CANONICAL METRICS ON 3-MANIFOLDS AND 4-MANIFOLDS

MICHAEL T. ANDERSON

In Memory of S.S. Chern

1. Introduction.

In this paper, we discuss recent progress on the existence of canonical metrics on manifolds in dimensions 3 and 4, and the structure of moduli spaces of such metrics. The existence of a "best possible" metric on a given closed manifold is a classical question in Riemannian geometry, attributed variously to H. Hopf and R. Thom, see [22] for an interesting perspective. A good deal of motivation for this question comes from the case of surfaces; the uniformization theorem in dimension 2 has a multitude of consequences in mathematics and physics. Further, there are strong reasons showing that the closest relations between geometry and topology occur in dimensions 2, 3 and 4.

The precise formulation of the question in dimension 3 is given by Thurston's Geometrization Conjecture. This conjecture describes completely when a given 3-manifold admits a canonical metric (defined to be a metric of constant curvature or more generally a locally homogeneous metric), and thus determines exactly what the obstructions are to the existence of such a metric. Moreover, it describes how an arbitrary 3-manifold decomposes into topologically essential pieces, each of which admits a canonical metric, resulting in the topological classification of 3-manifolds. The apparent solution of the Geometrization Conjecture by Perelman is one of the most spectacular breakthroughs in geometry and topology in the past several decades.

The Thurston picture will be reviewed in more detail in §2, for the light it sheds on what might be hoped for or expected in dimension 4. Since there is already considerable analysis and discussion of the details of Perelman's work elsewhere, we will not discuss this in any detail here. We do however give one application of his work, (since this does not seem to be widely known), namely the determination of the value of the Yamabe invariant or Sigma constant $\sigma(M)$ of all 3-manifolds M for which $\sigma(M) \leq 0$, cf. §2.

Thus, the bulk of the paper concerns dimension 4. Canonical metrics will be defined to be metrics minimizing, (or possibly just critical points for), one of the classical and natural curvature functionals $\mathcal{F}$ on the space of metrics $\mathbb{M}$ on a given oriented 4-manifold M:

$$\mathcal{R}^2,\ \mathcal{W}^2,\ \mathcal{W}^2_+,\ \mathcal{W}^2_-,\ \mathcal{R}ic^2. \tag{1.1}$$

These are respectively the square of L^2 norm of the Riemann curvature R, Weyl curvature W, its self-dual and anti-self-dual components, W_+, W_-, and Ricci curvature Ric. We will also consider, but in much less detail, the scalar curvature functionals

$$\mathcal{S}^2, -\mathcal{S}|_{\mathcal{Y}}, \tag{1.2}$$

given by the square of the L^2 norm of the scalar curvature s, and the restriction of the total scalar curvature to the space $\mathcal{Y}$ of unit volume Yamabe metrics on M. The Chern-Gauss-Bonnet theorem [17] relating the functionals in (1.1)-(1.2) plays a crucial rôle in the analysis to follow.

Partially supported by NSF Grant DMS 0305865.

Einstein metrics, satisfying

$$Ric_g = \frac{s}{n}g, \tag{1.3}$$

$n = dimM$, are critical points of all of the functionals in (1.1)-(1.2), and in many cases, Einstein metrics are minimizers. However, there are large classes of minimizers of $\mathcal{W}^2$, or the related $\mathcal{W}^2_+$, $\mathcal{W}^2_-$ which are not Einstein. Critical metrics of $\mathcal{W}^2$ are Bach-flat metrics, satisfying the Bach equations, and include conformally flat as well as half-conformally flat (self-dual or anti-self-dual) metrics and these classes of metrics are also often minimizers. It does not seem to be known if there are any other minimizers, or even critical points of the functionals in (1.1) or (1.2), which are not Einstein or half-conformally flat.

Just as in dimension 3, on a general 4-manifold, metrics minimizing a particular functional in (1.1) will not exist. Until relatively recently, the only known obstructions to the existence of Einstein metrics were the Hitchin-Thorpe inequality $\chi(M) \geq \frac{3}{2}|\tau(M)|$ between the Euler characteristic and signature of M, and Gromov's improvement of this, based on the simplicial volume. In the past decade or so, many further obstructions have been found by LeBrun and others, which show for instance that the existence of an Einstein metric on a 4-manifold M often depends strongly on the smooth structure of M as opposed to just the topological structure; we refer to [28], [29] for overviews of the current status of these issues. Nevertheless, one is far from having a comprehensive understanding of the obstructions to Einstein or half-conformally compact metrics on a given 4-manifold.

In §3, we survey in some detail the currently known results concerning the structure of the moduli space of Einstein metrics and moduli spaces $\mathcal{M}_{\mathcal{F}}$ of the functionals $\mathcal{F}$ in (1.1).

In §4, these results, and the methods used in their proof, are extended to prove a general result on the weak or idealized existence of minimizers of the functionals in (1.1). The main result is summarized as follows, but we refer to Theorem 4.10, both for the definitions involved and for a more precise formulation.

Theorem 1.1. *Let M be a closed, oriented 4-manifold and let $\mathcal{F}$ be either of the functionals $\mathcal{R}^2$ or $\mathcal{R}ic^2$ in (1.1). Then there exist minimizing sequences $\{g_i\}$ for $\mathcal{F}$ on the space $\mathbb{M}_1$ of unit volume metrics metrics on M which exhibit one of the following behaviors:*

(I). *The sequence $\{g_i\}$ converges in the Gromov-Hausdorff topology to a compact, oriented orbifold (V, g_0) associated to M, possibly reducible, with C^∞ metric g_0 on the regular set V_0, and having a C^0 extension across the singular points. One has*

$$\mathcal{F}(g_0) \leq \inf_{g\in\mathbb{M}_1} \mathcal{F}(g), \quad \text{and} \quad vol_{g_0}V = 1. \tag{1.4}$$

(II). *The sequence $\{g_i\}$ collapses everywhere, i.e. $inj_{g_i}(x) \to 0$, for all $x \in M$, and on the complement of a finite collection B_i of arbitrarily small balls $B_i = \cup_k B_{z_k}(\varepsilon_i)$, $\varepsilon_i \to 0$, the sequence $\{g_i\}$ collapses with locally bounded curvature along a sequence of F-structures on $M \setminus B_i$.*

(III). *The sequence $\{g_i\}$ converges in the pointed Gromov-Hausdorff topology to a maximal open orbifold Ω, possibly reducible and possibly empty, with C^∞ smooth metric g_0 on the regular set Ω_0, C^0 across singular points, and satisfying*

$$\mathcal{F}(g_0) \leq \inf_{g\in\mathbb{M}_1} \mathcal{F}(g), \quad \text{and} \quad vol_{g_0}\Omega \leq 1. \tag{1.5}$$

Any compact set $K \subset \Omega_0$ embeds in M, and if K is sufficiently large, the complement $M \setminus K$ carries an F-structure, metrically on the complement B_i of finitely many balls of arbitrarily small radius, as in (II).

In both cases (I) *and* (III), *the metric g_0 satisfies the Euler-Lagrange equation*

$$\nabla\mathcal{F} = 0. \tag{1.6}$$

A similar but slightly weaker result holds for the conformally invariant functionals $\mathcal{W}^2$ or $\mathcal{W}^2_\pm$, cf. Theorem 4.11.

These results give a general framework in which to study the existence of minimizers of one of the curvature functionals in (1.1) and, in situations where such metrics don't exist on the manifold M, a framework to try to understand what the obstructions to existence might be. Note that the general structure given by Theorem 1.1, and its analogue for the conformally invariant functionals, is the same for all functionals $\mathcal{F}$ in (1.1).

The results above also apply to the moduli spaces of minimizers, (or critical points) of $\mathcal{F}$, and in this context generalize recent results in [8], [45], cf. Theorem 4.15. A number of questions related to Theorem 1.1 are raised in §4, the most important being to what extent the domain Ω is topologically essential in M, analogous to the Thurston decomposition in dimension 3.

We point out one particular consequence of the proof of Theorem 1.1 here, related to an open question of Gromov and work of Rong [37]; again see Theorem 4.18 for more details.

