \le \sum 2\epsilon(n, k)^{-1} \int_{B_{r_j/10}} |F_A|^2 \le C(n, k) \int_{B_2} |F_A|^2 \le C(n, k)\Lambda, \tag{4.6}$$

which proves the content estimate.

Let us now focus on the hessian estimate of (4). Indeed, for this we use $(b3)$ to see that $r_A > \bar{r}(\Lambda, \delta)$ on $\mathcal{B} \subseteq B_{\delta^{-1}}$. Standard elliptic estimates then give us that $|\nabla^2 F_A| < C(n, k, \Lambda, \delta)$ pointwise on $\mathcal{B}$, which in particular implies the weaker L^1 estimate

$$\int_{\mathcal{B}} |\nabla^2 F_A| \le C(n, k, \Lambda, \delta)\mathrm{Vol}(B_{\delta^{-1}}), \tag{4.7}$$

as claimed.

We are now left with proving that (2) and (3) hold for δ' sufficiently small. So assume this is not the case for any δ', then we can find a sequence of connections A_i with bubble regions $\mathcal{B}_i \equiv B_1 \setminus B_{r_{i,x}}(\mathbf{b_i})$ as above with $\fint_{B_1} |F[\mathcal{L}_i]|^2 \to 0$. After passing to a subsequence we can limit $A_i \to A$ with $|F_{A_i}|^2 dv_{g_i} \to |F_A|^2 dv_g + \nu$ and $\mathcal{B}_i \to \mathcal{B} = B_1 \setminus B_{r_x}(\mathbf{b})$, and by using Theorem 2.3 we have that A defines a smooth Yang Mills connection on $\mathcal{L}^\perp$ and that ν is invariant under translation by $\mathcal{L}$.

Let us first focus on (2), and it is enough to prove this for $q = p$ as the other cases are verbatim. Indeed, viewing A as a connection on the four dimensional space $\mathcal{L}^\perp$ we have the equality

$$\nu[B_{\delta^{-1}}] + \int_{B_{\delta^{-1}}(0^\perp)} |F_A|^2 = \sum\left(\nu[B_{r_j^\perp/10}(b_j^\perp)] + \int_{B_{r_j^\perp}(b_j^\perp)} |F_A|^2\right) + \int_{\mathcal{B}\cap\mathcal{L}^\perp} |F_A|^2. \tag{4.8}$$

Using that $\theta_{\delta^{-1}}(p_i) \to \nu[B_{\delta^{-1}}] + \int_{B_{\delta^{-1}}} |F_A|^2$ with $\theta_{r_{i,j}}(b_{i,j}^\perp) \to \nu[B_{r_j/10}] + \int_{B_{r_j}(b_j^\perp)} |F_A|^2$ and $\omega_{n-4} \int_{\mathcal{B}_{p_i,i}} |F_{A_i}|^2 \to \int_{\mathcal{B}\cap\mathcal{L}^\perp} |F_A|^2$ this shows (2) must hold for sufficiently far in the sequence. Thus we can concentrate on (3).

To prove (3) let us apply Theorem 2.9 to the limit A on $\mathcal{L}^\perp \cap B_4$ to get points $c_1^A, \ldots, c_N^A \in \mathcal{L}^\perp \cap B_1$. By construction we have for $R \geq R(n, k, \Lambda, \epsilon)$ that

$$\int_{B_2 \cap B_{Rs_j}(c_j)} |F_A|^2 < \frac{\epsilon}{4} . \tag{4.9}$$

Since $B_{\delta^{-1}}(p)$ is δ-weakly flat we may apply Theorem 2.8 in order to see that $\int_{A_{1,\delta^{-1}}(p)} |F_A|^2 < C(k)\delta < \frac{\epsilon}{4}$, and hence

$$\int_{B_{\delta^{-1}} \cap B_{Rs_j}(c_j)} |F_A|^2 < \frac{\epsilon}{2} . \tag{4.10}$$

Finally, observing that $\text{supp}[\nu] \subseteq \bigcup_{\mathbf{b}} B_{r_b}(b)$ this gives us for sufficiently far in the sequence that

$$\int_{\mathcal{B}_p \cap B_{\delta^{-1}} \cap B_{Rs_j}(c_j)} |F_{A_i}|^2 < \epsilon , \tag{4.11}$$

which shows that (3) holds and thus finishes the proof. $\square$

5. δ-ANNULAR REGIONS

In this section we consider the second piece of our quantitative decompositions, namely the δ-annular regions. These are the regions which will turn out to be the most challenging to analyze, and the next several sections of this paper will be dedicated to proving the results stated in this section.

Quantitative bubble regions have the property that they have large curvature, but only on bounded domains. On the other hand, quantitative annular regions have small curvature, but over potentially an infinite number of scales. Recalling the notion of a weakly flat region in Definition 3.1 we define an annular region in the following manner:

Definition 5.1. If $K_M < \delta$ from (1.11), then we call $\mathcal{A} \subseteq B_2(p)$ a δ-annular region if there exists a closed subset $\mathcal{C} = \mathcal{C}_0 \cup \mathcal{C}_+ = \mathcal{C}_0 \cup \{x_i\}$, a radius function $r : \mathcal{C} \to \mathbb{R}^+$ with $0 < r_x \leq \delta$ on $\mathcal{C}_+$ and $r_x = 0$ on $\mathcal{C}_0$, and a $n - 4$ subspace $\mathcal{L}$ such that $\mathcal{A} \equiv B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ satisfies

   (a1) $\{B_{\tau^2 r_x}(x)\}$ are pairwise disjoint.
   (a2) For each $x \in \mathcal{C}$ and $r_x \leq r \leq 2$ we have that $B_r(x)$ is δ-weakly flat wrt $\mathcal{L}_x \equiv \mathcal{L} + x$.
   (a3) For each $x \in \mathcal{C}$ and $r_x \leq r$ with $B_{2r}(x) \subseteq B_2$ we have that $\mathcal{L}_x \cap B_r \subseteq B_{\tau r}(\mathcal{C})$ and $\mathcal{C} \cap B_r \subseteq B_{\delta r}(\mathcal{L}_x)$.
   (a4) $|\text{Lip } r_x| \leq \delta$.

For each $\tau \leq s \leq 1$ we define the regions $\mathcal{A}_s \equiv B_2 \setminus \overline{B}_{s \cdot r_x}(\mathcal{C})$ as well as the wedge regions $\mathcal{W}^\theta(x) \equiv \{y \in A_{r_x/2,2}(x) : d(y, \mathcal{L}_x) \geq \cos(\theta)\, d(y, x)\}$ and $\mathcal{W}_r^\theta(x) \equiv \mathcal{W}^\theta(x) \cap A_{\cos \theta\, r, r/\cos \theta}(x)$.

Remark 5.1. The constant $\tau = \tau_n = 10^{-10n}\omega_n$ is a dimensionally chosen constant designed to neutralize any errors obtained overlaps in covering constructions.

Remark 5.2. For a smooth connection we have $\mathcal{C}_0 = \emptyset$. For nonsmooth stationary connections one must allow for the possibility that $\mathcal{C}_0 \neq \emptyset$.

Remark 5.3. Note that ($a1$) and ($a4$) imply that $B_{10r_x}(x)$ intersects at most $C(n)$ other balls in the covering, all of which have radii which are in the range $[\frac{1}{2}r_x, 2r_x]$.

Associated to each annular region is its packing measure. In the same way in which $\mathcal{C}$ approximates the $n-4$ defect measure of the connection, we have that the packing measure approximates the $n-4$ Hausdorff measure on the support of this defect measure. Precisely we have the following:

Definition 5.2. Let $\mathcal{A} \equiv B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ be an annular region, then we define the associated packing measure

$$\mu = \mu_{\mathcal{A}} \equiv \sum_{x \in \mathcal{C}_+} r_x^{n-4} \delta_x + \lambda^{n-4}|_{\mathcal{C}_0}, \tag{5.1}$$

where $\lambda^{n-4}|_{\mathcal{C}_0}$ is the $n-4$-dimensional Hausdorff measure restricted to $\mathcal{C}_0$.

The goal of this section is two fold. We will first discuss some properties of annular regions. Our three main properties about annular regions will be to show that the packing measure is Ahlfor's regular, and that in the annular region we have apriori L^1 hessian and L^2 curvature bounds. These estimates will be the eventual key to both the global L^1 hessian estimate of Theorem 1.2 and the energy identity of Theorem 1.3. We will prove the Ahlfor's regularity statement in this section, however the curvatures estimates will not be proved until later in the paper, as there is a lot of new technical constructions needed in their proofs.

Our second main goal will be to prove the existence of annular regions. In order for an annular region to be useful in the end analysis, we will need to know many exist. In this section we will give some basic criteria used to construct *maximal* annular regions. The maximal property of the constructed annular regions will be crucial in the proof of the annulus/bubble decomposition in Section 11.

Let us now begin by stating our main structural result on the properties of annular regions:

Theorem 5.3 (Structure of Annular Regions)**.** *Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ satisfying (1.11) and $\fint_{B_4} |F_A|^2 \leq \Lambda$. For each $\epsilon > 0$ if $\delta < \delta(n, \Lambda, k, \epsilon)$ we then have:*

  (1) *For each $x \in \mathcal{C}$ and $r_x < r < 4$ we have that $A(n)^{-1} r^{n-4} \leq \mu(B_r(x)) \leq A(n) r^{n-4}$.*
  (2) *We have the estimate $\int_{\mathcal{A} \cap B_1} |\nabla^2 F| < \epsilon$.*
  (3) *We have the estimate $\int_{\mathcal{A} \cap B_1} |F_A|^2 < \epsilon$.*

With this in hand let us see what criteria may be used to build annular regions. A key result in the annulus/bubble decomposition of Theorem 11.1 is the $n-4$ content bound on the number of pieces to the decomposition. It is worth noting that for such an estimate to hold, one must be quite careful about the construction of the annular regions. Indeed, if one were to build annular regions which were much smaller than they need to be, it is possible the content estimate would fail. Therefore, we must also analyze what it means to build *maximal* annular regions, a concept which will be made precise in the following Theorem:

Theorem 5.4 (Existence of Annular Regions)**.** *Let A be a stationary Yang-Mills connection with $\fint_{B_4} |F_A|^2 \leq \Lambda$. For each $\epsilon > 0$ and $0 < \delta < \delta(n, k, \Lambda, \epsilon)$ there exists $\delta'(n, k, K, \Lambda, \delta)$ such that if*

(1) $\int_{B_8} |F[\mathcal{L}]|^2 < \delta'$,

(2) *For each $4^{-1}\delta \le r \le 4$ we have that $B_r(p)$ is δ-weakly flat wrt $\mathcal{L}$,*

then $\exists$ a δ-annular region $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ wrt $\mathcal{L}$ such that if we consider the set

$$\mathcal{C}^c \equiv \begin{cases} x \in \mathcal{C} : & B_r(x) \text{ is } \delta\text{-weakly flat for } \delta^4 r_x \le r \le r_x, \text{ or} \\ x \in \mathcal{C} : & \exists q \in \mathcal{L} \text{ with } \mathcal{L}_q^\perp \cap \bar{B}_{r_x}(x) \ne \emptyset \text{ and } \int_{\mathcal{L}_q^\perp \cap \mathcal{A}} |F_A|^2 > \epsilon, \end{cases} \tag{5.2}$$

then we have the estimate $\mu(\mathcal{C}^c \cap B_1) < \epsilon$.

5.1. **Proof of Ahlfor's Regularity of Annular Region.**

In this subsection we prove the Ahlfor's regularity of Theorem 5.3.1 . Thus throughout this section $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(x)$ is a δ-annular region and μ is its packing measure. To prove the result let us begin with the following:

Claim: The projection mapping $\pi : \mathcal{C} \to \mathcal{L}$, from the center points to the annular best plane $\mathcal{L}$, is a $1 + \epsilon_n$-bilipschitz map where $\epsilon_n < 100^{-1}$. That is, for $x, y \in \mathcal{C}$ we have that $(1 - \epsilon_n)|\pi(x) - \pi(y)| \le d(x, y) \le (1 + \epsilon_n)|\pi(x) - \pi(y)|$.

To prove the claim let us consider $x, y \in \mathcal{C}$ and let $r \equiv d(x, y)$. Then by condition $(a3)$ we have for $\mathcal{L}_x \equiv \mathcal{L} + x$ that $\mathcal{L}_x \cap \overline{B}_r \subseteq B_{\tau_n r}(\mathcal{C})$ and $\mathcal{C} \cap \overline{B}_r \subseteq B_{\tau_n r}(\mathcal{L})$. In particular, this gives us that

$$|\pi(x) - \pi(y)| - \tau_n r \le d(x, y) \le |\pi(x) - \pi(y)| + \tau_n r . \tag{5.3}$$

However, we have chosen our scale so that $r \equiv d(x, y)$, and therefore by rearranging we have

$$\frac{1}{1 + \tau_n}|\pi(x) - \pi(y)| \le d(x, y) \le \frac{1}{1 - \tau_n}|\pi(x) - \pi(y)| , \tag{5.4}$$

which finishes the proof of the claim. □

Now to finish the proof let us pick $x \in \mathcal{C}$ with $r_x \le r < 1$ such that $B_{2r}(x) \subseteq B_2$. Let us first prove the upper bound on $\mu(B_r(x))$. First note that by the bilipschitz condition we have that

$$\pi(\mathcal{C} \cap B_r(x)) \subseteq B_{2r}(\pi(x)) . \tag{5.5}$$

Further, since the collection of balls $\{B_{\tau^2 r_y}(y)\}$ are all disjoint, if we again use the bilipschitz condition we must have that the image balls $\{B_{\tau^3 r_y}(\pi(y))\}$ are disjoint in $\mathcal{L}$. But then this give us

$$\mu(B_r(x)) \equiv \sum_{y \in \mathcal{C}_+ \cap B_r(x)} r_y^{n-4} + \lambda^{n-4}(\mathcal{C}_0 \cap B_r(x)) \le \tau_n^{-3(n-4)}\Big(\sum_{y \in \mathcal{C}_+ \cap B_r(x)} (\tau_n^3 r_y)^{n-4} \Big) + \lambda^{n-4}(\mathcal{C}_0 \cap B_r(x))$$

$$\le C(n)\Big(\sum_{x \in \mathcal{C}_+ \cap B_r} \text{Vol}(B_{\tau^3 r_y}(\pi(y))) + Vol(\pi(\mathcal{C}_0 \cap B_r)) \Big) \le C(n)\text{Vol}(B_{2r}(\pi(x))) \le A(n)r^{n-4} , \tag{5.6}$$

where we have used the bilipschitz condition on π multiple times and that $\pi(B_r(x)) \subseteq B_{2r}(\pi(x))$.

In order to prove the lower bound we start with the following claim:

Claim: For each $x \in \mathcal{C}$ and $r_x \leq r < 1$ with $B_{2r}(x) \subseteq B_2$ we have that $B_{r_y}(\pi(\mathcal{C} \cap B_r(x))) = \bigcup_{y \in \pi(\mathcal{C} \cap B_r(x))} B_{r_y}(y) \supseteq B_{r/2}(\pi(x))$.

To prove the result let us assume it is false and let $y \in B_{r/2}(\pi(x)) \setminus B_{r_y}(\pi(\mathcal{C} \cap B_r(x)))$. Let us choose $z \in \pi(\mathcal{C} \cap B_r(x))$ such that

$$z \in \operatorname{argmin}\{d(z,y) : z \in \pi(\mathcal{C} \cap B_r(x))\}. \tag{5.7}$$

Let $s \equiv d(y,z)$, and note that by our assumption on y we have that $s \geq r_z$. But if we apply condition $(a3)$ to $B_{2s}(z)$, then we know that $\pi(\mathcal{C} \cap B_r(x))$ is $2\tau_n s$ dense in $B_{2s}(\pi(z)) \subseteq \mathcal{L}$, which contradicts that z is the closest point of $\pi(\mathcal{C})$ to y and satisfies $d(y,z) = s$. This proves the claim. □

With the claim in hand we can now easily prove the lower volume bound. Indeed, we have:

$$\mu(B_r(x)) = \sum_{y \in \mathcal{C} \cap B_r} r_y^{n-4} \geq C(n)^{-1} \sum_{y : \pi(y) \in B_{r/2}(\pi(x))} \operatorname{Vol}(B_{r_y}(\pi(y))) \geq C(n)^{-1} \operatorname{Vol}(B_{r/2}(\pi(x))) \geq A(n)^{-1} r^{n-4}, \tag{5.8}$$

which finishes the proof of the lower bound. □

5.2. Existence of Annular Regions.

In this section we deal with the issue of proving the existence of a δ-annular region which is maximal in a suitable sense. Let us pick some $\delta'' > 0$ which will be fixed later, and now we begin by defining the following for each $x \in B_2$:

$$s_x \equiv \inf_{s \leq 1} \left\{ \forall s \leq r \leq 1 \ r^{4-n} \int_{B_r(x)} |F[\mathcal{L}]|^2 < \delta'' \right\},$$

$$s_x' \equiv \inf_{s \leq 1} \left\{ \forall \delta^3 s \leq r \leq 1 \ B_r(x) \text{ is } \delta\text{-weakly flat} \right\},$$

$$t_x \equiv \max\{s_x, s_x'\}. \tag{5.9}$$

Now for each $q \in \mathcal{L} \cap B_1$ let $y_q \in \mathcal{L}_q^{\perp}$ and $t_q \geq 0$ be defined by

$$t_q \equiv \min_{y \in \mathcal{L}_q^{\perp}} t_y,$$

$$y_q \in \arg\min_{y \in \mathcal{L}_q^{\perp}} t_y. \tag{5.10}$$

Now let us define $\{B_{t_i}(y_i)\}$ be a maximal subset of $\{B_{t_q}(y_q)\}$ such that $\{B_{10^{-1}t_i}(y_i)\}$ are disjoint. Let us decompose this collection into two subcollections:

$$\{B_{t_i}(y_i)\} = \{B_{t_i'}(y_i')\} \cup \{B_{\tilde{t}_i}(\tilde{y}_i)\}, \tag{5.11}$$

where $\{B_{t_i'}(y_i')\}$ is the subcollection such that

$$t_i' = \max\{s_{y_i'}, s_{y_i'}'\} = s_{y_i'}' \geq 10 \sup_{y \in B_{t_i'}(y_i')} s_y, \tag{5.12}$$

and $\{B_{\tilde{t}_i}(\tilde{y}_i)\}$ are the remaining balls. Notice that if $\delta' < \delta'(n, \delta, \delta'')$ then by a standard maximal function argument we have the estimate

$$\sum \tilde{t}_i^{n-4} \leq \delta^2. \tag{5.13}$$

Let us now consider the following claim:

Claim: If $\delta'' < \delta''(n, k, \Lambda, \delta)$ and $y \in B_{1000t_i'}(y_i')$, then $c(k)t_i' < s_y' \le t_y$. Further, if $y \in \{y_j\}$ is a ball center then we have the two sided estimate $c(k)t_i' < t_y < 10^4 t_i'$.

The upper bound $t_y < 10^4 t_i'$ in the case y is a center point follows from the Vitali condition, therefore we will focus on the lower bound $c(k)t_i' < s_y'$ for general y, which will itself follow from Theorem 3.3. Note by the definition of t_i' we have that $B_{\delta t_i'}(y_i')$ is δ-weakly flat, but $B_r(y_i')$ is not δ-weakly flat for some $\frac{1}{2}\delta^3 t_i' < r < \delta^3 t_i'$.

Now let us look at the ball $B_{t_i'}(y)$. If we assume $s_y' \le c(k)t_i'$, then we have for all $c(k)\delta^3 t_i' \le r \le c(k)^{-1}\delta^3 t_i'$ that $B_r(y)$ is δ-weakly flat. Using Theorem 3.3 with $\kappa = 10^{-1}$ we have for $c(k)$ sufficiently small that $B_r(y)$ is $10^{-1}\delta$-weakly flat for all $10^{-1}\delta^3 t_i' \le r \le 10\delta^3 t_i'$. In particular, for $\delta'' \le \delta''(n, k, \Lambda, \delta)$ we have that there must exist a point $z \in \mathcal{L}_{y_i'}^\perp \cap B_{t_i'}$ such that $s_z' \le 2^{-1}t_i'$. However, by the definition of $B_{t_i'}(y_i')$ we also have that $s_z \le 10^{-1}t_i'$, and therefore $t_z \le 2^{-1}t_i'$. This contradicts that $t_{y_i'} = \min_{u \in \mathcal{L}_{y_i'}^\perp} t_u$, and thus proves the Claim. $\square$

To continue with the proof of Theorem 5.4 let us define the radii function

$$r_x \equiv \begin{cases} \delta^2 t_i & \text{if } x \in B_{\delta t_i}(y_i), \\ \delta d(x, \{y_i\}) & \text{if } x \notin \bigcup B_{\delta t_i}(y_i). \end{cases}$$

Note that $|\text{Lip } r_x| < \delta$. Now we play a similar game as in the previous covering and define for each $q \in \mathcal{L} \cap B_1$ the point $x_q \in \mathcal{L}_q^\perp$ and $r_q \ge 0$ by

$$r_q \equiv \min_{x \in \mathcal{L}_q^\perp} r_x,$$

$$x_q \in \arg\min_{x \in \mathcal{L}_q^\perp} r_x. \tag{5.14}$$

Now for some x_q let y_i be the center point such that either $x_q \in B_{\delta t_i}(y_i)$ or $r_q = \delta d(x, y_i)$, so that in either case we have that $r_q \ge \delta^2 t_i$. Since $B_r(y_i)$ is δ-weakly flat for $\delta^3 t_i \le r \le c(k)^{-1}$ we have by Theorem 3.3 the improved estimate that $B_r(y_i)$ is $10^{-1}\delta$-weakly flat for $c(k)\delta t_i \le r \le 1$. In particular, we have that $B_r(y_i)$ is $10^{-1}\delta$-weakly flat for $r_q \le r \le 1$. If $\delta'' \le \delta''(n, k, \Lambda, \delta)$ is sufficiently small we then conclude that $B_r(x_q)$ is itself δ-weakly flat for $r_q \le r \le 1$.

We now define our annular region by $\mathcal{C} \subseteq \{x_q\}$ so that $\{B_{r_x}(x)\}_{x \in \mathcal{C}}$ is a maximal subset of $\{B_{r_q}(x_q)\}_{q \in \mathcal{L}}$ such that $\{B_{\tau^2 r_x}(r_x)\}$ are disjoint. It is a straightforward, if somewhat tedious, exercise using the remarks of the previous paragraph to check for $\delta'' \le \delta''(n, k, \Lambda, \delta)$ that this defines a δ-annular region. We will focus then on the nontriviality of this annular region. That is, if we consider the sets

$$\mathcal{C}^{c,1} \equiv \{x \in \mathcal{C} : B_r(x) \text{ is } \delta\text{-weakly flat for } \delta^4 r_x \le r \le r_x\},$$

$$\mathcal{C}^{c,2} \equiv \{x \in \mathcal{C} : \exists q \in \mathcal{L} \text{ with } \mathcal{L}_q^\perp \cap \bar{B}_{r_x}(x) \ne \emptyset \text{ and } \int_{\mathcal{L}_q^\perp} |F_A|^2 > \epsilon\}, \tag{5.15}$$

then we want to see for $\delta' \le \delta'(n, k, \Lambda, \delta)$ that we have the estimates $\mu(\mathcal{C}^{c,1}), \mu(\mathcal{C}^{c,2}) < \frac{1}{2}\epsilon$. We begin by estimating $\mathcal{C}^{c,1}$. To accomplish this let us consider $x \in \mathcal{C} \cap B_{t_i'}(y_i')$, then by using the two sided bound of the

previous Claim, and particular that $c(k)t_i' \le t_j \le c(k)^{-1}t_i'$ for any other ball center in $B_{2t_i'}(y_i')$, we have the estimate

$$c(k)\delta\, t_i' \le r_x \le c(k)^{-1}\delta\, t_i'\,.$$

Additionally, we know by the Claim that $s_x \ge c(k)t_i'$, which is to say for some radius $r \ge \frac{1}{2}c(k)\delta^3 t_i'$ we must have that $B_r(x)$ is not δ-weakly flat. Combining this with the above estimate we see that for some radius $r \ge c(k)\delta^2 r_x$ that $B_r(x)$ is not δ-weakly flat, which for $\delta < \delta(k)$ implies that $x \notin \mathcal{C}^{c,1}$, and in particular gives us the inclusion

$$\mathcal{C}^{c,1} \subseteq \bigcup B_{\tilde{t}_j}(\tilde{y}_j)\,. \tag{5.16}$$

Finally, using (5.13) and the Ahlfor's regularity on μ proved in Section 5.1 we have the estimate

$$\mu(\mathcal{C}^{c,1}) \le \sum \mu(B_{\tilde{t}_j}(\tilde{y}_j)) \le C(n) \sum \tilde{t}_j^{n-4} \le C(n)\delta^2 < \delta\,. \tag{5.17}$$

Now we focus on estimating $\mathcal{C}^{c,2}$, which will itself depend on the curvature estimate of Theorem 5.3.3. So let us choose $\delta < \delta(n, k, \Lambda, \epsilon')$ such that Theorem 5.3 holds with $\epsilon' > 0$. Note then that we have

$$\int_{B_1(0^{\mathcal{L}})} \int_{\mathcal{L}_{\tilde{q}}^{\perp} \cap \mathcal{A}} |F_A|^2 = \int_{\mathcal{A} \cap B_1} |F_A|^2 < \epsilon'\,. \tag{5.18}$$

Using the first claim of subsection 5.1, where it is shown that the projection $\pi_{\mathcal{L}}$ is uniformly bilipschitz on $\mathcal{C}$, this implies that

$$\int_{B_1} \Big(\fint_{B_{r_x}(\pi(x))} \int_{\mathcal{L}_{\tilde{q}}^{\perp} \cap \mathcal{A}} |F_A|^2 \Big) d\mu[x] < C(n)\epsilon'\,. \tag{5.19}$$

In particular, if $\epsilon' \le c(n)\epsilon^2$ then away from a set $\mathcal{C}' \subseteq \mathcal{C} \cap B_1$ with $\mu(\mathcal{C}') < \frac{1}{2}\epsilon$ we have that $\fint_{B_{r_x}(\pi(x))} \int_{\mathcal{L}_{\tilde{q}}^{\perp} \cap B_1} |F_A|^2 < 10^{-1}\epsilon$. Now let us use that $B_r(x)$ is δ-weakly flat for each $r > r_x$ combined with standard elliptic estimates to see that $|\nabla F_A| < C(n)\delta r^{-3}$ on $A_{r,r/2}(\mathcal{L}_x)$. This in particular gives us for each $x \in \mathcal{C} \cap B_1$ that

$$\Big| \max_{q \in B_{r_x}(\pi(x))} \int_{\mathcal{L}_{\tilde{q}}^{\perp} \cap \mathcal{A}} |F_A|^2 - \fint_{B_{r_x}(\pi(x))} \int_{\mathcal{L}_{\tilde{q}}^{\perp} \cap \mathcal{A}} |F_A|^2 \Big| < C(n)\delta\, r_x\,. \tag{5.20}$$

For $x \notin \mathcal{C}'$ this gives us that

$$\max_{q \in B_{r_x}(\pi(x))} \int_{\mathcal{L}_{\tilde{q}}^{\perp} \cap \mathcal{A}} |F_A|^2 \le \fint_{B_{r_x}(\pi(x))} \int_{\mathcal{L}_{\tilde{q}}^{\perp} \cap \mathcal{A}} |F_A|^2 + C(n)\delta\, r_x < 10^{-1}\epsilon + C(n)\delta^2 < \epsilon\,. \tag{5.21}$$

Therefore we have that $\mathcal{C}^{c,2} \subseteq \mathcal{C}'$ and thus $\mu(\mathcal{C}^{c,2}) < \frac{1}{2}\epsilon$, which finishes the proof of the Theorem. $\square$

6. HARMONIC ϵ-GAUGE

Standard in any gauge problem is the need to choose a good coordinate system in order to study the equations. In the context of Yang Mills the standard gauge condition one prefers is the Coulomb gauge. Unfortunately, such a gauge will only exist locally and in general only when the underlying geometry is quite simple (e.g. when the curvature of the connection is small).

A key object of study in this paper are annular regions. Annular regions $\mathcal{A} \subseteq B_2$ are regions for which the connection looks very flat in a weak sense, but there is some curvature concentration on B_2 which is only

visible on small scales inside the singular balls. A Coulomb gauge will in general certainly not exist on the whole ball.

Instead in this section we will introduce a form of linearized Coulomb gauge associated to the induced vector bundle $E \to M$ coming from the orthogonal representation of $G \subseteq SO(k)$. This gauge will exist and solve an equation on the whole ball B_2, a point which will be important and useful in the analysis. On the other hand, this linearized gauge will only form a a legitimate (vector bundle) gauge on part of the ball. Recall that if A is a Yang-Mills connection on P then E is equipped with a metric connection ∇_A. In particular, we have the associated Laplace operator $\Delta_A : \Gamma(E) \to \Gamma(E)$. Let us begin by defining our notion of an ϵ-gauge on E:

Definition 6.1. We say that sections $V^1, \ldots, V^k \in \Gamma(B_r, E)$ form a harmonic ϵ-gauge on $B_r(x)$ if the following hold:

(1) $\Delta V^a = 0$.
(2) $|V^a| \leq 1 + \epsilon$.
(3) $\fint_{B_r} |\langle V^a, V^b \rangle - \delta^{ab}| < \epsilon$.
(4) $r^2 \fint_{B_r} |\nabla V|^2 < \epsilon^2$.

The goal of this section is to prove the existence of ϵ-gauges on balls B_2 which admit annular regions $\mathcal{A} \subseteq B_2$, and to prove that V^a forms an actual vector bundle gauge over the whole annular region $\mathcal{A}$. The difficulty of this second statement is that apriori the sections V^a, which have bounded norm, may have norm tending to zero near the singular balls or may be becoming linearly dependent. While one cannot say this doesn't happen, in fact in the whole ball it must happen, we will see that it cannot happen faster than at a small polynomial rate, and thus the sections remain a basis. In fact, we will show something much stronger, we will see that for every $x \in \mathcal{C}$ and $r \geq r_x$ that there exists a $k \times k$ matrix T such that $T \circ V$ is an ϵ-gauge on $B_r(x)$. The idea for this is related to the ideas of [CN15]. We will also discuss some applications of these estimates which will be useful later in the paper.

In the next section we will tackle the more refined estimates on ϵ-gauge's, which will tell us for *most* points $x \in \mathcal{C}$ that for any ball $B_r(x)$ with $r \geq r_x$, the V^a remain an ϵ-gauge, even without transformation. This result will be crucial in the proof of the energy identity and L^1 hessian estimate. However, our first main result of this section is the following, which begins by showing the existence of harmonic ϵ-gauge's on sufficiently symmetric balls:

Theorem 6.2. *Let A be a stationary Yang-Mills connection with $\fint_{B_4} |F_A|^2 \leq \Lambda$ and $B_4(p)$ a δ-weakly flat ball. For each $\epsilon > 0$ if $\delta < \delta(n, k, \Lambda, \epsilon)$, then there exists a harmonic ϵ-gauge $V^a \in \Gamma(B_2, E)$.*

With the existence of an ϵ-gauge established, we would like to understand how close to being a gauge the V^a actually are. The L^2 condition only concludes that the the V^a form a gauge on a set of almost full measure. This is actually pretty weak, and we would like to understand that the V^a form a gauge on all of $\mathcal{A}$. In fact, it will be important for us to understand much more than this and have some effective control on the behavior of the V^a at small scales. Motivated by the transformation theorem which is to come, let us make the following definition:

Definition 6.3. Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} \equiv B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$, and let $V^a \in \Gamma(B_4, E)$ be an ϵ-gauge. Given $x \in \mathcal{C}$ with $r_x \leq r \leq 2$ and $\theta \equiv 10^{-2}$ we define the $k \times k$ symmetric transformation matrix:

$$T = T(x, r) = \left(\fint_{\mathcal{W}_r^\theta(x)} \langle V^a, V^b \rangle \right)^{-1/2}. \tag{6.1}$$

Remark 6.1. Recall that the wedge regions are defined by $\mathcal{W}^\theta(x) \equiv \{y \in B_2 : d(y, \mathcal{L}_x) \geq \cos(\theta)\, d(y, x)\}$ and $\mathcal{W}_r^\theta(x) \equiv \mathcal{W}^\theta(x) \cap A_{\cos\theta\, r, r/\cos\theta}(x)$.

Remark 6.2. Apriori T may have infinite eigenvalues as defined, however we will see in Theorem 6.4 below that this is not the case.

Remark 6.3. For the sake of the theorems of this section one could have made the slightly simpler definition $T = (\fint_{B_r(x)} \langle V^a, V^b \rangle)^{-1/2}$, where one averages over a ball instead of a small portion of it. Taking the average away from the singular set becomes important in the next section, when we try and control in a more refined manner the behavior of T. In this case, if one were to average over all of B_r, then this adds small errors at every scale which may potentially pile up.