Theorem 1.2. *There is an $\varepsilon_0 > 0$, such that if M is a 4-manifold admitting a metric with*

$$\int_M |R|^2 \leq \varepsilon_0, \tag{1.7}$$

then M has an F-structure.

We do not attempt here to give a broad overview of results on canonical metrics on 4-manifolds, which would require a much longer article; thus many important topics are not discussed at all. Some important omissions include the existence of canonical Kähler-Einstein metrics, where a great deal more is known based on Yau's solution of the Calabi conjecture [47]. Similarly, extremal Kähler metrics and twistor theoretic techniques are not addressed. In fact, the relations between the canonical metric problem with complex and algebraic geometry are not considered, and it would be interesting to see if the conclusions of Theorem 1.1 can be strengthened in the context of Kähler metrics for instance.

Finally, all manifolds below are compact, connected and oriented, and of dimension 3 or 4, unless otherwise stated.

I would like to thank the referee for several useful comments on the manuscript.

2. 3-Manifolds.

In dimension 3, it is natural to define canonical metrics to be the metrics of constant curvature, or equivalently, Einstein metrics. Most 3-manifolds M do not admit an Einstein metric; in fact it is quite easy to see that essential spheres and tori obstruct the existence of an Einstein metric, (except tori in the very special case of flat 3-manifolds). Here, an embedded sphere S^2 in M is essential if it does not bound a 3-ball in M, while an embedded torus is essential if the embedding induces an injection of fundamental groups. So for example, a non-trivial connected sum $M_1 \# M_2$, or any circle bundle over a surface with infinite fundamental group, (which is not a flat 3-manifold), does not carry an Einstein metric.

A special case of the Thurston Geometrization Conjecture, (the most important case given Thurston's results on the conjecture [43]), is that the simplest essential surfaces embedded in M, namely spheres and tori, are the only obstructions to the existence of an Einstein metric. In fact, the conjecture states that a general 3-manifold may be naturally split along a suitable collection of such spheres (sphere decomposition) and tori (torus decomposition) into pieces, each of which admits a canonical geometric structure. A geometric structure is a mild generalization of an Einstein metric, namely a complete, locally homogeneous metric. There are eight types of geometries; the three of constant curvature and five which are products or twisted products of lower dimensional manifolds, (where the uniformization theorem for surfaces comes into play).

To describe the splitting in a bit more detail, the sphere decomposition is a decomposition into a connected sum of irreducible 3-manifolds, and has the form

$$M = (K_1\#...\#K_p)\#(L_1\#...\#L_q)\#(\#_1^r S^2 \times S^1), \tag{2.1}$$

where the K and L factors are irreducible and of infinite π_1 and finite π_1 respectively; irreducible means that every embedded S^2 bounds a 3-ball B^3 in the manifold. The torus decomposition is a splitting of a K-factor into a finite collection of disjoint open manifolds $K \setminus \overline{\mathcal{T}}$, where $\overline{\mathcal{T}}$ is a finite collection of disjoint, non-isotopic, essential tori in K such that each component of $K \setminus \overline{\mathcal{T}}$ has no essential tori not homotopic to boundary torus in $\mathcal{T}$.

Thurston's conjecture is the assertion that each L factor in (2.1) is a spherical space form, while each K factor has the form

$$K = H \cup_{\mathcal{T}} G, \tag{2.2}$$

where H is a finite union of complete hyperbolic 3-manifolds of finite volume, and G is a finite union of graph manifolds, with $\mathcal{T} \subset \overline{\mathcal{T}}$. Each component of G may be further decomposed as a union of circle bundles over surfaces with boundary, (Seifert fibered spaces); the resulting toral boundary components then essentially comprise $\overline{\mathcal{T}} \setminus \mathcal{T}$. Thus G is a union of Seifert fibered spaces with boundary, glued together by toral automorphisms. The Seifert fibered pieces of G carry product or twisted product geometries. There is one exception to the rule (2.2), namely when K is a 3-dimensional Sol-manifold, i.e. a finite cover of K is a non-trivial torus bundle over a circle.

In a remarkable series of papers [33]-[35], Perelman has apparently proved the Geometrization Conjecture. His work has gradually gained increasing acceptance among experts and it seems likely that full acceptance will occur in the near future. In addition to Perelman's papers, there are now a number of expositions of his work at various levels, and so the proof will not be discussed here; see also the general source [25].

For later purposes, there is one point worth explaining however. While the method, the Ricci flow with surgery, leads to the geometrization of the constant curvature (Einstein) factors in (2.2), it does not lead to the geometric structures on the graph manifold part G, (or the Sol geometry). Instead, the geometry of G that emerges is that of collapse along the circle fibers in the Seifert fibered spaces and collapse of the toral regions glueing them together, (or collapse of toral fibers in Sol manifolds). Thus, the basic configuration in the limit is a collection of Einstein metrics, together with a well-defined degeneration by collapse of the remaining parts of M.

The point worth emphasizing here is that although most 3-manifolds do not carry Einstein metrics, given Perelman's work one has a precise understanding of which 3-manifolds do, and how a general 3-manifold is obtained by assembling pieces having such canonical geometries.

Finally, the moduli space of Einstein metrics on 3-manifolds is completely understood; the spherical space-forms are rigid (Calabi), as are the hyperbolic manifolds of finite volume (Mostow and Mostow-Prasad). The moduli spaces of the remaining six geometries are basically determined by the moduli of constant curvature metrics on the underlying surfaces.

Application to the Sigma Constant.

Since it does not appear to be widely known at this time, we give an application of Perelman's work to the Sigma constant, also called the Yamabe invariant, of 3-manifolds. Thus, let $\mathcal{S}$ denote the Einstein-Hilbert action restricted to the space $\mathbb{M}_1$ of unit volume metrics on a given 3-manifold M;

$$\mathcal{S}(g) = \int_M s_g dV_g, \tag{2.3}$$

where s_g is the scalar curvature of g. $\mathcal{S}$ is bounded below in any given conformal class $[g]$ and the invariant $\sigma(M)$ is given by

$$\sigma(M) = \sup_{[g]\in\mathcal{C}}\{\inf_{[g]}\mathcal{S}(g)\} = \sup_{\gamma\in\mathcal{Y}} s_{\gamma}, \tag{2.4}$$

where $\mathcal{C}$ is the space of conformal classes and $\mathcal{Y}$ is the space of unit volume Yamabe metrics. Now suppose

$$\sigma(M) \leq 0. \tag{2.5}$$

It follows from classical work of Schoen-Yau or Gromov-Lawson that (2.5) occurs if the decomposition (2.1) contains at least one K factor; (Perelman's work implies that (2.5) occurs precisely when (2.1) contains at least one K factor).

We show that Perelman's work implies that when $\sigma(M) \leq 0$, $\sigma(M)$ is determined by the volume of the hyperbolic part of M, in that

$$|\sigma(M)| = 6(vol_{-1}H)^{2/3}, \tag{2.6}$$

where $vol_{-1}H$ is the volume of H with respect to the metric of constant curvature -1. In particular, the graph manifold part G and the positive parts S^3/Γ, $S^2 \times S^1$ if any, are invisible to $\sigma(M)$. Perelman's work answers affirmatively a conjecture of Schoen in [39], and its generalization in [6].