The following sums up the the use of the transformation matrices:

Theorem 6.4. *Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} \subseteq B_2(p)$ satisfying* (1.11) *and $\fint_{B_4} |F_A|^2 \leq \Lambda$, and let $V^a \in \Gamma(B_4, E)$ be a δ-gauge. For each $\epsilon > 0$, if $\delta < \delta(n, \Lambda, \epsilon)$, then for all $B_{2r}(x) \subseteq B_2$ with $x \in \mathcal{C}$ we have that*

(1) $T(x, r)$ *is nondegenerate, in fact $1 - \epsilon \leq T \leq (1 + \epsilon) r^{-\epsilon}$.*
(2) $\tilde{V}^a \equiv (T \circ V)^a = T_b^a V^b$ *is an ϵ-gauge on $B_r(x)$.*

The first immediate corollary is that V^a is everywhere nondegenerate on $\mathcal{A}$ and thus defines a vector bundle gauge. The second immediate corollary, which is itself not otherwise obvious, is that by using the uniform lower bound on $T(x, r)$ and standard elliptic estimates we have for each $x \in \mathcal{A}_{10^{-6}}$ the pointwise estimates

$$d(x, \mathcal{C})|\nabla V|(x) < C(n)\epsilon,$$
$$d(x, \mathcal{C})^2 |\nabla^2 V|(x) < C(n)\epsilon. \tag{6.2}$$

Indeed, the fact that $|\nabla V|$ and $|\nabla^2 V|$ are scale invariantly bounded follows from the global L^∞ estimate on $|V|$. However, the smallness of this bound is more subtle to prove, and will be directly used in Section 7.

6.1. Computation Properties and Basic Estimates of Harmonic Sections.
In this subsection we record some basic computation properties of harmonic sections, as well as some basic estimates over ϵ-regularity regions. Let us begin with the following computations for harmonic sections over a vector bundle E equipped with a Yang Mills connection:

Lemma 6.5. *Let $V \in \Gamma(B_r, E)$ satisfy $\Delta V = 0$, where E is equipped with a Yang Mills metric connection ∇_A. Then the following hold:*

(1) $\Delta|V|^2 = 2|\nabla V|^2 \geq 0$.

(2) $\Delta|\nabla V|^2 = 2|\nabla^2 V|^2 + 2F(\nabla V, \nabla V) + 2Rc(\nabla V, \nabla V)$.

Remark 6.4. We leave both as an easy exercise, but let us point out that the Yang-Mills condition plays a role in the second equation, as typically there is a div$F$ term which appears in this computation.

Let us now also record a basic estimate of ϵ-gauges inside ϵ-regularity regions:

Lemma 6.6. *Let $V \in \Gamma(B_{2r}, E)$ satisfy $\Delta V = 0$, where E is equipped with a Yang Mills metric connection ∇_A. Assume that we have the regularity scale estimate $r_A \geq 4r$. Then for $k \geq 1$ we have the estimate $r^k \sup_{B_r} |\nabla^k V| \leq C(n,k)\, r \fint_{B_{2r}} |\nabla V|$.*

Proof. Let us briefly outline this because it uses the Yang-Mills conditions in two ways. Since this is a scale invariant estimate we can assume $r = 1$ without any loss. First, as in Lemma 6.5 we may compute $\Delta \nabla V = F(\nabla V)$ using the Yang-Mills condition, and thus on $B_3(x)$ we have the inequality

$$\Delta|\nabla V| \geq -C(n)|\nabla V|, \tag{6.3}$$

so that $|\nabla V|$ satisfies a subharmonic inequality. We may therefore use the mean value theorem for nonnegative functions satisfying the above in order to conclude that $\sup_{B_{5/2}} |\nabla V| \leq C(n) \fint_{B_2} |\nabla V|$. Now we may use standard elliptic estimates on the equation $\Delta \nabla V = F(\nabla V)$ in order to conclude the result. $\quad\square$

6.2. **Proof of ϵ-Gauge Existence of Theorem 6.2.** We will prove the result by contradiction. Indeed, let us assume for some $\epsilon > 0$ that no such $\delta(n, k, \Lambda, \epsilon)$ exists. Then we can find a sequence of Yang-Mills connections A_i with $B_4(p_i)$ δ_i-weakly flat balls with $\delta_i \to 0$ for which there does exist an ϵ-gauge V^a on B_2.

After possibly composing with a rotation, there is no harm in assuming each weakly flat ball is with respect to the $n-4$ plane of symmetry $\mathcal{L}_i \equiv \mathcal{L} \equiv \mathbb{R}^{n-4} \times \{0\}$. Note that since $\delta_i \to 0$, we have that the underlying manifolds are converging

$$B_4(p_i) \xrightarrow{C^{1,\alpha}} B_4(0) \subseteq \mathbb{R}^n. \tag{6.4}$$

Additionally, we have on $B_4 \setminus B_{\delta_i}(\mathcal{L})$ that the curvature satisfies $|F_i| \to 0$. Thus if we restrict the bundles $E_i \to B_{\delta_i^{-1}} \setminus B_{\delta_i}(\mathcal{L})$ then after passing to a subsequence we can limit

$$E_i \to E,$$
$$A_i \to A, \tag{6.5}$$

where $E \to \mathbb{R}^n \setminus \mathcal{L}$ and A is a flat connection on E. In particular, we can pick global parallel sections $E^1, \ldots, E^k$ on E, and thus canonically extend E to the trivial bundle $\mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^n$.

Let us consider the convergence $B_{\delta_i^{-1}} \to \mathbb{R}^n$ and $E_i \to E$ in slightly more detail, as it will be useful in the coming analysis. Specifically, our convergence tells us that for all i sufficiently large we may find a $C^{2,\alpha}$

diffeomorphism and G-bundle maps such that the following diagram commutes:

$$
\begin{array}{ccc}
E_i & \xrightarrow{\ \ \varphi_E\ \ } & E \\
\pi_i \downarrow & & \downarrow \pi \\
B_{\delta_i^{-1}}(p_i) \setminus B_{\delta_i}(\mathcal{L}) & \xrightarrow{\ \ \varphi\ \ } & \mathbb{R}^n \setminus B_{\delta_i}(\mathcal{L})
\end{array}
\tag{6.6}
$$

Let us now define the sections $V_i^1, \ldots, V_i^k \in \Gamma(E_i, B_3)$ as the unique harmonic sections $\Delta V_i^a = 0$ in $B_3(p_i)$ which satisfy the Dirichlet boundary values

$$
V_i^a(x) = \begin{cases} \varphi_E^* E^a & \text{if } x \in \partial B_3(p_i) \setminus B_{\delta_i}(\mathcal{L}), \\ 0 & \text{if } x \in \partial B_3(p_i) \cap B_{\delta_i}(\mathcal{L}). \end{cases}
$$

Now let us see that for i sufficiently large this defines our ϵ-gauge. Let us begin by showing the pointwise estimate $|V_i^a| \le 1$. Indeed, we have on ∂B_3 that $|V_i^a| \le |\varphi_E^* E^a| \equiv 1$. On the other hand, by Lemma 6.5 we have that $|V|$ is a subharmonic function. Therefore by a maximum principle we have the estimate $|V_i^a| \le 1$ on all of B_3.

We will want to show the gradient estimate $\fint_{B_4} |\nabla V_i^a|^2 < \epsilon^2$. Let us begin by proving the related, but apriori weaker, estimate that if $B_{8r}(x) \subseteq B_3$ then $r^2 \fint_{B_r} |\nabla V_i^a|^2 < C(n)$. Indeed, let $\phi : B_{8r}(x) \to \mathbb{R}$ be a smooth cutoff function with

$$
\phi \equiv 1 \text{ on } B_{4r}, \quad \phi \equiv 0 \text{ outside } B_{7r},
$$
$$
r |\nabla \phi|, r^2 |\nabla^2 \phi| < C(n). \tag{6.7}
$$

Thus we can compute

$$
\fint_{B_{4r}} |\nabla V_i^a|^2 \le C(n) r^{-n} \int \phi |\nabla V_i^a|^2 = C(n) r^{-n} \int \phi \Delta |V_i^a|^2
$$
$$
= C(n) r^{-n} \int \Delta \phi |V_i^a|^2 \le C(n) r^{-2} \sup_{B_{7r}} |V_i^a|^2 \le C(n) r^{-2}, \tag{6.8}
$$

where in the middle equality we have used Lemma 6.5.1.

The global L^∞ bound on V_i^a combined with the curvature bound $|F_i| \to 0$ on $B_4(p_i) \setminus B_{\delta_i}(\mathcal{L})$ tells us that after passing to a subsequence we can limit

$$
V_i^a \to V^a \in \Gamma(E, B_3(0^n) \setminus \mathcal{L}), \tag{6.9}
$$

where the convergence is in $C^{1,\alpha}$ on compact subsets of $\overline{B_3}(0^n) \setminus \mathcal{L}$. Using $|V_i^a| \le 1$, $\Delta V_i^a = 0$ and (6.8) we therefore obtain

$$
|V^a| \le 1,
$$
$$
V^a \in H_{\text{loc}}^1(B_3),
$$
$$
\Delta V^a = 0 \text{ on } B_3 \setminus \mathcal{L}. \tag{6.10}
$$

Note that we have already identified E as the trivial $\mathbb{R}^k$ bundle over $\mathbb{R}^n$, and thus we may view $V^a : B_3 \setminus \mathcal{L} \to \mathbb{R}^k$ canonically.

Using the H^1_{loc} estimate above and that the set $\mathcal{L}$ has capacity zero we can therefore conclude that V^a extends a smooth solution of $\Delta V^a = 0$ over all B_3. On the other hand, since the convergence is smooth on compact subsets of $\overline{B_3} \setminus \mathcal{L}$ we know that $V^a = E^a$ on ∂B_3. The E^a are themselves harmonic, so by uniqueness if $V^a = E^a$ on ∂B_3 then $V^a = E^a$ on all of B_3. Thus we have concluded that

$$V_i^a \to E^a, \tag{6.11}$$

where the convergence is $C^{1,\alpha}$ on compact subsets of $\overline{B_3} \setminus \mathcal{L}$. In particular, we have that

$$\langle V_i^a, V_i^b \rangle \to \delta^{ab} \text{ pointwise on } B_3 \setminus \mathcal{L}. \tag{6.12}$$

Using the pointwise bound $|V_i^a| \le 1$ and that $\text{Vol}(B_3 \cap B_r(\mathcal{L})) \to 0$ as $r \to 0$ one therefore easily concludes the L^2 almost orthogonality

$$\fint_{B_3} \left| \langle V_i^a, V_i^b \rangle - \delta^{ab} \right| \to 0. \tag{6.13}$$

Finally, if we can show that $\fint_{B_2} |\nabla V_i^a|^2 \to 0$, then we will proved that for all i sufficiently large V_i^a forms an ϵ-gauge, which is our desired contradiction. Thus, for any $x \in B_2$ let us consider the ball $B_r(x) = B_{1/16}(x)$ with $B_{8r} \subseteq B_3$, and let us consider the cutoff function ϕ from (6.7). Then we can make a slightly more refined version of the previous computation in order to conclude

$$\begin{aligned}
\fint_{B_r(x)} |\nabla V_i^a|^2 &\le C(n) \int \phi |\nabla V_i^a|^2 = C(n) \int \phi \Delta |V_i^a|^2 \\
&= C(n) \int \phi \Delta \left(|V_i^a|^2 - 1 \right) = C(n) \int \Delta \phi \left(|V_i^a|^2 - 1 \right) \\
&\le C(n) \fint_{B_3} \left| |V_i^a|^2 - 1 \right| \to 0.
\end{aligned} \tag{6.14}$$

Since this hold for any ball $B_{1/16}(x)$ with $x \in B_2$ we have $\fint_{B_2} |\nabla V_i^a|^2 \to 0$ as claimed, which proves our desired contradiction and thus proves the theorem. $\square$

6.3. **Proof of Transformation Theorem 6.4.**

The proof will be by contradiction. Therefore let us assume for some $\epsilon > 0$ the result fails, and thus we can find a sequence of δ_i-annular regions $\mathcal{A}_i \subseteq B_2(p_i)$ with δ_i-gauges $V_i^1, \ldots, V_i^k \in \Gamma(B_4(p_i), E_i)$ such that the result fails for each i with $\delta_i \to 0$. Let us choose $x_i \in \mathcal{C}_i$ to be one of the points for which the result fails, and define

$$r_i \equiv \min\{ r_{x_i} \le r < 2 : \forall\, r \le s \le 2 \text{ we have that } (1) \text{ and } (2) \text{ hold on } B_s(x_i) \}. \tag{6.15}$$

We have by assumption that $r_i > r_{x_i}$, and therefore either (1) or (2) must fail for some radius $r > \frac{1}{2} r_i$. Notice that $r_i \to 0$ since $\delta_i \to 0$. Let $T_i \equiv T(x_i, r_i)$ and let us denote $\tilde{V}_i^a \equiv T_i \circ V_i$ to be the transformed sections. Since the result holds for r_i we have that $\tilde{V}_i^a$ is an ϵ-gauge on $B_{r_i}(x_i)$, but note that we cannot have that $\tilde{V}_i^a$ is a $10^{-2n} \epsilon$-gauge on $B_{r_i}(x_i)$. If this were to hold, then it is clear that (1) and (2) must still be satisfied for all $r \ge \frac{1}{2} r_i$, which is not the case.

Therefore let us rescale the geometry so that $B_{r_i}(x_i) \to B_2(\tilde{x}_i)$, so that $\tilde{V}_i^a \in \Gamma(B_{r_i^{-1}}(\tilde{x}_i), E_i)$ is an ϵ-gauge on $B_2(\tilde{x}_i)$. After rotation we may assume the best plane $\mathcal{L}_i \equiv \mathcal{L}$ for each annular region $\mathcal{A}_i$ is a constant. Let us begin with the following claim:

Claim: For each $1 \le r \le r_i^{-1}$ and $0 < s < 1$ we have the estimates

(1) $\sup_{B_r} |\tilde{V}_i| \le C(n)r^\epsilon$.

(2) $r^2 \fint_{B_r(\tilde{x}_i)} |\nabla \tilde{V}_i^a|^2 \le C(n)r^{2\epsilon}$.

(3) $\int_{B_r \cap B_s(\mathcal{L})} |\nabla \tilde{V}_i^a|^2 \le C(n)r^{n-4+2\epsilon}s^2$.

To prove the claim let us observe that for every $1 \le r \le r_i^{-1}$ we have by the definition of r_i that if $\tilde{T} \equiv \left(\fint_{W_r^\theta(\tilde{x}_i)} \langle \tilde{V}_i^a, \tilde{V}_i^b \rangle \right)^{-1/2}$, then by condition (2) we have that $\tilde{T} \circ \tilde{V}_i^a$ is an ϵ-gauge on $B_r(x_i)$. In particular,

$$\sup_{B_r} |\tilde{T} \circ \tilde{V}_i^a| \le 1 + \epsilon. \tag{6.16}$$

However, by condition (1), rescaled since our original ball B_{r_i} now has radius 2, we have the estimate

$$\frac{1}{2}r^{-\epsilon} \le \tilde{T} \le 2r^{-\epsilon}. \tag{6.17}$$

Plugging this into (6.16) we arrive at the first estimate of the claim. The second and third estimates are proved by a verbatim argument, so let us focus on the second. Choose a cutoff function ϕ so that $\phi \equiv 1$ on $B_r(\tilde{x}_i)$, $\phi \equiv 0$ outside of $B_{2r}(\tilde{x}_i)$ and $r|\nabla\phi|, r^2|\nabla^2\phi| \le C(n)$. Multiplying both sides of Lemma 6.5.1 by ϕ and integrating we arrive at

$$\fint_{B_r} |\nabla \tilde{V}_i^a|^2 \le C(n)r^{-n} \int \phi|\nabla \tilde{V}_i^a|^2 = C(n)r^{-n} \int \phi\Delta|\tilde{V}_i^a|^2$$

$$= C(n)r^{-n} \int \Delta\phi|\tilde{V}_i^a|^2$$

$$\le C(n)r^{-2} \sup_{B_{2r}} |\tilde{V}_i^a|^2 \le C(n)r^{-2}r^{2\epsilon}, \tag{6.18}$$

which finishes the proof of the claim. $\square$

Now as in the proof of Theorem 6.2 using that $\delta_i, r_i \to 0$ we may pass to a subsequence in order to limit our spaces

$$B_{r_i^{-1}}(\tilde{x}_i) \to \mathbb{R}^n,$$

$$E_i \to E \equiv \mathbb{R}^k \times \mathbb{R}^n, \tag{6.19}$$

where E is apriori a flat bundle over $\mathbb{R}^n \setminus \mathcal{L}$ which may be canonically extended to the trivial bundle $\mathbb{R}^k \times \mathbb{R}^n$. Using the estimates of the previous claim we may also pass to a subsequence to also limit

$$\tilde{V}_i^a \to \tilde{V}^a \in \Gamma(E, \mathbb{R}^n), \tag{6.20}$$

where the convergence is smooth on $\mathbb{R}^n \setminus \mathcal{L}$. By using (3) of the claim and that $\tilde{V}_i^a \to \tilde{V}^a$ smoothly on $\mathbb{R}^n \setminus \mathcal{L}$ we also see that

$$\fint_{B_r(x)} |\nabla \tilde{V}_i^a|^2 \to \fint_{B_r(x)} |\nabla \tilde{V}^a|^2, \tag{6.21}$$

so that the H^1 norms converge.