To prove (2.6), consider the quantity

$$S_{-}(M) = \sup_{g\in\mathbb{M}}\{s_{min}v^{2/3}(g)\}, \tag{2.7}$$

where the sup is taken over the space $\mathbb{M}$ of all metrics on M, $s_{min}(g) = \min_M s_g$ and v is the volume of (M, g). The product in (2.7) is scale invariant. It is easy to see that when $\sigma(M) \leq 0$, then

$$S_{-}(M) = \sigma(M). \tag{2.8}$$

Namely, since Yamabe metrics are of constant scalar curvature, one has $S_{-}(M) \geq \sigma(M)$. On the other hand, given any g, let $\widetilde{g} = u^4 g$ be a Yamabe metric of the same volume as g in $[g]$, so that u satisfies the Yamabe equation

$$u^5\widetilde{s} = -8\Delta u + su. \tag{2.9}$$

When $\widetilde{s} \leq 0$, the maximum principle implies that $\widetilde{s} \geq s_{min}$, (since $\max u \geq 1$). This proves (2.8), and so (2.6) follows from

$$|S_{-}(M)| = 6(vol_{-1}H)^{2/3}. \tag{2.10}$$

To prove (2.10), suppose first that M is irreducible, so the sphere decomposition (2.1) is trivial, $(M = K)$. Then

$$M = H \cup_{\mathcal{T}} G, \tag{2.11}$$

where the union is along incompressible tori $\mathcal{T}$. Now it is easy to construct a metric g_{ε} on M such that

$$s_{min}v^{2/3}(g_{\varepsilon}) \geq -6(vol_{-1}H)^{2/3} - \varepsilon, \tag{2.12}$$

for any given $\varepsilon > 0$. This can be done "by hand", by taking a truncation of the hyperbolic metric on H, joined with a highly collapsed metric on G; one can easily construct such metrics on G with $s \geq -6$, $vol \leq \varepsilon$, for any given $\varepsilon > 0$, and which smoothly glue onto the hyperbolic cusps sufficiently far down the cusps, cf. [5], [7] for further details. (If $H = \emptyset$, then this already implies (2.10)). Thus one has

$$S_{-}(M) \geq -6(vol_{-1}H)^{2/3} = -\tfrac{3}{2}(vol_{-1/4}H)^{2/3}. \tag{2.13}$$

Now suppose there is a metric g_0 on M such that $S_-(g_0) > -\frac{3}{2}(vol_{-1/4}H)^{2/3}$. Then start the Ricci flow on M with initial metric g_0. Perelman's work implies that the Ricci flow with surgery g_t exists for all time, and that the scale invariant quantity $S_-(g_t)$ is monotone non-decreasing in t, since $S_-(g_t) \leq 0$ for all t, cf. [34]. Hence as $t \to \infty$, $S_-(g_t) \to \widetilde{S} > -\frac{3}{2}(vol_{-1/4}H)^{2/3}$, with $\widetilde{S} \leq 0$.

On the other hand, as Perelman shows, the rescaled metrics $\widetilde{g}_t = t^{-1}g_t$ have the property that $s_{min}(\widetilde{g}_t) \to -\frac{3}{2}$ as $t \to \infty$, cf. [34]. Now the decomposition (2.11) is unique up to isotopy, (cf. [7]), and the metrics $\widetilde{g}_t$, when restricted to compact subsets of H, converge to the hyperbolic metric with curvature $-1/4$. Thus, one must have $\widetilde{V} = \liminf_{t\to\infty} vol(\widetilde{g}_t) \geq vol_{-1/4}H$. Hence, $\widetilde{S} = \limsup_{t\to\infty} s_{min}(\widetilde{g}_t) vol(\widetilde{g}_t)^{2/3} \leq -\frac{3}{2}(vol_{-1/4}H)^{2/3}$, which gives a contradiction.

If M is not irreducible, then M is a connected sum of positive factors S^3/Γ, $S^2 \times S^1$ and non-positive irreducible factors K_i. The work above shows that (2.10) holds on each K_i. One can perform the connected sum surgery by hand to increase s_{min} pointwise and with arbitrarily small change to the volume, cf. again [5], [7], so that (2.12) holds for general M. One may then apply exactly the same argument as before to prove that (2.10) holds, when $\sigma(M) \leq 0$. (The Ricci flow with surgery performs the sphere decomposition (2.1), and in particular disconnects the factors in (2.1) in finite time, while the K factors persist for infinite time). ■

In contrast, no applications of Perelman's ideas have yet been found to determine the Sigma constant of the positive 3-manifolds, i.e. S^3/Γ; cf. [1], [14] for some recent progress on this problem.

Observe that (2.10) shows that if (M, g) is any closed Riemannian 3-manifold with $\sigma(M) \leq 0$, then

$$s_g \geq -6 \Rightarrow vol_g M \geq vol_{-1}H, \tag{2.14}$$

where H is the hyperbolic part of M. This gives a very strong generalization of results of [13] in dimension 3, and extends their results from Ricci curvature to scalar curvature.

In fact, (2.14) can easily be generalized somewhat further. Let

$$\mathcal{S}_-^{3/2}(g) = \int |\min(s_g, 0)|^{3/2} dV_g. \tag{2.15}$$

Then it is easy to see that

$$|\sigma(M)|^{3/2} = \inf_{g\in\mathbb{M}_1} \mathcal{S}_-^{3/2}(g). \tag{2.16}$$

Namely, the definition (2.4) gives immediately $|\sigma(M)|^{3/2} \geq \inf_{g\in\mathbb{M}_1} \mathcal{S}_-^{3/2}(g)$. On the other hand, given any $g \in \mathbb{M}_1$, let γ be a unit volume Yamabe metric in $[g]$. Setting $g = u^4\gamma$, as in (2.9) one has $u^5 s_g = -8\Delta u + s_\gamma u$. Since s_γ is a non-positive constant, simple calculations give

$$|s_\gamma| = -\int s_\gamma dV_\gamma = -\int s_g u^4 dV_\gamma - 8\int u^{-1}\Delta u dV_\gamma = -\int s_g u^4 dV_\gamma - 8\int |d\log u|^2 dV_\gamma$$

$$\leq \int |\min(s_g, 0)| u^4 dV_\gamma \leq [\int |\min(s_g, 0)|^{3/2} u^6 dV_\gamma]^{2/3} = (\mathcal{S}^{3/2}(g))^{2/3}.$$

This gives $\inf \mathcal{S}^{3/2}(g) \geq |\sigma(M)|^{3/2}$, and so (2.16). Hence, (2.14) generalizes to

$$\mathcal{S}_-^{3/2}(g) \geq 6^{3/2} vol_{-1}H. \tag{2.17}$$

In fact, (2.17) reflects the behavior of metrics minimizing the functional $\mathcal{S}_-^{3/2}$, (or stronger functionals such as $\mathcal{S}^2$) on a given 3-manifold with $\sigma(M) \leq 0$. Thus, one may find minimizing sequences $\{g_i\}$ for $\mathcal{S}_-^{3/2}$ which crush essential 2-spheres in M to points, according to the sphere decomposition, and on each non-positive K-factor, converge to the complete hyperbolic metric on H, while collapsing the graph manifold part G with uniformly bounded curvature. The positive parts S^3/Γ and $S^2 \times S^1$ are invisible to $\mathcal{S}_-^{3/2}$. Thus, one can construct minimizing sequences for $\mathcal{S}_-^{3/2}$ which give a geometric decomposition of M, equivalent to the Thurston decomposition, (cf. [7] for further details).

This will be the main point of view in the analysis to follow in 4-dimensions.

3. 4-Manifolds: Moduli Spaces.

On 4-manifolds, it is less clear what a canonical metric should be. As discussed in §1, we will take the point of view of variational problems on the space of metrics $\mathbb{M}$ on a given 4-manifold M and define such a metric to be a minimizer, (or possibly a critical point), of one of the curvature functionals $\mathcal{F}$ in (1.1), i.e.

$$\mathcal{R}^2,\ \mathcal{W}^2,\ \mathcal{W}_+^2,\ \mathcal{W}_-^2,\ \mathcal{R}ic^2, \tag{3.1}$$

or the much weaker scalar curvature analogues,

$$\mathcal{S}^2,\ -\mathcal{S}|_{\mathcal{Y}}. \tag{3.2}$$

These functionals are all bounded below, and so in principle one can use direct methods in the calculus of variations to study the existence and properties of minimizers.

The basic problem is to understand the existence and moduli spaces of such metrics on a given manifold M. However, just as in dimension 3 as discussed in §2, one cannot expect an arbitrary 4-manifold admits a smooth metric minimizing one of the functionals $\mathcal{F}$. In fact, the situation in dimension 4 is much more complicated than that in 3 dimensions. While numerous obstructions to the existence of minimizers of a given $\mathcal{F}$ are known, cf. [28], [30] and further references therein, there is no general conjecture as to what an exact and complete set of obstructions is, i.e. there is currently no analog of the Thurston geometrization conjecture.