Now using that E is flat and trivial we may view $\tilde{V}^a : \mathbb{R}^n \to \mathbb{R}^k$, and by the estimates of the previous Claim we have

(1) $\sup_{B_r} |\tilde{V}| \leq C(n) r^\epsilon$.
(2) $r^2 \fint_{B_r(x_i)} |\nabla \tilde{V}|^2 \leq C(n) r^\epsilon$.

Since the convergence of $\tilde{V}_i^a \to \tilde{V}$ is smooth on $\mathbb{R}^n \setminus \mathcal{L}$ we have that $\tilde{V}$ is harmonic on $\mathbb{R}^n \setminus \mathcal{L}$. However, since $\tilde{V} \in H^1_{\text{loc}}$ by (2) above and since $\mathcal{L}$ is a set with zero capacity, we have that $\tilde{V}^a$ extends to a smooth harmonic function on all of $\mathbb{R}^n$.

However, since $|\tilde{V}|$ is growing at most at a small polynomial rate, we have by Liouville's theorem that

$$|\nabla \tilde{V}^a| \equiv 0 \,. \tag{6.22}$$

In particular, we have that $\langle \tilde{V}^a, \tilde{V}^b \rangle = constant$ for each a, b. However, by construction we also have that

$$\fint_{W_2^\theta(0)} \langle \tilde{V}^a, \tilde{V}^b \rangle = \delta^{ab} \,. \tag{6.23}$$

Combining these points we get that

$$\langle \tilde{V}^a, \tilde{V}^b \rangle = \delta^{ab} \,, \tag{6.24}$$

on all of $\mathbb{R}^n$. In particular, $\tilde{V}^a$ is a 0-gauge. Recall now that while $\tilde{V}_i^a$ is an ϵ-gauge on B_2, by the construction of r_i it is not a $10^{-2n}\epsilon$-gauge. However, using (6.21) and that $\tilde{V}^a$ is a 0-gauge we see that for i sufficiently large this is our desired contradiction, and thus we have proved the Theorem. $\square$

7. ϵ-GAUGE'S ON ANNULUS REGIONS

In the previous section we showed the existence of ϵ-gauge's on δ-annular regions and proved some basic estimates. In this section we study more carefully the properties of such ϵ-gauge's and prove our main analytic estimates. There are two main results we wish to prove and discuss in this section. The first is that we will see that on most of the δ-annular region our ϵ-gauge is a legitimate vector Coulomb gauge which is ϵ-orthonormal. Precisely:

Theorem 7.1. *Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ and $\fint_{B_2} |F_A|^2 \leq \Lambda$, and let V be a δ-gauge on B_4. For each $\epsilon > 0$ if $\delta < \delta(n, k, \Lambda, \epsilon)$ then there exists a subset $\mathcal{C}_\epsilon \subseteq \mathcal{C} \cap B_1$ such that*

(1) $\mu(\mathcal{C}_\epsilon) \geq (1 - \epsilon)\mu(\mathcal{C} \cap B_1)$.
(2) *For each $x \in \mathcal{C}_\epsilon$ and $r_x \leq r \leq 1$ we have that V is an ϵ-gauge on $B_r(x)$.*

In fact the above result will eventually follow from the scale invariant gradient estimate discussed in the next theorem, which is where most of the work of this section focuses. Precisely, we have the following:

Theorem 7.2. *Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ and $\fint_{B_2} |F_A|^2 \leq \Lambda$, and let V be a δ-gauge on B_4. For each $\epsilon > 0$ if $\delta < \delta(n, k, \Lambda, \epsilon)$ then we have the following estimates:*

$$\int_{\mathcal{A}_{10^{-4}} \cap B_{3/2}} r_A^{-3} |\nabla V|, \quad \int_{\mathcal{A}_{10^{-4}} \cap B_{3/2}} r_A^{-2} |\nabla^2 V|, \quad \int_{\mathcal{A}_{10^{-4}} \cap B_{3/2}} |\nabla^4 V| < \epsilon.$$

(7.1)

7.1. Annular Green's Function.
In this subsection we introduce and study the Green's function $G_{\mathcal{A}}$ associated to an annular region. Recall that $G_x(y) \sim \alpha_n |x - y|^{2-n}$ is the standard Green's function, which is the solution of $-\Delta G_x = \delta_x$. Since we work under the assumption that $K \leq \delta$, by (1.7) and standard estimates, we have that there exists a constant C for which $C^{-1} d(x, y)^{2-n} \leq G_x(y) \leq C d(x, y)^{2-n}$, and also $C^{-1} d(x, y)^{1-n} \leq |\nabla G_x(y)| \leq C d(x, y)^{1-n}$. The annular version of the Green's function $G_{\mathcal{A}}$ satisfies the following:

Definition 7.3. Let $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ be a δ-annular region with packing measure μ. Then we define:
 (1) The annular Green's function $G_{\mathcal{A}}(y) \equiv \int G_x(y) \, d\mu[x]$, which is the global solution of $-\Delta G_{\mathcal{A}} = \mu$.
 (2) The annular distance function $b(y) = b_{\mathcal{A}}(y)$ which is defined by the formula $G_{\mathcal{A}} \equiv b^{-2}$.

Notice in the above that if one viewed $\mathcal{A}$ as a perfect annulus $\mathcal{A} \equiv B_2(0^n) \setminus \mathbb{R}^{n-4}$, then $b_{\mathcal{A}}(y) \propto d(y, \mathbb{R}^{n-4})$ would be the distance to the singular set. Therefore b is our smooth approximation to such a distance. Let us see that this is a fair interpretation in the general case:

Lemma 7.4. *Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ satisfying $\fint_{B_2} |F_A|^2 \leq \Lambda$, and let $b(y)$ be the annular distance function. Then if $\delta < \delta(n, k, \Lambda)$ then there exists $C(n)$ such that the following hold:*
 (1) $C^{-1} d(y, \mathcal{C}) < b(y) < C d(y, \mathcal{C})$ *for all $y \in \mathcal{A}_{10^{-6}}$.*
 (2) $C^{-1} < |\nabla b| < C$ *on $\mathcal{A}_{10^{-6}}$.*

Remark 7.1. Recall that $\mathcal{A}_s \equiv B_2 \setminus \overline{B}_{s \cdot r_x}(\mathcal{C})$ is the extended annulus.

Proof. Let $y \in \mathcal{A}_{10^{-6}}$ with $x \in \mathcal{C}$ the closest point of $\mathcal{C}$ to y and $r \equiv 2d(x, y)$. Let us denote the sequence of scales $r_\alpha \equiv 2^\alpha r$, then we can write

$$G_\mu(y) = \int G_z(y) \, d\mu[z] \sim \int_{B_r(x)} G_z(y) \, d\mu[z] + \sum_{\alpha \geq 0} \int_{A_{r_{\alpha+1}, r_\alpha}(x)} G_z(y) \, d\mu[z].$$

(7.2)

Thus our upper and lower bounds are derived from the estimates

$$G_\mu(y) \leq C(n) \sum_\alpha r_\alpha^{2-n} \cdot r_\alpha^{n-4} = C(n) \sum 2^{-2\alpha} r^{-2} \leq C(n) r^{-2} = C(n) d(x, \mathcal{C})^{-2},$$

$$G_\mu(y) \geq \int_{B_r(x)} \alpha_n d^{2-n}(y, z) \, d\mu[z] \geq C(n)^{-1} r^{2-n} r^{n-4} = C(n)^{-1} r^{-2},$$

(7.3)

where we have used the Ahlfor's upper bounds proved in theorem 5.3.

The upper bound on the gradient estimate is proved similarly with

$$|\nabla G_\mu|(y) \le C(n) \int_{B_r(x)} \alpha_n d^{1-n}(y, z) \, d\mu[z] + C(n) \sum_{\alpha \ge 0} \int_{A_{r_{\alpha+1}, r_\alpha}(x)} \alpha_n d^{1-n}(y, z) \, d\mu[z]$$

$$\le C(n) \sum r_\alpha^{1-n} \cdot r_\alpha^{n-4} \le C(n) r^{-3} = C(n) d(x, \mathcal{C})^{-3} . \tag{7.4}$$

The lower bound on the gradient takes a little bit more work. Let us consider the radial vector at y given by $v = \nabla d(x, y)$. Note that for every $z \in \mathcal{C} \cap B_r(x)$ that $\nabla_v G_z(y) > C^{-1} d(z, y)^{1-n}$. Further, by condition $(a3)$ we have that for every $\beta > 0$ if $\delta < \delta(\beta)$ then for every $z \in B_{r_\beta}(x) \cap \mathcal{C}$ we have that $\nabla_v G_z(y) > 0$. In particular, if this holds for a given β then we can estimate

$$|\nabla G_\mu|(y) \ge \nabla_v G_\mu(y) \ge \int_{B_r(x)} \nabla_v G_z(y) \, d\mu[z] + \sum_{0 \le \alpha \le \beta} \int_{A_{r_{\alpha+1}, r_\alpha}(x)} \nabla_v G_z(y) \, d\mu[z]$$

$$+ \sum_{\beta+1 \le \alpha} \int_{A_{r_{\alpha+1}, r_\alpha}(x)} \nabla_v G_z(y) \, d\mu[z]$$

$$> 2C(n)^{-1} r^{1-n} r^{n-4} - C(n) \sum_{\alpha \ge \beta+1} r_\beta^{1-n} r_\beta^{n-4} ,$$

$$\ge (2C(n)^{-1} - C(n)2^{-\beta}) r^{-3} . \tag{7.5}$$

Thus if $\beta = \beta(n)$ then we obtain the estimate

$$|\nabla G_\mu|(y) \ge C(n)^{-1} r^{-3} , \tag{7.6}$$

which completes the proof of the Lemma.   □

The following straightforward but useful computations are at the heart of what we will use the annular distance functions for:

Lemma 7.5. *Let* $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ *be a* δ-*annular region with packing measure* μ *and annular distance function* $b(y)$. *If* f *is a smooth function let us define* $S(r) = r \cdot r^{-3} \int_{b=r} f|\nabla b|$. *Then we have*

$$r \frac{d}{dr} \left(r \frac{d}{dr} S \right) = S + \int_{b=r} \Delta f \, |\nabla b|^{-1} . \tag{7.7}$$

Proof. Let us note that $\Delta b = \frac{3}{b} |\nabla b|^2$ and that the mean curvature of the level set $b = r$ is given by

$$H_{b=r} = \text{div}\left(\frac{\nabla b}{|\nabla b|} \right) = 3 \frac{|\nabla b|}{b} - \frac{\langle \nabla b, \nabla |\nabla b| \rangle}{|\nabla b|^2} . \tag{7.8}$$

One can then compute

$$r \frac{d}{dr} S = S + r^{-1} \int_{b \le r} \Delta f . \tag{7.9}$$

Applying $r \frac{d}{dr}$ again leads to the result.   □

7.2. **Super Convexity for Scale Invariant L^1 Gradient.** In this subsection we derive a superconvexity estimate for the gradient $|\nabla V|$ of an ϵ-gauge on an annular region. This estimate will turn out to be the key technical tool in the proofs of Theorem 7.2 and Theorem 7.1.

We will begin by defining a convenient cutoff function associated to an annular region. For this let $0 \leq \phi(s) \leq 1$ be a fixed smooth cutoff with $\phi \equiv 0$ for $|s| \leq \frac{8}{10}$, $\phi \equiv 1$ for $|s| \geq \frac{9}{10}$ and with the estimates $|\frac{d^k}{ds^k}\phi| \leq C(k)$. For each $x \in B_2$ and $0 < r \leq 10$ we can then define $\phi_{x,r}(y) \equiv \phi(r^{-2}d^2(x,y))$. Associated to an annular region $\mathcal{A} \subseteq B_2(p)$ we then define the cutoff

$$\phi_{\mathcal{A}}(y) \equiv (1 - \phi_{p,2})(y) \cdot \prod_{x \in \mathcal{C}} \phi_{x,10^{-5}r_x}(y) = (1 - \phi_{p,2}(y)) \cdot \tilde{\phi}_{\mathcal{A}}(y). \tag{7.10}$$

Using $(a1) \rightarrow (a4)$ and Remark 5.3 it is easy to check the following properties of the cutoff

$$\phi_{\mathcal{A}} \equiv 1 \text{ in } \mathcal{A}_{10^{-5}},$$
$$\text{supp}\,|\nabla\tilde{\phi}_{\mathcal{A}}| \subseteq B_{10^{-5}r_x}(\mathcal{C}) \setminus B_{10^{-6}r_x}(\mathcal{C}) \subseteq \mathcal{A}_{10^{-6}},$$
$$|\nabla^{(k)}\phi_{\mathcal{A}}| \leq C(n,k)r_x^{-k} \text{ in each } B_{r_x}(x). \tag{7.11}$$

Let us begin with the main computation of this subsection:

Proposition 7.6. *Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ satisfying $\fint_{B_2} |F_A|^2 \leq \Lambda$, and let V be a δ-gauge on B_4. Let us define the scale invariant quantity $S(r) = r \cdot r^{-3}\int_{b=r} |\nabla V|\,\varphi_{\mathcal{A}}\,|\nabla b|$, then for each $\epsilon > 0$ if $\delta < \delta(n,k,\epsilon,\Lambda)$ then:*

$$r\frac{d}{dr}\Big(r\frac{d}{dr}S\Big) \geq (1 - \epsilon)^2 S - e(r), \tag{7.12}$$

where $|e(r)| \leq \epsilon\mu(\{x : C^{-1}r \leq r_x \leq Cr\}) + Cr$ for $C = C(n)$.