Nevertheless, it is natural to try to find a geometric decomposition of M with respect to one of these functionals. Thus, as discussed at the end of §2, one can try to see if minimizing sequences decompose the manifold into pieces, (analogous to (2.1) or (2.2)), on some of which they converge to smooth limits and others on which they degenerate in a well-defined way.

The single most important fact allowing one to develop such a theory on the existence, or the structure of moduli spaces of such functionals, is Chern's generalization of the Gauss-Bonnet theorem; in dimension 4, this is

$$\frac{1}{8\pi^2}\int \{|R|^2 - |z|^2\}dV = \chi(M), \tag{3.3}$$

where $z = Ric - \frac{s}{4}g$ is the tracefree Ricci curvature. The expression (3.3) is equivalent to

$$\frac{1}{8\pi^2}\int \{|W|^2 - \frac{1}{2}|z|^2 + \frac{1}{24}s^2\}dV = \chi(M). \tag{3.4}$$

This gives one L^2 control of the full curvature R (or W) in terms of L^2 control of Ric. Chern-Weil theory and the signature theorem also give the relation

$$\frac{1}{12\pi^2}\int \{|W_+|^2 - |W_-|^2\}dV = \tau(M), \tag{3.5}$$

where $\tau(M)$ is the signature of M. Combining (3.4) and (3.5) gives

$$\frac{1}{2\pi^2}\int \{|W_+|^2 - \frac{1}{4}|z|^2 + \frac{1}{48}s^2\}dV = 2\chi(M) + 3\tau(M). \tag{3.6}$$

The functionals (3.1) are all scale-invariant in dimension 4. In the following, we will always work on the space $\mathbb{M}_1$ of unit volume metrics on M, unless stated otherwise.

In this section, we study the structure of the moduli spaces of minimizers or critical points of the functionals in (3.1). This serves as an introduction as to what one can expect for existence results, which are discussed in §4.

(A). Einstein Moduli Spaces.

We begin with the case of Einstein metrics, which are critical points of all the functionals in (3.1) and (3.2). Let $\mathcal{M} = \mathcal{M}_E$ denote the moduli space of unit volume Einstein metrics on M. Since for instance the functional $\mathcal{S}^2$ is critical on $\mathcal{M}$ and Einstein metrics have constant scalar curvature, the scalar curvature $s_g : \mathcal{M} \to \mathbb{R}$ is constant on components of $\mathcal{M}$. By (3.3), $\mathcal{R}^2$ is constant on all of $\mathcal{M}$, while by (3.4), $\mathcal{W}^2$ is again constant on components of $\mathcal{M}$.

The first general result on the structure of the moduli space $\mathcal{M}$ of unit volume Einstein metrics on a given 4-manifold M was obtained in [2], [10], [32]; a partial result along these lines was also obtained in [44] in the special case of Kähler-Einstein metrics with $c_1 > 0$. Overall, the picture resembles somewhat Uhlenbeck's results on the moduli space of self-dual Yang-Mills fields.

To describe this, an Einstein orbifold (V, g) associated to M is defined to be a (4-dimensional) orbifold, with a finite number of singular points $\{q_i\}$, each having a neighborhood homeomorphic to the cone $C(S^3/\Gamma)$, where $\Gamma \neq \{e\}$ is a finite subgroup of $SO(4)$. Let $V_0 = V \setminus \cup q_k$ be the regular set of V. Then g is a smooth Einstein metric on V_0, which extends smoothly over $\{q_k\}$ in local finite covers. The manifold M is a resolution of V in the sense that there is a continuous surjection $\pi : M \to V$ such that $\pi|_{\pi^{-1}(V_0)} : \pi^{-1}(V_0) \to V_0$ is a diffeomorphism onto V_0. In particular, V is compact.

Then the result is that the completion $\hat{\mathcal{M}}$ of $\mathcal{M}$ in the Gromov-Hausdorff topology consists of $\mathcal{M}$ together with unit volume Einstein orbifold metrics associated to M. Moreover, the completion is locally compact, in that any sequence g_i of unit volume Einstein metrics on M, bounded in the Gromov-Hausdorff topology, has a subsequence converging to an Einstein orbifold associated to M.

In analogy to the Uhlenbeck completion of the moduli space of Yang-Mills instantons, orbifold singularities arise from the bubbling off of gravitational instantons, i.e. complete non-flat Ricci-flat metrics (N, g_∞) which are ALE, (asymptoticaly locally Euclidean), in that the metric g_∞ is asymptotic to a flat cone $C(S^3/\Gamma)$ at infinity. There is at least one such ALE space associated to each singularity; however, in general there may be finitely many such spaces, arising at different blow-up scales. All such blow-up limits N are topologically embedded in M and a simple Mayer-Vietoris argument shows that most of the rational homology of any such N injects in the homology of M, in that

$$0 \to H_k(N, \mathbb{R}) \to H_k(M, \mathbb{R}), \tag{3.7}$$

for $k = 1, 2$.

Such ALE spaces (N, g_∞) have nontrival 2^{nd} Betti number, and V is obtained from M by collapsing essential cycles in $H_2(M, \mathbb{R})$ to points. In particular, if $b_2(M) = 0$, then there are no orbifold singularities and so $V = M$. This is the case for example M is a rational homology sphere, (with any differentiable structure). Also, the proof of the smoothness of g across the orbifold singularities in a local uniformization follows the lines of proof of Uhlenbeck's removable singularity theorem, cf. [10]. Finally, since each ALE space (N, g_∞) has a definite amount of curvature in L^2, one has a uniform bound on the number of orbifold singularities depending only on $\chi(M)$, by the Chern-Gauss-Bonnet theorem (3.3).

The completion in the Gromov-Hausdorff topology is equivalent to the completion with respect to a diameter bound, so that all metrics $g \in \hat{\mathcal{M}}$ satisfy

$$vol_g M = 1, \quad diam_g M \leq D, \tag{3.8}$$

for some $D = D(g) < \infty$.

However, there is a marked difference compared with the Uhlenbeck completion. Namely, the frontier $\hat{\mathcal{M}} \setminus \mathcal{M}$ should perhaps not be considered as a boundary, but instead as a filling in of missing pieces in $\mathcal{M}$. For example, in the case of K3 surfaces, the moduli space has dimension 57, and the frontier consists of subvarieties of codimension 3 in $\hat{\mathcal{M}}$, cf. [12]; it does not form a natural boundary as a "wall" past which $\hat{\mathcal{M}}$ cannot be continued. This is the case in all known examples, although it is unknown if this holds in general.

A main point of the Uhlenbeck completion is that the completion is compact. Consider the components of $\mathcal{M}$ for which

$$s_g \geq s_0 > 0. \tag{3.9}$$

Myers' theorem then implies (3.8) holds, (with $D = D(s_0)$), and so the completion $\hat{\mathcal{M}}$ of this part of $\mathcal{M}$ in the Gromov-Hausdorff topology is compact. However, this is certainly not the case when $s_g \leq 0$, (consider for example flat tori or products of hyperbolic surfaces). Thus, one needs to consider what happens when the Gromov-Hausdorff distance goes to infinity.

In [4] a more complete theory of the global behavior of $\mathcal{M}$ was developed, which from a broad perspective has a strong resemblance with the moduli space of constant curvature metrics on surfaces. (The compactness of the part of $\hat{\mathcal{M}}$ for which (3.9) holds corresponds to the compactness of the moduli space of Einstein metrics on S^2). In studying the boundary of Teichmüller space or the Riemann moduli space, one of the most natural metrics is the Weil-Petersson metric. This has a natural generalization to all dimensions, since it is just the restriction of the natural L^2 metric on $\mathbb{M}$ to the moduli space. Thus, we consider the completion of $\mathcal{M}$ with respect to the L^2 metric.

To describe the results, one needs the following definition. A domain Ω (i.e. an open 4-manifold) weakly embeds in M, $\Omega \subset\subset M$, if for any compact subdomain $K \subset \Omega$, there is a smooth embedding $F = F_K : K \to M$. The same definition applies if Ω is an orbifold, with the obvious modification that the corresponding part of M is a resolution of K.

The completion $\overline{\mathcal{M}}$ of $\mathcal{M}$ with respect to the L^2 metric on $\mathbb{M}$ is a complete Hausdorff metric space, whose frontier $\partial\mathcal{M}$ consists of two parts: the orbifold part $\partial_o\mathcal{M}$ and the cusp part $\partial_c\mathcal{M}$.

(I). $\partial_o\mathcal{M}$ consists of compact Einstein orbifolds of unit volume associated to M. The partial completion $\mathcal{M} \cup \partial_o\mathcal{M}$ is locally compact. This is the same as the situation described before.

(II). An element in the cusp boundary $\partial_c\mathcal{M}$ is given by a pair (Ω, g), where Ω is a non-empty maximal orbifold domain Ω weakly embedded in M. The domain Ω consists of a finite number of components Ω_k called cusps, each with a bounded number, (possibly zero), of orbifold singularities. The metric g is a complete Einstein metric on Ω, with

$$vol_g\Omega = 1, \tag{3.10}$$

and outside a compact set K, Ω carries an F-structure along which g collapses with locally bounded curvature as one goes to infinity in Ω; thus as $x \to \infty$ in Ω,

$$inj(x) \to 0 \ \text{ and } \ (|R|inj)^2(x) \to 0, \tag{3.11}$$

where $inj(x)$ is the injectivity radius at x.

To describe the behavior of the region $M \setminus K$, let g_i be a sequence in $\mathcal{M}$ with $g_i \to g$ in the L^2 metric. Then $M \setminus K$ also carries an F-structure on the complement of a finite number of arbitrarily small balls. Thus, there exists a finite collection of points $z_j \in M$, and a sequence $\varepsilon_i \to 0$ such that outside $B_{z_j}(\varepsilon_i)$, $M \setminus K$ has an F-structure. If one chooses an exhaustion K_i of Ω, then $M \setminus K_i$ collapses everywhere, and collapses with locally bounded curvature away from the singular points $\{z_j\}$.

The convergence in (I) is also in the Gromov-Hausdorff topology, while that in (II) is also in the pointed Gromov-Hausdorff topology, for a suitable collection of base points. Further, cusps, i.e.

Case (II), can occur only on the components of $\mathcal{M}$ for which there is a constant s_0 such that

$$s_g \leq s_0 < 0. \tag{3.12}$$

The L^2 completion $\overline{\mathcal{M}}$ is not compact in general, as seen explicitly in the cases of flat metrics on tori, or Ricci-flat metrics on $K3$. On the $K3$ surface, the L^2 metric on $\overline{\mathcal{M}}$ is the complete metric of finite volume on the non-compact locally symmetric space $\Gamma \setminus SO(3,19)/(SO(3) \times SO(19))$, cf. [4], [12], so that $\overline{\mathcal{M}} = \mathcal{M} \cup \partial_o \mathcal{M}$. (Of course if (3.9) holds, then $\overline{\mathcal{M}}$ is compact). The behavior of $\overline{\mathcal{M}}$ at infinity is described as follows:

(III). Suppose g_i is a divergent sequence in $\overline{\mathcal{M}}$ such that

$$s_{g_i} \to 0, \quad \text{as} \quad i \to \infty. \tag{3.13}$$

Then $\{g_i\}$ collapses everywhere with locally bounded curvature, i.e. (3.11) holds, metrically on the complement of finitely many singular points $\{z_j\}$. (Thus, $\Omega = \emptyset$ in the context of (II)).

Suppose instead that g_i is a divergent sequence in $\overline{\mathcal{M}}$ such that

$$s_{g_i} \leq s_0 < 0, \quad \text{as} \quad i \to \infty. \tag{3.14}$$

Then $\{g_i\}$ either has the same behavior as in (III) or (II), where Ω may instead have possibly infinitely many components, (of total volume at most 1).

Recently, building on the work in [4], Cheeger-Tian [16] have improved the statement above, and proved:

(IV). Suppose instead that g_i is a divergent sequence in $\overline{\mathcal{M}}$ such that

$$s_{g_i} \leq s_0 < 0, \quad \text{as} \quad i \to \infty.$$

Then $\{g_i\}$ has the same behavior as in (II). Moreover, the collapse with locally bounded curvature is actually collapse with uniformly bounded curvature: $|R| \leq \Lambda$, for some $\Lambda = \Lambda(M) < \infty$, away from the singular points.

To complete the analogy with the case of surfaces, it is natural to conjecture, (cf. [4]), that in fact Case (IV) does not occur, i.e. when (3.14) holds, $\overline{\mathcal{M}}$ is compact. We recall here that the completion of the moduli space of hyperbolic metrics on a surface with respect to the Weil-Petersson (L^2) metric is compact, and agrees with the Deligne-Mumford compactification.

There are many open questions regarding the structure of $\mathcal{M}_E$ that remain unanswered. Among the most basic are the following:

(i). Does $\mathcal{M}$ or $\overline{\mathcal{M}}$ have finitely many components? This is open even for the portion of $\mathcal{M}$ satisfying (3.9) where $\overline{\mathcal{M}}$ is compact. Apriori, one could have a sequence of metrics g_i in distinct components of $\mathcal{M}$ which converge to an orbifold metric $(V, g) \in \overline{\mathcal{M}}$ in a limit component.

(ii). If g_i is a sequence in $\mathcal{M}$ converging to $(V, g) \in \overline{\mathcal{M}}$ as above, does there exist a curve $\gamma(t)$ in $\overline{\mathcal{M}}$, with $\gamma(0) = g$ and $\gamma(t_i) = g_i$, for some sequence $t_i \to 0$?

(iii). The occurence of orbifold singularities is closely related to the topology of M, via (3.7). If it is possible topologically for orbifold singularities to occur, do they in fact occur in $\overline{\mathcal{M}}$?

(iv). Is there any relation between the topology of Ω and the topology of M, i.e. is Ω in any way topologically essential in M?

(v). What can be said about the structure of the singularities in the collapsing situation of Case (III), or in the complement of Ω in Case (II)? A concrete example of collapsing metrics on the $K3$ surface is described in detail in [23]; the singularities here are modeled on the Ooguri-Vafa metrics, which are periodic versions of the Taub-NUT metrics.

(B) General Moduli spaces.

The first result (I) discussed above, on the completion of the moduli space of Einstein metrics on M^4 with respect to the Gromov-Hausdorff topology, has recently been generalized to the functionals in (3.1). Thus, let $\mathcal{F}$ denote any of the functionals in (3.1), and consider the moduli space $\mathcal{M}_{\mathcal{F}}$ of critical points of $\mathcal{F}$ of unit volume. As before, $\mathcal{F}$ is constant on each component of $\mathcal{M}_{\mathcal{F}}$.

First, the definition of orbifold needs to enlarged somewhat, in that one allows a neighborhood of a singular point q to be given by a finite collection of cones $C(S^3/\Gamma_j)$, with vertex q, (not just a single cone). For emphasis, sometimes such orbifolds will be called reducible, while the orbifolds as previously defined will be called irreducible. In particular, an orbifold V is reducible if and only if the regular set V_0 has more than one component near each singular point. Further, some or all of the finite groups Γ_j may be trivial, corresponding to cones on S^3 and so 4-balls. Also, from now on, the metric g on V is only asserted to be C^0 across each singularity q in local uniformizations of the cones.

Choose $\nu_0 > 0$ and $\Lambda < \infty$, and let $\mathcal{M}_{\mathcal{F}}(\nu_0, \Lambda)$ denote the portion of $\mathcal{M}_{\mathcal{F}}$ consisting of all unit volume metrics g such that, for $r \leq 1$,

$$volB_x(r) \geq \nu_0 r^4, \quad \text{and} \quad \int |R|^2 \leq \Lambda, \tag{3.15}$$

where $B_x(r)$ is the geodesic r-ball about x in (M, g). Then for any $\nu_o > 0$ and $\Lambda < \infty$, the closure of $\mathcal{M}_{\mathcal{F}}(\nu_0, \Lambda)$ in the Gromov-Hausdorff topology consists of $\hat{\mathcal{M}}_{\mathcal{F}} = \mathcal{M}_{\mathcal{F}} \cup \mathcal{M}_o$, where $\mathcal{M}_o$ consists of orbifold singular metrics (V, g), as in (I), with the modifications in the definition of orbifolds discussed above; in particular, the orbifold V may be reducible. Here, in the case of the conformally invariant functionals $\mathcal{W}^2$ and $\mathcal{W}^2_{\pm}$, the representative metrics $g \in [g]$ are assumed to be unit volume Yamabe metrics. This result was proved in [8] and [45]. Under certain conditions, e.g. the presence of a uniform Sobolev inequality in addition to (3.15), the orbifolds are irreducible.