Proof. Let us consider $\epsilon' > 0$, which will eventually be chosen by $\epsilon' = \epsilon'(n,\epsilon)$. We can use Lemma 7.5 in order to compute

$$r\frac{d}{dr}\Big(r\frac{d}{dr}S\Big) = S + \int_{b=r} \Big(\Delta|\nabla V|\,\phi + 2\langle\nabla|\nabla V|,\nabla\phi\rangle + |\nabla V|\Delta\phi\Big)|\nabla b|^{-1}. \tag{7.13}$$

Using Lemma 6.5 and that our cutoff satisfies (7.10) $\text{supp}\,\phi_{\mathcal{A}} \subseteq \mathcal{A}_{10^{-6}}$ we have that

$$\Delta|\nabla V|\,\phi \geq -C(n)\delta\,d(x,\mathcal{C})^{-2}|\nabla V|\phi. \tag{7.14}$$

If we combine this with Lemma 7.4 and integrate we arrive at

$$\int_{b=r} \Delta|\nabla V|\,\phi \geq -C(n)\delta\,b^{-2}\int_{b=r} |\nabla V|\,\phi \geq -\epsilon S. \tag{7.15}$$

In order to estimate the other error terms let us recall that we can write $\phi_{\mathcal{A}} = \phi_{p,2} \cdot \tilde{\phi}_{\mathcal{A}}$ as in (7.10). Using Theorem 6.4 with $\epsilon' > 0$ together with (6.2) and (7.11) we can therefore write

$$2|\langle\nabla|\nabla V|,\nabla\phi_{\mathcal{A}}\rangle|(y) \leq C(n)\epsilon'b^{-2} + C(n)\epsilon'\sum_{x \in \mathcal{C} \cap B_{18/10}} r_x^{-3}\chi[B_{C(n)r_x}(x) \cap \{C(n)^{-1}r_x < b(y) < C(n)r_x\}],$$

$$|\nabla V||\Delta\phi|(y) \leq C(n)\epsilon'b^{-1} + C(n)\epsilon'\sum_{x \in \mathcal{C} \cap B_{18/10}} r_x^{-3}\chi[B_{C(n)r_x}(x) \cap \{C(n)^{-1}r_x < b(y) < C(n)r_x\}]. \tag{7.16}$$

Integrating these last two terms and using Lemma 7.4 we arrive at

$$\int_{b=r} 2|\langle \nabla|\nabla V|, \nabla \phi \rangle| \le C(n)\epsilon' b + C(n)\epsilon' \sum_{x \in B_{18/10} \cap \{r_x \in [C^{-1}r, Cr]\}} r^{n-4} \le C(n)\epsilon' \mu(\{C^{-1}r < r_x < Cr\}),$$

$$\int_{b=r} |\nabla V| |\Delta \phi| \le C(n)\epsilon' b^2 + C(n)\epsilon' \mu(\{C^{-1}r < r_x < Cr\}). \tag{7.17}$$

Choosing $\epsilon' = C(n)^{-1}\epsilon$ we have arrived at our conclusion. $\qquad \square$

7.3. **Dini Estimates and Superconvexity.** In order to exploit Proposition 7.6 we will apply a maximum principle and study solutions of the underlying superconvex equation. The following tells us how to estimate the Dini integral of solutions:

Proposition 7.7. *For each $R > 0$ and $e(r)$ the solution of $r\frac{d}{dr}\left(r\frac{d}{dr}\bar{S}\right) = (1 - \epsilon)^2 \bar{S} - e(r)$ with $\bar{S}(0) = \bar{S}(R) = 0$ satisfies the Dini estimate*

$$\int_0^R \frac{\bar{S}(r)}{r} \le \frac{1}{(1 - \epsilon)^2} \int_0^R \frac{e}{s} + \frac{R^{-1+\epsilon}}{(1 - \epsilon)^2} \int_0^R s^{-\epsilon} e. \tag{7.18}$$

Proof. Observe that $r^{1-\epsilon}$ and $r^{-1+\epsilon}$ are solutions to the homogeneous equation. With this one can check that an explicit solution to $r\frac{d}{dr}\left(r\frac{d}{dr}\bar{S}\right) = (1 - \epsilon)^2 \bar{S} - e(r)$ under the conditions $\bar{S}(0) = \bar{S}(R) = 0$ is given by

$$\bar{S}(r) = \frac{1}{2(1-\epsilon)}\left(r^{-1+\epsilon} \int_0^r s^{-\epsilon} e + r^{1-\epsilon} \int_r^R s^{-2+\epsilon} e - \left(\frac{r}{R}\right)^{1-\epsilon} R^{-1+\epsilon} \int_0^R s^{-\epsilon} e\right). \tag{7.19}$$

From this we have the explicit computation

$$\int_0^R \frac{\bar{S}(r)}{r} = \frac{1}{2(1-\epsilon)}\left(\int_0^R r^{-2+\epsilon} \int_0^r s^{-\epsilon} e + \int_0^R r^{-\epsilon} \int_r^R s^{-2+\epsilon} e - \int_0^R r^{-\epsilon} R^{-2+2\epsilon} \int_0^R s^{-\epsilon} e\right). \tag{7.20}$$

Estimating each of these terms is similar, so let us just focus on the first. Indeed, by changing the order of integration we arrive at

$$\int_0^R \int_0^r r^{-2+\epsilon} s^{-\epsilon} e = \int_0^R \int_s^R r^{-2+\epsilon} s^{-\epsilon} e = \frac{-1}{1 - \epsilon} \int_0^R (R^{-1+\epsilon} - s^{-1+\epsilon}) s^{-\epsilon} e$$

$$= \frac{1}{1 - \epsilon}\left(\int_0^R s^{-1} e - R^{-1+\epsilon} \int_0^R s^{-\epsilon} e\right). \tag{7.21}$$

Arguing in a verbatim manner with the other terms leads to the conclusion of the lemma.

$\qquad \square$

Our main corollary of the above is the following, which gives a Dini estimate for our L^1 Hessian:

Corollary 7.8. *Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} = B_2 \setminus \bar{B}_{r_x}(\mathcal{C})$ satisfying $\fint_{B_2} |F_A|^2 \le \Lambda$, and let V be a δ-gauge on B_4. Let us define the scale invariant quantity $S(r) = r \cdot r^{-3} \int_{b=r} |\nabla V| \varphi_{\mathcal{A}} |\nabla b|$. Then for each $\epsilon > 0$ if $\delta < \delta(n, k, \Lambda, \epsilon)$ we have the estimate*

$$\int_0^\infty \frac{S(r)}{r} < \epsilon. \tag{7.22}$$

Proof. Recall from Proposition 7.6 that $S(r)$ solves the differential inequality

$$r\frac{d}{dr}\left(r\frac{d}{dr}S\right) \geq (1-\epsilon)^2 S - e(r), \tag{7.23}$$

where we have that

$$e(r) \leq C(n)\mu\left(\{x \in \mathcal{C} \cap B_{19/10} : C(n)^{-1}r \leq r_x \leq C(n)r\}\right) + C(n)r. \tag{7.24}$$

Note also that by the construction of the cutoff $\phi_{\mathcal{A}}$ that $S(0) = S(R) = 0$ where $R \leq R(n)$. Let us then consider the solution of

$$r\frac{d}{dr}\left(r\frac{d}{dr}\bar{S}\right) = (1-\epsilon)^2\bar{S} - e(r),$$
$$\bar{S}(0) = \bar{S}(R) = 0. \tag{7.25}$$

Note by a maximum principle applied to $\bar{S} - S$ we immediately yield the inequality $S(r) \leq \bar{S}(r)$. On the other hand, let us note that

$$\int_0^R \frac{e}{s} \leq C(n)\mu\{B_{18/10}\} + C(n) \leq C(n), \tag{7.26}$$

where in the last inequality we have used Theorem 5.3. Thus by applying Proposition 7.7 with $\epsilon' > 0$ we have the estimate

$$\int_0^\infty \frac{S}{r} \leq \int_0^R \frac{\bar{S}}{r} \leq C(n)\epsilon' < \epsilon, \tag{7.27}$$

where in the last line we have chosen $\epsilon' \equiv C(n)^{-1}\epsilon$. This completes the proof of the corollary. $\square$

7.4. **Proof of the L^1 gradient estimate of Theorem 7.2.** Using the coarea formula we may write

$$\int_{\mathcal{A}\cap B_{3/2}} d_{\mathcal{C}}^{-3}|\nabla V| \leq \int d_{\mathcal{C}}^{-3}|\nabla V|\phi_{\mathcal{A}} = \int_0^R \int_{b=r} d_{\mathcal{C}}^{-3}|\nabla V|\phi_{\mathcal{A}}|\nabla b|^{-1}, \tag{7.28}$$

where b is the μ-Green's distance associated to $\mathcal{A}$ and $R \leq R(n)$. Using Lemma 7.4 we can estimate this by

$$\int_{\mathcal{A}\cap B_{3/2}} d_{\mathcal{C}}^{-3}|\nabla V| \leq C(n)\int_0^R r^{-3}\int_{b=r}|\nabla V|\phi_{\mathcal{A}}|\nabla b| = C(n)\int_0^R \frac{S(r)}{r}. \tag{7.29}$$

Using Corollary 7.8 with $\epsilon' \equiv C(n)^{-1}\epsilon$ we then arrive at the desired estimate

$$\int_{\mathcal{A}\cap B_{3/2}} d_{\mathcal{C}}^{-3}|\nabla V| \leq C(n)\int_0^R \frac{S(r)}{r} < \epsilon, \tag{7.30}$$

as claimed. The other estimates then follow by combining this with Lemma 6.6. $\square$

7.5. **Transformation Estimates.** Recall from Theorem 6.4 that if $x \in \mathcal{C}$ then for every $r_x \leq r < 1$ there exists a matrix $T = T(x, r)$, given explicitly in Definition 6.3, such that $T \circ V$ is still an ϵ-gauge. The key to Theorem 7.1 is to see that for most $x \in \mathcal{C}$ and for all $r_x < r \leq 1$, this matrix is in fact close to the identity. In this subsection we see how to use the gradient estimate of Theorem 7.2 in order to control the transformation matrix T, which will be used in the next section is order finish the proof of Theorem 7.1.

The main technical result of this subsection is the following:

Proposition 7.9. *Given a δ-annular region* $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ *and a δ-gauge* $V^a \in \Gamma(B_4, E)$, *then for* $x \in \mathcal{C}$
and $r_x \leq r < 1$ *with* $\int_{W^{\pi/4}(x)} r_A^{1-n} |\nabla V| < \epsilon_n$ *we have for every* $\epsilon > 0$ *that if* $\delta < \delta(n, k, \Lambda, \epsilon)$ *then*

$$\left| Id - T(x, r) \right| < \epsilon + C(n) \int_{W^{\pi/4}(x)} r_A^{1-n} |\nabla V|. \tag{7.31}$$

In order to prove the above Proposition let us begin with the following Lemma, which is a more local version of the above:

Lemma 7.10. *Given a δ-annular region* $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ *and a δ-gauge* $V^a \in \Gamma(B_4, E)$, *then for* $x \in \mathcal{C}$ *and* $r_x \leq r < 1$ *we have for* $\gamma \equiv \frac{4}{5}$ *that*

$$\left| Id - T(x, \gamma \cdot r) T(x, r)^{-1} \right| < C(n) r \fint_{W_r^{\pi/4}(x)} |\nabla (T_{x,r} \circ V)| \tag{7.32}$$

Proof. Let us denote $\tilde{V}^a \equiv T_{x,r} \circ V^a$, and so we are trying to estimate $\left| Id - \tilde{T}(x, \gamma r) \right|$ where $\tilde{T}(x, \gamma \cdot r) \equiv$
$\left(\fint_{W_{\gamma \cdot r}^\theta(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right)^{-1/2}$ with $\theta \equiv 10^{-2}$. Using the definition of $\tilde{V}^a$ it is equivalent for us to estimate

$$\left| Id - \fint_{W_{\gamma \cdot r}^\theta(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right| = \left| \fint_{W_r^\theta(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W_{\gamma \cdot r}^\theta(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right|. \tag{7.33}$$

To accomplish this let us observe the relations

$$B_{10^{-2}r}\left(W_r^\theta(x) \right), B_{10^{-2}r}\left(W_{\gamma r}^\theta(x) \right) \subseteq W_r^{\pi/4}(x). \tag{7.34}$$

We will use a Poincaré to then conclude the result. Indeed, let us estimate

$$\left| \fint_{W_r^\theta(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W_r^{\pi/4}(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right| \leq \fint_{W_r^\theta(x)} \left| \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W_r^{\pi/4}(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right|,$$

$$\leq \frac{\text{Vol}(W_r^{\pi/4}(x))}{\text{Vol}(W_r^\theta(x))} \fint_{W_r^{\pi/4}(x)} \left| \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W_r^{\pi/4}(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right|,$$

$$\leq C(n) \fint_{W_r^{\pi/4}(x)} \left| \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W_r^{\pi/4}(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right|,$$

$$\leq C(n) r \fint_{W_r^{\pi/4}(x)} \left| \nabla \langle \tilde{V}^a, \tilde{V}^b \rangle \right|,$$

$$\leq C(n) r \fint_{W_r^{\pi/4}(x)} \left| \nabla \tilde{V} \right|, \tag{7.35}$$

where in the last line we have used the L^∞ bounds on $\tilde{V}$. A verbatim computation also gives

$$\left| \fint_{W_{\gamma r}^\theta(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W_r^{\pi/4}(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right| \leq C(n) r \fint_{W_r^{\pi/4}(x)} \left| \nabla \tilde{V} \right|. \tag{7.36}$$

Combining (7.35) with (7.36) we are able to estimate (7.33) by

$$\left| Id - \fint_{W^\theta_{\gamma \cdot r}(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right| = \left| \fint_{W^\theta_r(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W^\theta_{\gamma \cdot r}(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right|,$$

$$\leq \left| \fint_{W^\theta_r(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W^{\pi/4}_r(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right| + \left| \fint_{W^\theta_{\gamma r}(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle - \fint_{W^{\pi/4}_r(x)} \langle \tilde{V}^a, \tilde{V}^b \rangle \right|,$$

$$\leq C(n) r \fint_{W^{\pi/4}_r(x)} \left| \nabla \tilde{V} \right|, \tag{7.37}$$

which finishes the proof of the Lemma. $\qquad \square$

With the Lemma in hand we now finish the proof of Proposition 7.9:

Proof of Proposition 7.9. Let $x \in \mathcal{C}$ be such that $\int_{W^{\pi/4}(x)} r_A^{1-n} |\nabla V| < \epsilon_n$ holds, where ϵ_n will be chosen shortly. Let us consider the sequence of scales $s_\alpha \equiv \gamma^\alpha$ together with the associated matrices $T_\alpha \equiv T(x, s_\alpha)$. We will only prove the result for the T_α with $s_\alpha \geq r_x$. Using that the $T_\alpha \circ V^a$ are ϵ-gauges on $B_{s_\alpha}(x)$ the result easily extends to all radii $r \geq r_x$. Note also that since V^a is a δ-gauge we have for $\delta < \delta(n, k, \epsilon)$ that $|Id - T_0| < \epsilon$.

Now let us remark on the following. If $\alpha \geq 0$ is such that $|T_\alpha| \leq 2$, then by applying Lemma 7.10 we have the estimate

$$\left| T_\alpha - T_{\alpha+1} \right| \leq C(n) s_\alpha \fint_{W^{\pi/4}_{s_\alpha}(x)} \left| \nabla V \right|$$

$$\leq C(n) \int_{W^{\pi/4}(x) \cap A_{2^{-1} s_\alpha, 2 s_\alpha}} r_A^{1-n} \left| \nabla V \right|. \tag{7.38}$$

Iterating on this we see that if $T_\beta \leq 2$ for all $\beta \leq \alpha$ then we get the estimate

$$\left| T_{\alpha+1} - T_0 \right| \leq C(n) \sum_{\beta \leq \alpha} \int_{W^{\pi/4}(x) \cap A_{2^{-1} s_\beta, 2 s_\beta}} r_A^{1-n} \left| \nabla V \right|,$$

$$\leq C(n) \int_{W^{\pi/4}(x) \cap A_{2^{-1} s_\alpha, 2}} r_A^{1-n} \left| \nabla V \right|,$$

$$\leq C(n) \epsilon_n < 10^{-2}, \tag{7.39}$$

where in the last line we have chosen ϵ_n sufficiently small in a manner which depends only on dimension.

Combining this with the estimate $|T_0 - Id| < \epsilon_n$ we conclude that if $s_\alpha \geq r_x$ is such that $\left| T_\beta \right| \leq 2$ for all $\beta \leq \alpha$, then in fact if $s_{\alpha+1} \geq r_x$ then $\left| T_\beta \right| \leq 2$ for all $\beta \leq \alpha + 1$. Therefore we have the estimate

$$|T_\alpha| \leq 2 \text{ for all } s_\alpha \geq r_x. \tag{7.40}$$

In particular, we then we get that (7.39) holds for all $s_\alpha \geq r_x$, from which we get the estimate

$$\left| T_\alpha - Id \right| \leq |T_0 - Id| + C(n) \int_{W^{\pi/2}(x)} r_A^{1-n} \left| \nabla V \right|,$$

$$\leq \epsilon + C(n) \int_{W^{\pi/2}(x)} r_A^{1-n} \left| \nabla V \right|, \tag{7.41}$$

which finishes the proof of the Proposition. $\qquad \square$

7.6. **Proof of Theorem 7.1.** We finish the proof of Theorem 7.1 in this subsection. Let us begin by observing with $r_\alpha \equiv 2^{-\alpha}$ that if $\mathrm{supp}\,(f) \subseteq \mathcal{A}_{10^{-6}}$ then we have the following:

$$
\int \Big(\int_{W_x^{\pi/4}} d(x, \mathcal{C})^{1-n} |f| \Big) d\mu[x] \leq \int \Big(\sum_{r_\alpha \geq r_x} \int_{W_x^{\pi/4} \cap A_{\frac{1}{2}r_\alpha, 2r_\alpha}(\mathcal{C})} d(x, \mathcal{C})^{1-n} |f| \Big) d\mu[x]
$$

$$
\leq C(n) \int \Big(\sum_{r_\alpha \geq r_x} r_\alpha^{1-n} \int_{W_x^{\pi/4} \cap A_{\frac{1}{2}r_\alpha, 2r_\alpha}(\mathcal{C})} |f| \Big) d\mu[x] ,
$$

$$
\leq C(n) \sum_{r_\alpha} r_\alpha^{1-n} \mu(B_{4r_\alpha}(x)) \int_{A_{\frac{1}{2}r_\alpha, 2r_\alpha}(\mathcal{C})} |f| ,
$$

$$
\leq C(n) \sum_{r_\alpha} r_\alpha^{-3} \int_{A_{\frac{1}{2}r_\alpha, 2r_\alpha}(\mathcal{C})} |f| ,
$$

$$
\leq \int_{B_2} d(x, \mathcal{C})^{-3} |f| . \tag{7.42}
$$

With this in hand can finish Theorem 7.1. Indeed, by applying (7.42) to $f = |\nabla V| \chi_{\mathcal{A} \cap B_{3/2}}$ and by using Theorem 7.2 with $(\epsilon')^2$ we obtain the estimate

$$
\int_{B_1} \Big(\int_{W_x^{\pi/4}} d(x, \mathcal{C})^{1-n} |\nabla V| \Big) d\mu[x] \leq C(n) \int_{\mathcal{A} \cap B_{3/2}} r_A^{-3} |\nabla V| \leq C(n) (\epsilon')^2 . \tag{7.43}
$$

In particular, let us consider the set $\mathcal{C}' \equiv \{x \in B_1 \cap \mathcal{C} : \int_{W_x^{\pi/4}} d(x, \mathcal{C})^{1-n} |\nabla V| < \epsilon'\}$, then we see that $\mu(B_1 \setminus \mathcal{C}') \leq C(n)\epsilon'$. On the other hand, by applying Proposition 7.9 with ϵ' we see that for $\epsilon' < c(n, k)\epsilon$ that $\mathcal{C}' \subseteq \mathcal{C}_\epsilon$, and thus we have finished the proof of the Theorem. □

8. Proof of Curvature Estimates of Theorem 5.3

In this section we finish the proof of the L^1 hessian and L^2 curvature estimates on annular regions. The strategy of the proof is to use the scale invariant gradient estimates on V from Theorem 7.2 to show the estimates on $\mathcal{A}$ wherever V remains close to being an orthogonal basis. By Theorem 7.1 this is everything except a set of small $n - 4$ content, and therefore we can recover the rest and start the estimate over in a manner for which the inductive errors give rise to a geometric series. More slowly, let us start with the following, which is the main tool in our inductive construction:

Lemma 8.1. *Let A be a stationary Yang-Mills connection on a δ-annular region $\mathcal{A} = B_2 \setminus \overline{B}_{r_x}(\mathcal{C})$ satisfying $\fint_{B_2} |F_A|^2 \leq \Lambda$, and let V be a δ-gauge on B_4. For each $\epsilon > 0$ if $\delta < \delta(n, k, \Lambda, \epsilon)$ then there exists a collection $\{B_{r_j}(x_j)\}$ with $x_j \in \mathcal{C}$, $r_j > r_{x_j}$ and $\tilde{\mathcal{A}} = \mathcal{A} \setminus \bigcup B_{r_j}(x_j)$ such that*

(1) $B_{r_i/10}(x_i) \cap B_{r_j/10}(x_j) = \emptyset$.
(2) $\sum r_j^{n-4} < \epsilon$.
(3) $\int_{\tilde{\mathcal{A}} \cap B_1} |\nabla^2 F|, \int_{\tilde{\mathcal{A}} \cap B_1} |F|^2 < \epsilon$

Proof. For each $x \in \mathcal{C}$ and $\epsilon' > 0$, which will be fixed later, let us define the radius

$$
\bar{r}_x \equiv \min\{r_x \leq \bar{r} < 2 : |\langle V^a, V^b \rangle - \delta^{ab}| < \epsilon' \text{ in } W^{\pi/4}(x) \cap A_{\bar{r}, 2}(x)\} . \tag{8.1}
$$

With this let us consider the set

$$\bar{\mathcal{C}} \equiv \{x \in \mathcal{C} \cap B_1 : \bar{r}_x > r_x\}. \tag{8.2}$$

Note from Theorem 7.1 that for $\delta < \delta(n, k, \Lambda, \epsilon')$ that $\mu(\bar{\mathcal{C}}) < \epsilon'$. Let us start with the following claim:

Claim: We have the estimate $\mu(\bar{\mathcal{C}} \cap B_{\bar{r}_x/10}(x)) > c(n)\bar{r}_x^{n-4}$.

Indeed, let us consider two cases. If $\bar{r}_x < 2r_x$ then this is clear simply because $\mu(\bar{\mathcal{C}} \cap B_{\bar{r}_x/10}(x)) \geq \mu(x) > c(n)r_x^{n-4} \geq c(n)\bar{r}_x^{n-4}$.

On the other hand, if $\bar{r}_x \geq 2r_x$, then let $y \in W^{\pi/4}(x) \cap \partial B_{\bar{r}_x}(x)$ be such that $|\langle V^a, V^b \rangle - \delta^{ab}| = \epsilon'$. Let $x' \in \mathcal{C} \cap \bar{B}_{\bar{r}_x/20}(x)$ be one of the points of $\mathcal{C} \cap \bar{B}_{\bar{r}_x/20}(x)$ that is closest to y. Note then for every $z \in B_{\bar{r}_x/100}(x')$ that $y \in W^{\pi/4}(z)$ and with $d(y, z) > r_z$. This last inequality holds using $(a4)$ because $d(z, y) > \frac{3}{4}d(x, y) \geq \frac{3}{2}r_x > r_z$. In particular, we see that $B_{\bar{r}_x/100}(x') \subseteq \bar{\mathcal{C}}$ and thus $\mu(\bar{\mathcal{C}} \cap B_{\bar{r}_x/10}(x)) \geq \mu(B_{\bar{r}_x/100}(x')) > c(n)\bar{r}_x^{n-4}$ by the Ahlfor's regularity of Theorem 5.3. This finishes the proof of the Claim. $\square$

Now let us choose a Vitali subcovering $\{B_{\bar{r}_j}(x_j)\}$ with $x_j \in \bar{\mathcal{C}}$. Then using the above claim we can estimate

$$\sum \bar{r}_j^{n-4} \leq C(n) \sum \mu(\bar{\mathcal{C}} \cap B_{\bar{r}_j/10}(x_j)) \leq C(n)\mu(\bar{\mathcal{C}}) \leq C(n)\epsilon' < \epsilon, \tag{8.3}$$

where in the last inequality we have assumed $\epsilon' < C(n)^{-1}\epsilon$.

We claim now that our desired curvature estimates hold on the set $\tilde{\mathcal{A}} = \mathcal{A} \setminus \bigcup B_{r_j}(x_j)$. To prove this, note that, by Theorem 7.2, we have the estimates

$$\int_{\mathcal{A} \cap B_1} |\nabla^4 V| + \int_{\mathcal{A} \cap B_1} |\nabla^2 V|^2 + \int_{\mathcal{A} \cap B_1} r_A^{-3}|\nabla V| + \int_{\mathcal{A} \cap B_1} r_A^{-2}|\nabla^2 V| < \epsilon'. \tag{8.4}$$

Note that by the definition of curvature we have the estimates

$$|F(V)| \leq 2|\nabla^2 V|,$$
$$|\nabla^2 F(V)| \leq 2|\nabla^4 V| + 2|\nabla F||\nabla V| + |F||\nabla^2 V|. \tag{8.5}$$

In particular, if we are at a point $y \in \mathcal{A}$ such that $|\langle V^a, V^b \rangle - \delta^{ab}| < \epsilon'$ almost form an orthonormal basis, then we have the estimates

$$|F|(y) \leq 3 \sum_a |\nabla^2 V_a|(y),$$
$$|\nabla^2 F|(y) \leq 3 \sum_a \left(|\nabla^4 V_a|(y) + r_A^{-3}|\nabla V_a| + r_A^{-2}|\nabla^2 V_a| \right), \tag{8.6}$$

where we have used the scale invariant estimates $r_A^2|F|, r_A^3|\nabla F| << 1$ in the above. This estimate holds for each $y \in \bar{\mathcal{A}}$, and therefore we have that

$$\int_{\bar{\mathcal{A}} \cap B_1} |\nabla^2 F| \leq 3 \sum_a \left(\int_{\mathcal{A} \cap B_1} |\nabla^4 V_a| + \int_{\mathcal{A} \cap B_1} r_A^{-3}|\nabla V_a| + \int_{\mathcal{A} \cap B_1} r_A^{-2}|\nabla^2 V_a| \right) < 5\epsilon',$$
$$\int_{\bar{\mathcal{A}} \cap B_1} |F|^2 \leq 3 \sum_a \left(\int_{\mathcal{A} \cap B_1} |\nabla^2 V_a|^2 \right) < 3\epsilon', \tag{8.7}$$

which for $\epsilon' < \frac{1}{10}\epsilon$ finishes the proof of the Lemma. □

With the above in hand let us now finish the proof of Theorem 5.3:

Proof of Theorem 5.3. The proof of the curvature estimates of Theorem 5.3 is now just an inductive application of Lemma 8.1. Indeed, let us apply Lemma 8.1 with $\epsilon' = \epsilon'(n, \epsilon)$ which will be chosen later. Then we can write

$$\mathcal{A} \cap B_1 \subseteq \mathcal{A}_0 \cup \bigcup B_{r_{j_0}}(x_{j_0}), \tag{8.8}$$

such that

$$\int_{\mathcal{A}_0 \cap B_1} |\nabla^2 F|, \quad \int_{\mathcal{A}_0 \cap B_1} |F|^2, \quad \sum r_{j_0}^{n-4} < \epsilon'. \tag{8.9}$$

Now note that the restriction of $\mathcal{A}$ to $B_{2r_j^0}(x_j^0)$ is a δ-annular region for each j, and therefore we may again apply Lemma 8.1 to each of the balls $B_{2r_j^0}(x_j^0)$ to obtain

$$\mathcal{A} \cap B_{r_{j_0}}(x_{j_0}) \subseteq \mathcal{A}_{j_0,1} \cup \bigcup B_{r_{j_0,j_1}}(x_{j_0,j_1}), \tag{8.10}$$

such that

$$\int_{\mathcal{A}_{j_0,1} \cap B_{r_{j_0}}} |\nabla^2 F|, \quad \int_{\mathcal{A}_{j_0,1} \cap B_{r_{j_0}}} |F|^2, \quad \sum_{j_1} r_{j_0,j_1}^{n-4} < \epsilon' \, r_{j_0}^{n-4}. \tag{8.11}$$

Thus by defining

$$\mathcal{A}_1 \equiv \mathcal{A}_0 \cup \bigcup_{j_0} \mathcal{A}_{j_0,1},$$

$$\{B_{r_{j_1}}(x_{j_1})\} \equiv \bigcup_{j_0} \{B_{r_{j_0,j_1}}(x_{j_0,j_1}), \tag{8.12}$$

we have the covering

$$\mathcal{A} \cap B_1 \subseteq \mathcal{A}_1 \cup \bigcup B_{r_{j_1}}(x_{j_1}) \tag{8.13}$$

such that

$$\int_{\mathcal{A}_1 \cap B_1} |\nabla^2 F| \leq \epsilon' + \epsilon' \sum r_{j_0}^{n-4} \leq \epsilon' + (\epsilon')^2,$$

$$\int_{\mathcal{A}_1 \cap B_1} |F|^2 \leq \epsilon' + (\epsilon')^2,$$

$$\sum r_{j_1}^{n-4} \leq \epsilon' \sum r_{j_0}^{n-4} \leq (\epsilon')^2. \tag{8.14}$$

Now we can observe that each $\mathcal{A} \cap B_{2r_{j_1}}$ is a δ-annular region and repeat this process again. Again, if we repeat this process I times then we arrive at a covering

$$\mathcal{A} \cap B_1 \subseteq \mathcal{A}_I \cup \bigcup B_{r_{j_I}}(x_{j_I}) \tag{8.15}$$

such that

$$\int_{\mathcal{A}_I \cap B_1} |\nabla^2 F|, \quad \int_{\mathcal{A}_I \cap B_1} |F|^2 \leq \sum_{j=1}^{I} (\epsilon')^j, \quad \sum r_{j_I}^{n-4} \leq (\epsilon')^I. \tag{8.16}$$

In particular, we clearly get the much weaker estimate that $r_{j_I} \to 0$ as $I \to \infty$, and thus we have that the $\mathcal{A}_I$ form an exhaustion of $\mathcal{A}$. Passing to the limit we then get the estimate

$$\int_{\mathcal{A} \cap B_1} |\nabla^2 F|, \quad \int_{\mathcal{A} \cap B_1} |F|^2 \quad \le \sum_{j=1}^{\infty} (\epsilon')^j \le 2\epsilon'. \tag{8.17}$$

Letting $\epsilon' < \frac{1}{2}\epsilon$ this concludes the proof of our Theorem. $\qquad\square$

9. QUANTITATIVE BUBBLE TREE AND QUANTITATIVE ENERGY IDENTITY

In this section we discuss our quantitative bubble tree and quantitative energy identity theorems. As the names suggest, these are both quantitative and higher dimensional versions of the more classical notions from dimension four. Both the classical bubble tree and energy identity's are traditionally discussed for limiting sequences of Yang Mills connections, however the quantitative versions hold on a fixed Yang Mills connection, and so from this point of view the results are new even for dimension four. We will see that the quantitative nature of the estimates are crucial for the applications.

Let us begin with our quantitative bubble tree decomposition:

Theorem 9.1 (Quantitative Bubble Tree). *Let A be a stationary connection with $\fint_{B_4} |F_A|^2 \le \Lambda$. If $B_4(p)$ is δ'-weakly flat wrt $\mathcal{L}$ for $\delta' < \delta'(n, k, \Lambda, \delta, \epsilon)$, then we have*

$$B_1(p) \subseteq \bigcup_a (\mathcal{A}_a \cap B_{r_a}) \cup \bigcup_b \mathcal{B}_b \cup \bigcup_c B_{r_c}(x_c), \tag{9.1}$$

such that

   (a) $\mathcal{A}_a \subseteq B_{2r_a}(x_a)$ *are δ-annular regions with respect to $\mathcal{L}$.*
   (b) $\mathcal{B}_b \subseteq B_{\delta^{-1}r_b}(x_b)$ *are δ-bubble regions with respect to $\mathcal{L}$.*
   (c) $\sum_a r_a^{n-4} + \sum_b r_b^{n-4} \le C(n, k, \Lambda)$ *and* $\sum_c r_c^{n-4} < \epsilon$.

To state the quantitative energy identity will pick points $q \in \mathcal{L}$ on the plane of symmetry and consider the slice $\mathcal{L}_q^\perp$. One can view the first part of the next result as a sliced version of the quantitative bubble tree theorem. The final part of the next theorem contains the real content of the quantitative energy identity, and tells one how to compute the energy at a point by summing energies of bubbles.

Theorem 9.2 (Quantitative Energy Identity). *Under the conditions and decomposition of Theorem 9.1 if $\delta < \delta(n, k, \Lambda, \epsilon)$ and $\delta' < \delta'(n, k, \Lambda, \delta)$ then $\exists \, \mathcal{G}_\epsilon \subseteq \mathcal{L} \cap B_1$ with $\mathrm{Vol}(B_1 \setminus \mathcal{G}_\epsilon) < \epsilon$ such that for each $q \in \mathcal{G}_\epsilon$ the covering*

$$\mathcal{L}_q^\perp \cap B_1(q) \subseteq \mathcal{A}_q \cup \mathcal{B}_q = \bigcup_a \mathcal{A}_{q,a} \cup \bigcup_b \mathcal{B}_{q,b}$$

$$= \bigcup_a \left(\mathcal{L}_q^\perp \cap \mathcal{A}_a \cap B_{r_a} \right) \cup \bigcup_b \left(\mathcal{L}_q^\perp \cap \mathcal{B}_b \cap B_{r_b} \right), \tag{9.2}$$

satisfies

(a) $\int_{\mathcal{A}_q} |F_A|^2 < \epsilon$.

(b) $\left| \int_{B_1(q)} |F_A|^2 - \omega_{n-4} \int_{\mathcal{B}_q} |F_A|^2 \right| < \epsilon$.

(s) $\#\{\mathcal{A}_{q,a}\}, \ \#\{\mathcal{B}_{q,b}\} < N(n,k,\Lambda)$.