Of course the condition (3.15) rules out any collapse behavior. In §4, we discuss analogs of the general degeneration for Einstein metrics for the moduli spaces $\mathcal{M}_{\mathcal{F}}$, (cf. Theorem 4.15).

At this point, it is useful to consider the following simple example, which illustrates some strong differences between the functionals $\mathcal{R}ic^2$ or $\mathcal{R}^2$, and $\mathcal{W}^2$.

Example 3.1. Let $M = \Sigma \times S^1$, where Σ is a compact hyperbolic 3-manifold and let g_μ be a product metric on M of the form

$$g_\mu = \mu^2 g_{-1} + \mu^{-6} g_{S^1(1)}, \tag{3.16}$$

where g_{-1} is the hyperbolic metric on Σ. The metrics g_μ are conformally flat, $W = 0$, and so give a curve in the moduli space $\mathcal{M}_{\mathcal{W}}$ of minimizers of $\mathcal{W}^2$ on M.

The scalar curvature of g_μ is given by $s_{g_\mu} = -6\mu^{-2}$ while the volume is $vol_{g_\mu} M = 2\pi vol_{g_{-1}} \Sigma$. As $\mu \to 0$, one thus has

$$\mathcal{S}^2(g_\mu) = \int_M s^2_{g_\mu} dV_{g_\mu} \to \infty. \tag{3.17}$$

In particular, the L^2 norm of the curvature R of g_μ diverges to infinity, in strong contrast to the situation of metrics with bounds on $\mathcal{R}ic^2$ or $\mathcal{Z}^2$.

As $\mu \to 0$, the factor Σ collapses to 0 volume, causing the curvature to blow up, while the S^1 factor expands, to preserve a fixed volume. This behavior is completely different from the kind of limits one sees in the moduli space of Einstein metrics in (I)-(IV) above. In the opposite direction where $\mu \to \infty$, the hyperbolic factor expands, (to a flat metric), while the circle S^1 collapses; this behavior is of course consistent with the collapse behavior discussed above.

One may also perform the same construction with (Σ, g_{-1}) replaced by (S^3, g_{+1}), preserving conformal flatness. The estimate (3.17) holds, where now the scalar curvature diverges to $+\infty$ instead of $-\infty$. Again the metrics g_μ behave badly as $\mu \to 0$, and collapse in a quite different manner than the Einstein case.

Here however, the divergence of (3.17) and g_μ is due to only to a bad choice of gauge for the conformal class. While the metrics g_μ are Yamabe metrics in the negative case, they are not Yamabe in the positive case, since for instance there is a uniform upper bound on the scalar curvature of Yamabe metrics with a fixed volume. In a Yamabe gauge, the conformal classes $(S^3 \times S^1, [g_\mu])$ have a uniform bound on the L^2 norm of curvature, and converge to the round metric on S^4, with two antipodal points identified, giving rise to an orbifold singularity consisting of two cones on S^3 joined at the vertex, cf. [39], [8].

4. 4-Manifolds: Existence Issues.

One would like to extend the results above on the moduli spaces toward an existence theory for metrics minimizing one of the functionals $\mathcal{F}$ in (3.1), (or (3.2)). From the point of view of direct methods in the calculus of variations, this requires understanding the limiting behavior of minimizing sequences of metrics g_i for $\mathcal{F}$ on a given 4-manifold M.

Consider for instance $\mathcal{R}^2$, with absolute minimum realized by Einstein metrics, (if such exist), by the Chern-Gauss-Bonnet theorem (3.3). Any sequence $g_i \in \mathcal{M}_E$ is trivially a minimizing sequence for $\mathcal{R}^2$, and so general minimizing sequences will exhibit at least the degenerations described in (I)-(IV) of §3. In other words, a general existence theory of minimizers of $\mathcal{F}$ must include a full description of the moduli space $\mathcal{M}_{\mathcal{F}}$.

However, a general minimizing sequence $\{g_i\}$ say for $\mathcal{R}^2$ will not satisfy any particular PDE, (as is the case on the moduli space). At best, the bound on $\mathcal{R}^2$ gives $L^{2,2}$ control of the metrics g_i in some local coordinate charts. This is troublesome in dimension 4, since an $L^{2,2}$ bound does not give a pointwise (L^∞) bound. With only such weak control, it does not seem possible to say anything reasonable about some limiting behavior of $\{g_i\}$.

A different approach to a rather general existence theory was introduced and developed by Taubes, both for the existence of self-dual Yang-Mills instantons [40], and self-dual metrics, (which of course are minimizers of $\mathcal{W}_-^2$), [41], cf. also [42]. The idea here is to glue together exact solutions on pieces of a manifold to obtain a global approximate solution, and show the approximation can be perturbed into a global exact solution. Ingeniously, to accomplish this Taubes introduces a slightly stronger norm than $L^{2,2}$ norm, which does embed in L^∞. Further comments on this approach will follow later in §4.

The problem that $L^{2,2}$ does not embed in L^∞ can be overcome if instead one minimizes the L^{2p} norm of the curvature tensor, for any $p > 1$. In fact, there is a general "convergence theorem" in this context, which shows that many of the main features of the results in §3 for the moduli space $\mathcal{M}_E$ hold for metrics with just L^{2p} bounds on the Ricci curvature.

Theorem 4.1. *(Convergence under L^p curvature bounds.) Let $\{g_i\}$ be a sequence of unit volume metrics on M such that*

$$\int_M |Ric|^{2p} \leq \Lambda, \tag{4.1}$$

for some $p > 1$ and $\Lambda < \infty$.

Then a subsequence of $\{g_i\}$ converges in the pointed Gromov-Hausdorff topology to a maximal orbifold domain Ω, possibly empty, weakly embedded in M, with $L^{2,2p}$ smooth metric g on the regular set Ω_0. The convergence is in the weak $L^{2,2p}$ topology on Ω_0 and so in particular g is locally C^α on Ω_0, $\alpha = 1 - (2-p)/p$. The orbifold singularities are all irreducible, with local group $\Gamma \neq \{e\}$, and the metric extends C^0 across orbifold singularities. One has

$$vol_g \Omega \leq 1. \tag{4.2}$$

Outside a sufficiently large compact set $K \subset \Omega$, Ω carries an F-structure along which g collapses with locally bounded curvature in L^{2p} on approach to $\partial\Omega$. The complement $M \setminus K$ also carries an F-structure on the complement of a finite number of balls $B_{z_j}(\varepsilon_i)$ with respect to g_i, where $\varepsilon_i \to 0$

as $i \to \infty$. If K_i is an exhaustion of Ω, then $M \setminus K_i$ collapses everywhere, and collapses with locally bounded curvature in L^{2p} away from the singular points $\{z_j\}$.

We point out that in general, the metric g will not be complete on Ω, and the domain Ω could have infinitely many components. On the other hand, Ω may be compact, so that $\Omega = M$ or Ω is an orbifold V. Further, one may have $\Omega = \emptyset$, in which case $\{g_i\}$ collapses M along a sequence of F-structures with locally bounded curvature in L^{2p} on the complement of a bounded number of arbitrarily small balls.

Also, here and in the following, we will not detail the choice of (the collection of) base points used for the pointed Gromov-Hausdorff topology. Some choices of base points lead to limits consisting of components of Ω, while other choices lead to the collapsed part on the complement of Ω.

For the proof of Theorem 4.1, we first need the following definitions, cf. [3], [5].