9.1. **Proof of the Quantitative Bubble Tree Decomposition and the Quantitative Energy Identity.** In this subsection we will prove both Theorem 9.1 and Theorem 9.2 simultaneously. Let us first pick $\epsilon' > 0$ which will be fixed later. Now since the ball $B_4(p)$ is δ'-weakly flat we have the estimate

$$\fint_{B_4} |F_A[\mathcal{L}]|^2 \leq \delta', \tag{9.3}$$

where we will choose δ' sufficiently small later on in the proof. Now for every point $x \in B_2$ let us define the radius

$$m_x \equiv \inf\{0 < r < 1 : \ r^{4-n} \int_{B_r(x)} |F[\mathcal{L}]|^2 \leq \delta''\}, \tag{9.4}$$

where $\delta'' = \delta''(n,k,\Lambda,\delta)$ is chosen so that Theorem 4.2, Theorem 5.4 and Theorem 5.3 hold with $\epsilon' > 0$. Additionally, let us consider the set

$$\mathcal{M} \equiv \{x \in B_2 : m_x > 0\}. \tag{9.5}$$

Note for each ball $B_{m_x}(x)$ with $m_x > 0$ we have the estimate

$$\int_{B_{m_x}(x)} |F[\mathcal{L}]|^2 = \delta'' m_x^{n-4}. \tag{9.6}$$

Now let us consider the covering $\{B_{10m_x}(x)\}_{x \in \mathcal{M}}$ of $\mathcal{M}$, and then we can take a Vitali subcovering

$$\mathcal{M} \subseteq \bigcup_c B_{r_{c,0}}(x_{c,0}) \equiv \bigcup_s B_{r_s}(x_s), \tag{9.7}$$

where $r_s = 10m_{x_s}$ and $\{B_{r_s/10}(x_s)\}$ are disjoint. We give this collection two names because the remaining c-balls which will built later in this construction will be done by different means with a different purpose, so that although in the end we will group them altogether we keep them distinct for now for intuition simplicity. Note now that we have the estimate

$$\sum_s r_s^{n-4} = 10^{n-4} \sum_s \left(\frac{r_s}{10}\right)^{n-4} = 10^{n-4} \delta''^{-1} \sum_s \int_{B_{r_s/10}(x_s)} |F[\mathcal{L}]|^2$$

$$\leq 10^{n-4} \delta''^{-1} \int_{B_3} |F[\mathcal{L}]|^2 = 10^{n-4} (\delta'')^{-1} \delta' < \delta, \tag{9.8}$$

where in the last line we have chosen $\delta' < \delta'(n,k,\Lambda,\delta)$.

In particular, we now know that for every $x \in B_2 \setminus \bigcup_s B_{r_s}(x_s)$ and all $0 < r < 2$ that we have the estimate

$$r^{4-n} \int_{B_r(x)} |F[\mathcal{L}]|^2 < \delta''. \tag{9.9}$$

Intuitively, using Theorem 4.2 and Theorem 5.4 this is telling us that every ball not strictly contained in $\bigcup_s B_{r_s}(x_s)$ is either a bubble region or an annular region. Finishing the proof of Theorem 9.1 is about making this intuition precise and keeping track of the estimates along the way.

To make this precise we will build a sequence of coverings. If $E \equiv \sup_{B_1(p)} \theta_1 = \bar{\theta}_1(p)$, then the ith covering of the sequence looks like

$$B_1(p) \setminus \bigcup_c B_{r_{c,i}}(x_{c,i}) \subseteq \bigcup_a (\mathcal{A}_{a,i} \cap B_{r_{a,i}}(x_{a,i})) \cup \bigcup_b \mathcal{B}_{b,i} \cup \bigcup_w B_{r_{w,i}}(x_{w,i}), \tag{9.10}$$

and will satisfy

(a) $\mathcal{A}_{a,i} \subseteq B_{r_{a,i}}(x_{a,i})$ are δ-annular with $\int_{\mathcal{L}_q^\perp \cap \mathcal{A}_i} |F_A|^2 < C(n,k,\Lambda,i)\epsilon'$ for $q \in \mathcal{L} \cap B_1(p) \setminus \bigcup_c B_{r_{c,i}}(x_{c,i})$.

(b) $\mathcal{B}_{b,i} \subseteq B_{\delta^{-1} r_{b,i}}(x_{b,i})$ are δ-bubble regions.

(c) For $q \in \mathcal{L}$ with $\mathcal{L}_q^\perp \cap \bigcup_c B_{r_{c,i}}(x_{c,i}) = \emptyset$ if $\mathcal{B}_q \equiv \mathcal{L}_q^\perp \cap \bigcup_b \mathcal{B}_{b,i}$ then there exists $\{x'_{w,i}\} \subseteq \{x_{w,i}\}$ with $B_{r'_{w,i}}(x'_{w,i}) \cap \mathcal{L}_q^\perp \neq \emptyset$ such that $\left| \theta_1(q) - \omega_{n-4} \int_{\mathcal{B}_q} |F_A|^2 - \sum \theta_{r_{w,i}}(x'_{w,i}) \right| < C(n,k,\Lambda,i)\epsilon'$.

(w) $B_r(x_{w,i})$ are δ-weakly flat wrt $\mathcal{L}$ for $4^{-1}\delta r_{w,i} \leq r \leq 4 r_{w,i}$ and satisfy $\bar{\theta}_{r_{w,i}}(x_{w,i}) \leq E - \epsilon_{n,k} \cdot i$.

(s) $\sum_a r_{a,i}^{n-4} + \sum_b r_{b,i}^{n-4} + \sum_w r_{w,i}^{n-4} < C(n,k,\Lambda,i)$ and $\sum_c r_{c,i}^{n-4} \leq C(n,k,\Lambda,i)\epsilon'$.

Let us begin by observing that if we can build this sequence of coverings then we have proved the Theorems. Indeed, for $i > \Lambda \epsilon_{n,k}^{-1} = i(n,k,\Lambda)$ we have that there are no w-balls in the covering, as their energy would necessarily be negative. However if $\epsilon' < c(n,k,\Lambda)\epsilon$ it is then clear from $(a) \rightarrow (s)$ that the covering in (9.10) would then satisfy the conditions of Theorems 9.1 and 9.2.

Our inductive covering will itself follow from a series of covering constructions. Let us begin with the following claim, which is the first step in the process:

Claim: Let $B_s(x)$ be δ-weakly flat wrt $\mathcal{L}$ for $4^{-1}\delta r \leq s \leq 4r$, then we have the decomposition $B_r(x) \setminus \bigcup_s B_{r_s}(x_s) \subseteq \mathcal{A} \cup \bigcup_b \mathcal{B}_b \cup \bigcup_c B_{r_c}(x_c) \cup \bigcup_w B_{r_w}(x_w)$ such that

(a) $\mathcal{A} \subseteq B_{2r}(x)$ is a δ-annular region with $\int_{\mathcal{L}_q^\perp \cap \mathcal{A}} |F_A|^2 < \epsilon'$ for $q \in \mathcal{L} \cap B_r(x) \setminus \bigcup_c B_{r_c}(x_c)$.

(b) $\mathcal{B}_b \subseteq B_{\delta^{-1} r_b}(x_b)$ are δ-bubble regions.

(c) For $q \in \mathcal{L}$ with $\mathcal{L}_q^\perp \cap \left(\bigcup_c B_{r_c}(x_c) \cup \bigcup_s B_{r_s}(x_s) \right) = \emptyset$ and $\mathcal{B}_q \equiv \mathcal{L}_q^\perp \cap \bigcup_b \mathcal{B}_b$, there exists $\{x'_w\} \subseteq \{x_w\}$ with $B_{r'_w}(x'_w) \cap \mathcal{L}_q^\perp \neq \emptyset$ such that $\left| \theta_r(q) - \omega_{n-4} \int_{\mathcal{B}_q} |F_A|^2 - \theta_{r_w}(x'_w) \right| < \epsilon'$.

(w) $B_r(x_w)$ is δ-weakly flat wrt $\mathcal{L}$ for $4^{-1}\delta r_w \leq r \leq 4 r_w$ and satisfies $\bar{\theta}_{r_w}(x_w) \leq \bar{\theta}_r(x) - \epsilon_{n,k}$.

(s) $\sum_b r_b^{n-4} + \sum_w r_w^{n-4} < C(n,k,\Lambda) r^{n-4}$ and $\sum_c r_c^{n-4} < \epsilon' r^{n-4}$.

To prove the claim will require an application of Theorem 5.4 and Theorem 4.2. Indeed, let us first apply Theorem 5.4 in order to build an annular region $\mathcal{A} = B_r(x) \setminus \bar{B}_{r_x}(\mathcal{C})$. Using the estimates on $\mathcal{C}^c$ we can cover

$$B_r(x) \setminus \bigcup_s B_{r_s}(x_s) \subseteq \mathcal{A} \cup \bigcup_{b'} B_{r_{b'}}(x_{b'}) \cup \bigcup_c B_{r_c}(x_c), \tag{9.11}$$

where $B_{r_{b'}}(x_{b'})$ are δ weakly flat but $B_s(x_{b'})$ is not δ-weakly flat for some $\delta^4 r_{b'} \leq s \leq r_{b'}$. We also have the estimates $\sum r_{b'}^{n-4} < C(n,k,\Lambda) r^{n-4}$ and $\sum r_c^{n-4} < \epsilon' r^{n-4}$. To each b'-ball we apply Theorem 4.2 in order to

cover

$$B_{r_{b'}} \subseteq \mathcal{B}_b \cup \bigcup_w B_{r_{b,w}}(x_{b,w}), \tag{9.12}$$

where with $r_b \equiv \delta r_{b'}$ we have that $\mathcal{B}_b \subseteq B_{\delta^{-1}r_b}(x_b)$ is a δ-bubble region, and $B_s(x_{b,w})$ are δ-weakly flat for $\delta r_{b,w} < s < r_{b,w}$ with $\bar{\theta}_{r_{b,w}}(x_{b,w}) < \bar{\theta}_{\delta^{-1}r_b}(x_b) - \epsilon_{n,k} < \bar{\theta}_r(x) - \epsilon_{n,k}$ and $\sum_w r_{b,w}^{n-4} \leq C(n,k,\Lambda)r_b^{n-4}$. By taking a union over all b' we obtain the covering

$$B_r(x) \setminus \bigcup_s B_{r_s}(x_s) \subseteq \mathcal{A} \cup \bigcup_b \mathcal{B}_b \cup \bigcup_w B_{r_w}(x_w) \cup \bigcup_e B_{r_e}(x_e), \tag{9.13}$$

which satisfy the estimates $\sum r_b^{n-4} + \sum_w r_w^{n-4} \leq C(n,k,\Lambda)r^{n-4}$ and $\sum r_c^{n-4} < \epsilon' r^{n-4}$. Condition (a) of the claim follows from the definition of $\mathcal{C}^c$ in Theorem 5.4, while condition (c) follows by combining this with (2) of Theorem 4.2, and thus we have finished the proof of the Claim. $\square$

With the Claim in hand, we are ready to finish the construction of the inductive covering of (9.10), which will itself finish the proof. Note first that for $i = 0$ we may take the trivial covering $B_1(p) = B_{r_{w,0}}(x_{w,0})$, where $x_{w,0} \equiv p$ and $r_{w,0} \equiv 1$. Thus we have the base step of the inductive construction. Now having built the covering at stage i, let us see how to build the covering at stage $i + 1$. More precisely, observe that for each ball $B_{r_{w,i}}(x_{w,i})$ we may apply the Claim. If we do this to each w-ball then we arrive at a new covering

$$B_1(p) \setminus \bigcup_s B_{r_s}(x_s) \subseteq \bigcup_a (\mathcal{A}_{a,i+1} \cap B_{r_{a,i}}(x_{a,i+1})) \cup \bigcup_b \mathcal{B}_{b,i+1} \cup \bigcup_c B_{r_{c,i+1}}(x_{c,i+1}) \cup \bigcup_w B_{r_{w,i+1}}(x_{w,i+1}), \tag{9.14}$$

where $\bar{\theta}_{r_{w,i+1}}(x_{w,i+1}) < \bar{\theta}_{r_{w,i}}(x_{w,i}) - \epsilon_{n,k} < E - \epsilon_{n,k}(i+1)$ and we have the estimates

$$\sum_a r_{a,i+1}^{n-4} + \sum_b r_{b,i+1}^{n-4} + \sum_w r_{w,i+1}^{n-4} \leq C(n,k,\Lambda,i) + C(n,k,\Lambda) \sum_w r_{w,i}^{n-4} \leq C(n,k,\Lambda,i+1),$$

$$\sum_c r_{c,i+1}^{n-4} \leq C(n,k,\Lambda,i)\epsilon' + \epsilon' \sum_w r_{w,i}^{n-4} \leq C(n,k,\Lambda,i+1)\epsilon'. \tag{9.15}$$

Conditions (a), (b), (w), and now (s) are all clearly satisfied by the new covering, while (c) follows from the (i)-inductive hypothesis of (c) combined with (c) of the Claim. This finishes the proof of the inductive covering, and hence of Theorem 9.1 and Theorem 9.2. $\square$

10. PROOF OF ENERGY IDENTITY

In this section we use the quantitative energy identity from Theorem 9.2 in order to finish the proof of the energy identity itself. Thus recall our setup that $A_i \to A$ with $|F_{A_i}|^2 dv_{g_i} \to |F_A|^2 dv_g + \nu$, where $\nu = e(x)\lambda_S^{n-4}$ is the $n - 4$ rectifiable defect measure. For each $\epsilon, \delta > 0$ let us consider the subset $E_{\epsilon,\delta} \subseteq \text{supp}[\nu]$ defined by $x \in E_{\epsilon,\delta}$ if there exists a $n - 4$ plane $\mathcal{L}_x$ through x and points $q_i \to x$ such that the following hold:

(1) $\exists \delta$-bubbles $\mathcal{B}_{q_i,\mathbf{b}} \subseteq B_{\delta^{-1}r_{q_i,b}}(x_{q_i,b})$ with $x_{q_i,b} \in \mathcal{L}_{q_i}^{\perp}$
(2) $\#\{\mathcal{B}_{q_i,\mathbf{b}}\} \leq N(n,\Lambda)$.
(3) If $\mathcal{B}_{q_i} \equiv \bigcup \mathcal{L}_{q_i}^{\perp} \cap \mathcal{B}_{q_i,\mathbf{b}}$ then $\left| \int_{\mathcal{B}_{q_i}} |F_{A_i}|^2 - e(x) \right| < \epsilon$.

Our goal is to show that for each $\epsilon > 0$ there exists $\delta_\epsilon > 0$ such that $E_{\epsilon,\delta}$ is a set of full measure in supp$[\nu]$. Let us first observe that if this is the case then we have finished the proof of the energy identity itself. Indeed, consider the collection of points $E_0 \equiv \bigcap_{\epsilon>0} E_{\epsilon,\delta_\epsilon}$, which is itself a set of full measure. Then for such a point $x \in E_0$ we can find a sequence $q_i \to x$ with $\left| \int_{\mathcal{B}_{q_i}} |F_A|^2 - e(x) \right| \to 0$. Since $\#\{\mathcal{B}_{q_i,b}\} \leq N$ we can apply Theorem 4.3.3 in order to find points $c_{i,b} \in \mathcal{B}_{q_i}$ and radii $s_{i,b} > 0$ such that for all $\eta > 0$ and $R \geq R(n, k, \Lambda, \eta)$ we have

$$\left| \int_{\mathcal{B}_{q_i}} |F_{A_i}|^2 - \int_{\mathcal{B}_{q_i} \cap \bigcup B_{Rs_{i,b}}(x_{i,b})} |F_{A_i}|^2 \right| < \eta, \tag{10.1}$$

so that

$$\left| e(x) - \int_{\mathcal{B}_{q_i} \cap \bigcup B_{Rs_{i,b}}(x_{i,b})} |F_{A_i}|^2 \right| < \eta. \tag{10.2}$$

Indeed, the pointed limits $s_{i,b}^{-1} A_i \to B_b \in \mathcal{B}[x]$ are then bubbles at $x \in E_0$ such that the energy identity

$$e(x) = \sum_b \int_{\mathcal{L}^\perp} |F_{B_b}|^2, \tag{10.3}$$

holds. Since E_0 is a set of full measure this finishes the proof of the Theorem.

Thus we need to show that each $E_{\epsilon,\delta}$ is a set of full measure. To accomplish this let us remark that by the definition of ν being rectifiable we have that a.e. $x \in \text{supp}\{\nu\}$ is such that the unique tangent measure at x is $e(x)\lambda_{\mathcal{L}}^{n-4}$, for some $n - 4$ plane $\mathcal{L}$. In particular, for any $\delta' > 0$ and all $0 < r < r(x, \delta')$ fixed sufficiently small that with i sufficiently large we have

$$\left| r^{4-n} \int_{B_r(x)} |F_{A_i}|^2 - \omega_{n-4} e(x) \right| < \delta',$$

$$B_{4r}(x) \text{ is } \delta'\text{-weakly flat.} \tag{10.4}$$

Let us choose $\delta' > 0$ sufficiently small that we can apply Theorem 9.2 with $\frac{1}{2}\epsilon$ for some $\delta \leq \delta(n, k, \Lambda, \epsilon)$. Let us consider the sets $\mathcal{G}_{\epsilon,i} \subseteq \mathcal{L} \cap B_r(x)$ coming from the Theorem and consider their Hausdorff limit $\mathcal{G}_{\epsilon,i} \to \mathcal{G}_\epsilon \subseteq \mathcal{L} \cap B_r(x)$. Using (10.4) above and the content of Theorem 9.2 it is immediate that $\mathcal{G}_\epsilon \subseteq E_{\epsilon,\delta}$. Since $0 < r < r(x, \delta')$ was arbitrary, we have in particular that x is a ϵ-density point of $E_{\epsilon,\delta}$. Since $x \in \text{supp}[\nu]$ was a set of full measure, we have that $E_{\epsilon,\delta} \subseteq \text{supp}[\nu]$ is a set of full measure as well. This completes the proof. $\square$

11. ANNULUS/BUBBLE DECOMPOSITION

In this section we introduce one last quantitative covering result, which we call the annulus/bubble decomposition. The decomposition will split a ball into two pieces, one of which are bubble regions with uniformly bounded curvature, and the other are annulus regions. The most important aspect of this decomposition for us will be the effective $n - 4$ content estimates that come with the covering, which will be used quite crucially in the proof of our global estimates on the hessian. Our main result in this subsection is the following:

Theorem 11.1 (Annulus/Bubble Decomposition). *Let A be a stationary Yang-Mills connection satisfying* (1.11) *and* $\fint_{B_2} |F_A|^2 \leq \Lambda$. *Then for each* $\delta > 0$ *we can write*

$$B_1(p) \subseteq \bigcup_a (\mathcal{A}_a \cap B_{r_a}) \cup \bigcup_b \mathcal{B}_b, \tag{11.1}$$

such that

    (a) $\mathcal{A}_a \subseteq B_{2r_a}(x_a)$ *are* δ-*annular regions.*
    (b) $\mathcal{B}_b \subseteq B_{\delta^{-1}r_b}(x_b)$ *are* δ-*bubble regions.*
    (s) $\sum_a r_a^{n-4} + \sum_b r_b^{n-4} \leq C(n, k, K, \Lambda, \delta).$

Remark 11.1. If the connection A is singular in the sense of [TT04] or [RPc] then one must allow an extra piece to this decomposition, so that $B_1(p) \subseteq \bigcup_a (\mathcal{A}_a \cap B_{r_a}) \cup \bigcup_b \mathcal{B}_b \cup \mathcal{S}$, where $\mathcal{S}$ is a closed set with $n-4$ measure zero. We shall point out in the proof where this happens.

Note several major differences between the annulus/bubble and the quantitative bubble tree decomposition. On the positive side, the original ball B_1 does not need to be weakly flat, and there are no c-balls in the covering. Therefore every point lies in either an annular region or a bubble region. On the negative side, the annular and bubble regions are not with respect to some fixed $\mathcal{L}$. This decomposition will be the more appropriate one in order to prove the L^1 hessian estimate, while as we have seen the quantitative bubble tree decomposition is the appropriate one to prove the energy identity.

11.1. **Proof of Theorem 11.1.** The proof of Theorem 11.1 is by recursively applying the quantitative bubble tree decomposition of Theorem 9.1 with the weakly flat decomposition of Theorem 3.4. Let us begin with the following claim:

Claim 1: Let A be a stationary Yang-Mills connection satisfying (1.11) and $\fint_{B_4} |F_A|^2 \leq \Lambda$. Then for each $\delta > 0$ we can write

$$B_1(p) \subseteq \bigcup_a (\mathcal{A}_a \cap B_{r_a}) \cup \bigcup_b \mathcal{B}_b \cup \bigcup_d B_{r_d}(x_d), \tag{11.2}$$

such that

    (a) $\mathcal{A}_a \subseteq B_{2r_a}(x_a)$ are δ-annular regions.
    (b) $\mathcal{B}_b \subseteq B_{\delta^{-1}r_b}(x_b)$ are δ-bubble regions.
    (d) $B_{r_d}(x_d)$ satisfy $\bar{\theta}_{r_d}(x_d) \leq \bar{\theta}_1(x) - \eta(n, k, K, \Lambda, \delta).$
    (s) $\sum_a r_a^{n-4} + \sum_b r_b^{n-4} + \sum_d r_d^{n-4} \leq C(n, k, \Lambda, \delta).$

To prove the Claim let us pick $\delta'(n, k, \Lambda, \delta)$ so that the quantitative bubble tree decomposition of Theorem 9.1 holds with $\epsilon = 10^{-6n}\omega_n$. We will produce a sequence of coverings of the form

$$B_1(p) \subseteq \bigcup_a (\mathcal{A}_{a,i} \cap B_{r_{b,i}}) \cup \bigcup_b \mathcal{B}_{b,i} \cup \bigcup_c B_{r_{c,i}}(x_{c,i}) \cup \bigcup_d B_{r_{d,i}}(x_{d,i}), \tag{11.3}$$

such that in addition to (a), (b), and (d) holding we will also have the estimates

    (c-i) $\sum_c r_{c,i}^{n-4} \leq \epsilon^i.$
    (s-i) $\sum_a r_{a,i}^{n-4} + \sum_b r_{b,i}^{n-4} + \sum_d r_{d,i}^{n-4} \leq C(n, k, \Lambda, \delta) \sum_{j=0}^{i-1} \epsilon^j.$

Let us first observe that if we can build this sequence of coverings then we have proved the Claim. Indeed, this sequence differs from our desired covering in the claim only by the existence of c-balls. However, by the content estimate we have that $\sup_c r_{c,i} \to 0$ as $i \to \infty$. In particular, for i sufficiently (but uncontrollably) large we have that any c-ball must define a smooth bubble region[1], and therefore we can take the c-balls to be empty in the covering for i sufficiently large, which produces the covering of our claim.

Thus let us now focus on building this sequence. Note that for $i = 0$ we can let the covering be defined by the single ball $B_1(p) \equiv B_{r_c}(x_c)$. Therefore our goal will be to produce the covering at the $i + 1$ stage given that we have produced the covering at stage i.

To accomplish this let us focus on each c ball $B_{r_{c,i}}(x_{c,i})$ in the covering. Let us now consider two options, either the ball $B_{4r_{c,i}}$ is δ'-weakly flat or it is not. In the first case we may apply the quantitative bubble tree decomposition of Theorem 9.1 in order to produce the covering:

$$B_{r_{c,i}}(x_{c,i}) \subseteq \bigcup_a (\mathcal{A}_{c,i,a'} \cap B_{r_{c,i,a'}}) \cup \bigcup_b \mathcal{B}_{c,i,b'} \cup \bigcup_c B_{r_{c,i,c'}}(x_{c,i,c'}) . \tag{11.4}$$

which satisfy the estimates $\sum_{c'} r_{c,i,c'}^{n-4} \le \epsilon r_{c,i}^{n-4}$ and $\sum_{a'} r_{c,i,a'}^{n-4} + \sum_{b'} r_{c,i,b'}^{n-4} \le C(n,k,\Lambda,\delta) r_{c,i}^{n-4}$. On the other hand, if $B_{4r_{c,i}}$ is not δ'-weakly flat then we may apply Theorem 3.4 in order to write

$$B_{r_{c,i}}(x_{c,i}) \subseteq \bigcup_c B_{r_{c,i,c'}}(x_{c,i,c'}) \cup \bigcup_d B_{r_{c,i,d'}}(x_{c,i,d'}) . \tag{11.5}$$

with the estimates $\sum_{c'} r_{c,i,c'}^{n-4} \le \epsilon r_{c,i}^{n-4}$ and $\sum_{d'} r_{c,i,d'}^{n-4} \le C(n,k,\Lambda,\delta) r_{c,i}^{n-4}$. If we take the union of all these coverings of every c-ball, and collect different pieces together, we arrive at the $i + 1$ covering

$$B_1(p) \subseteq \bigcup_a (\mathcal{A}_{a,i+1} \cap B_{r_{b,i+1}}) \cup \bigcup_b \mathcal{B}_{b,i+1} \cup \bigcup_c B_{r_{c,i}}(x_{c,i+1}) \cup \bigcup_d B_{r_{d,i+1}}(x_{d,i+1}) , \tag{11.6}$$

with the desired estimates

$$\sum_c r_{c,i+1}^{n-4} \le \epsilon \sum_c r_{c,i}^{n-4} \le \epsilon^{i+1} ,$$

$$\sum_a r_{a,i+1}^{n-4} + \sum_b r_{b,i+1}^{n-4} + \sum_d r_{d,i+1}^{n-4} \le \sum_a r_{a,i}^{n-4} + \sum_b r_{b,i}^{n-4} + \sum_d r_{d,i}^{n-4} + C(n,k,K,\Lambda) \sum_c r_{c,i}^{n-4}$$

$$\le C(n,k,\Lambda,\delta) \sum_{j=0}^{i-1} \epsilon^j + C\epsilon^i = C \sum_{j=0}^i \epsilon^j , \tag{11.7}$$

which therefore finishes the inductive step of the construction and hence the Claim. □

With the Claim in hand let us now return to finish the proof of the Theorem. Indeed, we will see that the proof of the Theorem is just a repeated application of the Claim. More precisely, let us produce a sequence of coverings

$$B_1(p) \subseteq \bigcup_a (\mathcal{A}_{a,i} \cap B_{r_{b,i}}) \cup \bigcup_b \mathcal{B}_{b,i} \cup \bigcup_d B_{r_{d,i}}(x_{d,i}) , \tag{11.8}$$

which in addition to satisfying (a) and (b) will satisfy the conditions:

(d-i) $B_{r_d}(x_d)$ satisfy $\bar{\theta}_{r_d}(x_d) \le \bar{\theta}_1(x) - \eta(n,k,K,\Lambda,\delta) i$.

---

[1]This is not true if A is not smooth. However, the estimate $(c - i)$ allows one to hausdorff limit the c-balls $\{B_{r_{c,i}}(x_{c,i})\}$ to a closed $n - 4$ measure zero set $\mathcal{S}$, which is the additional piece of the covering in the singular case.

(s-i) $\sum_a r_{a,i}^{n-4} + \sum_b r_{b,i}^{n-4} + \sum_d r_{d,i}^{n-4} \le C(n,k,\Lambda,\delta,i)$.

Let us first observe that once we have proved the existence of this sequence of coverings then we are done. In fact, for $i > \eta^{-1}\Lambda$ there cannot be any d-balls in the covering as any such ball would have negative energy. Therefore for such an i we have produced the desired covering of the Theorem.

Thus let us concentrate on proving the existence of this sequence of coverings. To produce the covering for $i = 1$ we simply apply the Claim. We now construct the covering at the $i + 1$ stage inductively, using that we have already built the covering at the ith stage. To accomplish this let us focus on each d-ball $B_{r_{d,i}}(x_{d,i})$ in the covering and apply the Claim to this ball. Taking a union produces the $i + 1$ covering. Hence, we have finished the proof of Theorem 11.1.

12. PROOF OF THEOREM 1.2

In this section we put together the annulus/bubble decomposition of Theorem 11.1 with the annulus structure of Theorem 5.3 and the bubble structure of Theorem 4.3 in order to complete the proof of the L^1 hessian estimate on the curvature. Indeed, let us choose $\delta(n,k,\Lambda)$ sufficiently small so that Theorem 5.3 holds for $\epsilon_n = 10^{-n}$. Then we can estimate our L^1 norm of the hessian by

$$\int_{B_1(p)} |\nabla^2 F| \le \sum_a \int_{\mathcal{A}_a \cap B_{r_a}} |\nabla^2 F| + \sum_b \int_{\mathcal{B}_b} |\nabla^2 F|. \tag{12.1}$$

On the δ-bubble regions we have by Theorem 4.3 that

$$\int_{\mathcal{B}_b} |\nabla^2 F| \le C(n,k,\Lambda,\delta)\, r_b^{n-4}. \tag{12.2}$$

On the other hand, by Theorem 5.3 we have on the annular regions $\mathcal{A}_a$ the scale invariant estimate

$$\int_{\mathcal{A}_a \cap B_{r_a}} |\nabla^2 F| \le \epsilon_n\, r_a^{n-4}. \tag{12.3}$$

Putting these together with (12.1) we obtain

$$\int_{B_1(p)} |\nabla^2 F| \le \epsilon_n \sum_a r_a^{n-4} + C\delta \sum_b r_b^{n-4} \le C(n,k,\Lambda)\Big(\sum_a r_a^{n-4} + \sum_b r_b^{n-4}\Big) \le C(n,k,\Lambda), \tag{12.4}$$

where the last inequality is due to the $n-4$ content estimate of Theorem 11.1. This completes the proof.[2] □

REFERENCES

[CN13] J. CHEEGER and A. NABER, Quantitative stratification and the regularity of harmonic maps and minimal currents, *Comm. Pure Appl. Math.* **66** (2013), 965–990. MR 3043387. Zbl 1269.53063. http://dx.doi.org/10.1002/cpa.21446. Available at http://arxiv.org/abs/1107.3097.

[CN15] J. CHEEGER and A. NABER, Regularity of Einstein manifolds and the codimension 4 conjecture, *Ann. of Math. (2)* **182** (2015), 1093–1165. MR 3418535. http://dx.doi.org/10.4007/annals.2015.182.3.5. Available at http://dx.doi.org/10.4007/annals.2015.182.3.5.

---

[2]If the connection A is singular in the sense of [TT04] or [RPc] one must also show $\nabla^2 f$ is measurable as a distribution. This may be done using that the singular set has $n-4$ measure zero together with the monotonicity formula.

[Don89]   S. K. DONALDSON, Compactification and completion of Yang-Mills moduli spaces, in *Differential geometry (Peñíscola, 1988)*, *Lecture Notes in Math.* **1410**, Springer, Berlin, 1989, pp. 145–160. MR 1034277. http://dx.doi.org/10.1007/BFb0086420. Available at http://dx.doi.org/10.1007/BFb0086420.

[DT98]   S. K. DONALDSON and R. P. THOMAS, Gauge theory in higher dimensions, in *The geometric universe (Oxford, 1996)*, Oxford Univ. Press, Oxford, 1998, pp. 31–47. MR 1634503.

[JN]   W. JIANG and A. NABER, L2 curvature bounds on manifolds with bounded ricci curvature, *preprint*. Available at http://arxiv.org/abs/1605.05583.

[LR02]   F.-H. LIN and T. RIVIÈRE, Energy quantization for harmonic maps, *Duke Math. J.* **111** (2002), 177–193. MR 1876445. http://dx.doi.org/10.1215/S0012-7094-02-11116-8. Available at http://dx.doi.org/10.1215/S0012-7094-02-11116-8.

[NV]   A. NABER and D. VALTORTA, Rectifiable-reifenberg and the regularity of stationary and minimizing harmonic maps, *accepted on Annals of Mathematics*. Available at http://annals.math.princeton.edu/articles/10010.

[Pri83]   P. PRICE, A monotonicity formula for Yang-Mills fields, *Manuscripta Math.* **43** (1983), 131–166. MR 707042. http://dx.doi.org/10.1007/BF01165828. Available at http://dx.doi.org/10.1007/BF01165828.

[Riv02]   T. RIVIÈRE, Interpolation spaces and energy quantization for Yang-Mills fields, *Comm. Anal. Geom.* **10** (2002), 683–708. MR 1925499. http://dx.doi.org/10.4310/CAG.2002.v10.n4.a2. Available at http://dx.doi.org/10.4310/CAG.2002.v10.n4.a2.

[RPa]   T. RIVIÈRE and M. PETRACHE, Partial regularity of yang-mills stationary weak connections in high dimensions, *preprint*.

[RPb]   T. RIVIÈRE and M. PETRACHE, The resolution of the yang-mills plateau problem in super-critical dimensions, *preprint*. Available at http://cvgmt.sns.it/paper/2177/.

[RPc]   T. RIVIÈRE and M. PETRACHE, The space of weak connections in high dimensions, *preprint*.

[TT04]   T. TAO and G. TIAN, A singularity removal theorem for Yang-Mills fields in higher dimensions, *J. Amer. Math. Soc.* **17** (2004), 557–593. MR 2053951. http://dx.doi.org/10.1090/S0894-0347-04-00457-6. Available at http://dx.doi.org/10.1090/S0894-0347-04-00457-6.

[Tia00]   G. TIAN, Gauge theory and calibrated geometry. I, *Ann. of Math. (2)* **151** (2000), 193–268. MR 1745014. http://dx.doi.org/10.2307/121116. Available at http://dx.doi.org/10.2307/121116.

[Uhl82]   K. K. UHLENBECK, Removable singularities in Yang-Mills fields, *Comm. Math. Phys.* **83** (1982), 11–29. MR 648355. Available at http://projecteuclid.org/euclid.cmp/1103920742.

[Yu]   W. YU, Quantitative stratification of stationary yang-mills fields, *preprint*.