Definitions. Let (N, g) be a Riemannian 4-manifold. Let $B_x(r)$ denote the geodesic r-ball about x in N and for $y \in B_x(r)$, let $D_y(s) = B_x(r) \cap B_y(s)$, $s \leq r$. The *μ-volume radius* of (N, g) at x is

$$\nu(x) = \nu_\mu(x) = \sup\{r : \frac{vol D_y(s)}{s^4} \geq \mu\}, \tag{4.3}$$

for all $y \in B_x(r)$. The parameter $\mu > 0$ may be freely chosen, but from now on we assume μ is fixed, say $\mu = 10^{-2}$.

The *L^q curvature radius* $\rho^{(q)}(x)$ is the largest radius such that for $y \in B_x(\rho^{(q)}(x))$, one has

$$\frac{s^{2q-4}}{vol D_y(s)} \int_{D_y(s)} |R|^q \leq c_0. \tag{4.4}$$

We assume that $q \geq 2$, so that $L^{2,q} \subset C^\alpha$, $\alpha = 1 - (4-q)/q$ when $q > 2$. The parameter $c_0 = c_0(q) > 0$ will be chosen to be sufficiently small. An important case is $q = 2$, and we set $\rho^{(2)} = \rho$.

The $L^{2,q}$ harmonic radius $r_h^{2,q}$ at x is the radius of the largest geodesic ball about x for which there exists a harmonic coordinate chart on $B_x = B_x(r_h^{2,q}(x))$ in which the metric components g_{ij} satisfy

$$e^{-C}\delta_{ij} \leq g_{ij} \leq e^{C}\delta_{ij}, \quad \text{as bilinear forms,} \tag{4.5}$$

and

$$(r_h^{2,q})^\lambda ||\partial^2 g_{ij}||_{L^q(B_x)} \leq C, \tag{4.6}$$

where $\lambda = (2q-4)/q$; here one assumes $q > 2$.

The parameter C is a fixed constant, e.g. $C = 1$. The radii ν, $\rho^{(q)}$ and $r_h^{2,q}$ all scale as distances under rescalings of the metric.

For $q > 2$, one has $\rho^{(q)}(x) \geq c \cdot r_h^{2,q}(x)$, for a fixed numerical constant c, (depending only on c_0, C). More importantly, there is constant $c_1 > 0$, depending only on c_0, C and a lower bound ν_0 for $\nu(x)$, such that

$$\rho^{(q)}(x) \leq c_1 r_h^{2,q}(x). \tag{4.7}$$

Thus, (4.7) holds on scales bounded above by the volume radius. In the following, it will be assumed that c_0 is chosen sufficiently small so that the ball $B = B_x(\min(\nu(x), \rho^{(q)}(x))$ is diffeomorphic to a ball in a flat manifold and the metric is C-close to the flat metric on B.

Finally, in harmonic coordinates, the Ricci curvature is an elliptic operator in the metric; in fact

$$-\tfrac{1}{2}Ric_{ij} = \Delta_g g_{ij} + Q_{ij}(g, \partial g), \text{ where } \Delta_g = g^{ab}\partial_a\partial_b. \tag{4.8}$$

Hence, if the metric is controlled in C^α, for some $\alpha > 0$, then an L^q bound for Ric implies an $L^{2,q}$ bound for g, and hence an L^q bound for the full curvature R.

Proof of Theorem 4.1. Let $\{g_i\}$ be any sequence of unit volume metrics on M satisfying (4.1); in the following we will usually write g for any of the metrics g_i. As will be seen, a crucial point is that the bound (4.1), together with the Chern-Gauss-Bonnet theorem (3.3) gives a bound

$$\int |R|^2 \leq \Lambda', \tag{4.9}$$

where Λ' depends only on Λ and $\chi(M)$.

Pick any $\nu_0 > 0$, and let M^{ν_0} be the ν_0-thick part of (M, g), $(g = g_i)$, given by

$$M^{\nu_0} = \{x \in M : \nu(x) > \nu_0\}. \tag{4.10}$$

Then M^{ν_0} is an open submanifold in M, and it follows from [3] that is $L^{2,2p}$ orbifold compact, (away from its boundary), in that (M^{ν_0}, g_i) has a subsequence converging to a limit orbifold domain (Ω_{ν_0}, g), embedded in M; (more precisely, there is a domain in M which is a resolution of (Ω_{ν_0}, g)). The metric g is an $L^{2,2p}$ metric, C^0 across the orbifold singularities, and Ω_{ν_0} has finitely many components. There is a uniform bound on the number of orbifold singular points on Ω_{ν_0}, depending only on $\chi(M)$, for the same reasons as discussed following (3.7); namely each orbifold singularity is associated to a finite number of Ricci-flat ALE spaces, each of which contributes a definite amount to the L^2 norm of the curvature in (4.9). Similarly, the orbifold Ω_{ν_0} is irreducible, with local group $\Gamma \neq \{e\}$.

Let ν_j be a decreasing sequence with $\nu_j \to 0$ as $j \to \infty$. The argument above applies to each $M^{\nu_j} \subset M^{\nu_{j+1}}$ and taking a diagonal subsequence of the double sequence (i, j) gives a maximal limit orbifold domain (Ω, g), having the structure described in Theorem 4.1. As noted above, Ω may be compact, in which case Ω is denoted by V and V is an orbifold associated to M. At the other extreme, one may have $\Omega = \emptyset$; in this case $M^{\nu_0} = \emptyset$, for any given ν_0, provided i is sufficiently large.

Next, consider the ν_0-thin part of (M, g_i), i.e.

$$M_{\nu_0} = \{x \in M : \nu(x) \leq \nu_0\}. \tag{4.11}$$

By the discussion above, we may assume that ν_0 is arbitrarily small, in that $\nu_0 = \nu_0(i) \to 0$ sufficiently slowly, as $i \to \infty$. We divide M_{ν_0} into two further subdomains. Thus, fix a large constant $K < \infty$, and let

$$U = \{x \in M_{\nu_0} : \rho(x) > K\nu(x)\}, \text{ and } W = \{x \in M_{\nu_0} : \rho(x) \leq K\nu(x)\}. \tag{4.12}$$

Thus W corresponds to the set where the curvature may concentrate in L^2, in the scale of the volume radius, cf. also [4, p.63] for the analogous description in the case of Einstein metrics.

Lemma 4.2. *In a subsequence, the set W tends metrically to finitely many points $\{z_j\}$, as $i \to \infty$.*

Proof: For $z_i \in W$, rescale the metrics g_i so that $\nu(z_i) = 1$, i.e. set $\hat{g}_i = \nu(z_i)^{-2} g_i$, so that $\hat{\rho}(z_i) \leq K$. By the definition of L^2 curvature radius in (4.4), one has on $\hat{g}_i$,

$$\int_{B(\hat{\rho})} |\hat{R}|^2 = c_0 \frac{volB(\hat{\rho})}{\hat{\rho}^4}, \tag{4.13}$$

where $\hat{\rho} = \hat{\rho}(z_i)$, $B(\hat{\rho}) = B_{z_i}(\hat{\rho})$. If $\hat{\rho} \leq 1$, then $volB(\hat{\rho}) \geq \mu\hat{\rho}^3$, since $\hat{\nu} = 1$, while if $\hat{\rho} > 1$, $volB(\hat{\rho}) \geq volB(1) \geq \mu$. Thus

$$\int_{B(\hat{\rho})} |\hat{R}|^2 \geq c_1, \tag{4.14}$$

with c_1 depending only on c_0. μ, and K. Hence, by scale-invariance and (4.9), there is a bounded number of such balls, with bound depending only on c_0, μ, K and Λ'. Since $\nu_i(z_i) \to 0$, as $i \to \infty$, $\rho(z_i) \to 0$, and so these balls converge metrically to a finite number of points as $i \to \infty$. ∎

Consider now the complement U. For any given $x_i \in U$, we work again in the scale $\hat{g}_i$ where $\nu = 1$, so that $\hat{\rho}(x_i) \geq K$. In this scale, the bound (4.1) becomes

$$\int_M |\hat{Ric}|^{2p} \leq \Lambda \nu_i^{4p-4} \to 0 \ \text{ as } \ i \to \infty. \tag{4.15}$$

When the L^{2p} norm of the Ricci curvature is sufficiently small, one has a volume comparison result, cf. [36] for instance, which gives

$$\hat{\nu}(y_i) \geq \mu_1 \hat{\nu}(x_i) = \mu_1, \tag{4.16}$$

for all $y_i \in B_{x_i}(K)$, where μ_1 depends only on K and μ. Thus, the ball $B_{x_i}(K)$ in the metric $\hat{g}_i$ is everywhere non-collapsed as $i \to \infty$. By (4.7), the $L^{2,2p}$ harmonic radius is uniformly bounded below on $B_{x_i}(K)$, and hence the metrics $\hat{g}_i$ are precompact in the $L^{2,2p}$ and C^α topologies. As discussed following (4.7), c_0 is chosen sufficiently small so that the metric is close to the flat metric. Since $\hat{\nu}(x_i) = 1$ and K is large, each $B_{x_i}(K)$ is thus close to a ball of radius K in a non-trivial flat manifold $\mathbb{R}^4/\Gamma$, where Γ is a discrete group of Euclidean isometries acting freely on $\mathbb{R}^4$.

Hence, associated to every point in U, there is a neighborhood of a definite size, which is diffeomorphic to $\mathbb{R}^4/\Gamma$, for some $\Gamma \neq \{e\}$, and on which one has uniform bounds on the metric in local harmonic coordinates in $C^\alpha \cap L^{2,2p}$. Following the collapse theory of Cheeger-Gromov [15], it is shown in [5] that these elementary F-structures piece together to give a global F-structure on U. (In [5], this is done in dimension 3, but the proof given works the same in all dimensions, given $C^\alpha \cap L^{2,q}$ control of the metric, for some $q > n = dimM$).

Finally, collapse at $\{x_i\}$ with locally bounded curvature in L^{2p} means $\nu(x_i) \to 0$ and the scale-invariant quanitity $(\nu(x_i))^{4p-4} \int |R|^{2p} \to 0$, as $i \to \infty$. This follows from the results established above. ■

Remark 4.3. (i). An essentially immediate consequence of the proof of Theorem 4.1 is that if

$$\inf_{\mathbb{M}_1} \int |Ric|^{2p} = 0,$$

then either M admits a Ricci-flat orbifold metric, i.e. $\Omega = M$ or $\Omega = V$, or if not, then $\Omega = \emptyset$ so that M carries an F-structure on the complement of finitely many, arbitrarily small metric balls. To see this, suppose $\Omega \neq \emptyset$. We then claim (Ω, g) is necessarily complete. The incompleteness of (Ω, g) is caused (only) by the collapse of the metric in finite distance. However, (Ω, g) is Ricci-flat, and the Bishop-Gromov volume comparison theorem rules out collapse of the volumes of local balls within finite distance. This proves that (Ω, g) is complete. On the other hand, a result of Calabi and Yau implies that a complete Ricci-flat metric on an open manifold has infinite volume, contradicting (4.2).

(ii). The lower semi-continuity of the functional implies of course that $\int_\Omega |Ric|^{2p} \leq \Lambda$. In local harmonic coordinates, the metric g is C^α and satisfies (4.8). Hence

$$R \in L^{2p}_{loc},$$

on the regular set Ω_0 of Ω.

(iii). We point out if the bound (4.1) is replaced by the stronger bound $\int |R|^{2p} \leq \Lambda$, then the same proof shows that the limit (Ω, g) has no orbifold singularities, and there are no singularities in the collapsed part either. This is because all singularities under the bound (4.1) arise from rescalings (blow-downs) of complete, non-compact Ricci-flat 4-manifolds, (ALE in case of orbifold singularities). Under a bound on $\int |R|^{2p}$, these cannot arise, since all blow-up limits under such a bound are necessarily flat, cf. (4.15).

Now consider perturbations of one of the functionals $\mathcal{F}$ in (3.1) in the direction of the L^{2p} norm of the Ricci curvature, (for example). We first consider $\mathcal{F} = \mathcal{R}ic^2$, and set

$$\mathcal{R}ic_\varepsilon^{2p} = \int |Ric|^2 + \varepsilon \int (1 + |Ric|^2)^p, \tag{4.17}$$

for $p > 1$ and $\varepsilon > 0$. Similar perturbations of $\mathcal{W}^2$ and $\mathcal{R}^2$ will be discussed later. The perturbation (4.17) is analogous to the α-energy perturbation of Sacks-Uhlenbeck [38] in their study of harmonic maps of surfaces into Riemannian manifolds; perturbations of this type for curvature functionals were studied in detail in dimension 3 in [5], [7].

The functional $\mathcal{R}ic_\varepsilon^{2p}$ is a C^∞ smooth functional on space $\mathbb{M}_1$ of metrics of volume 1. The idea is to obtain an existence result for minimizers of $\mathcal{R}ic_\varepsilon^{2p}$, and then pass to a limit $\varepsilon \to 0$, (or $p \to 1$), to obtain an existence result for minimizers of $\mathcal{R}ic^2$. To begin, the following is an immediate consequence of Theorem 4.1:

Corollary 4.4. *For any $\varepsilon > 0$ and $p > 1$, there exists a minimizing pair $(\Omega_\varepsilon, g_\varepsilon)$ for $\mathcal{R}ic_\varepsilon^{2p}$. The pair $(\Omega_\varepsilon, g_\varepsilon)$ has the properties given in Theorem 4.1 together with*

$$\mathcal{R}ic_\varepsilon^{2p}(g_\varepsilon) \leq \inf_{\mathbb{M}_1} \mathcal{R}ic_\varepsilon^{2p}. \tag{4.18}$$

■

The Euler-Lagrange equations of $\mathcal{R}ic_\varepsilon^{2p}$ on $\mathbb{M}_1$ at a metric g are:

$$D^*Dh - 2\delta^*\delta h - (\delta\delta h)g + \mathcal{P}_R = 0, \tag{4.19}$$

where $h = fRic$, $f = (1 + \varepsilon p(1 + |Ric|^2)^{p-1})$ and $\mathcal{P}_R$ is a curvature term, given by

$$\mathcal{P}_R = -2R(h) + \frac{1}{2}[|Ric|^2 + \varepsilon(1 + |Ric|^2)^p + c_\varepsilon]g, \tag{4.20}$$

where c_ε is a constant; ($c_\varepsilon = \varepsilon \inf_{\mathbb{M}_1} [(p-1)(\int (1 + |Ric|^2))^p + p(\int (1 + |Ric|^2))^{p-1})$). We note that $(\delta^*\omega)(X, Y) = \frac{1}{2}[(\nabla_X\omega)(Y) + (\nabla_Y\omega)(X)]$, so that $\delta\omega = -tr\delta^*\omega$.

This formula is easily derived from standard formulas, cf. [12]. Briefly, the variation of the integrand of $\mathcal{R}ic_\varepsilon^{2p}$ is

$$(1 + 2\varepsilon p(1 + |Ric|^2)^{p-1})[2\langle Ric'(g'), Ric\rangle - 2\langle Ric^2, g'\rangle] + \tfrac{1}{2}[|Ric|^2 + \varepsilon(1 + |Ric|^2)^p + c]\langle g', g\rangle.$$

One has

$$2Ric'(k) = D^*Dk - 2\delta^*\delta k - D^2(trk) - 2R(k) + Ric \circ k + k \circ Ric, \tag{4.21}$$

and if $(Ric')^*$ denotes the adjoint of Ric', then

$$2(Ric')^*(k) = D^*Dk - 2\delta^*\delta k - (\delta\delta k)g - 2R(k) + Ric \circ k + k \circ Ric. \tag{4.22}$$

Combining the various terms gives (4.19).

When $\varepsilon = 0$, $h = Ric$ and (4.19) becomes the Euler-Lagrange equation $\nabla\mathcal{R}ic^2 = 0$, i.e.

$$D^*DRic - 2\delta^*\delta Ric - (\delta\delta Ric)g - 2R(h) + \tfrac{1}{2}|Ric|^2 g = 0. \tag{4.23}$$

Now a minimizer $(\Omega_\varepsilon, g_\varepsilon)$ from Corollary 4.4 is a weak $L^{2,2p}$ solution of the Euler-Lagrange equation (4.19), i.e. (4.19) holds when viewed as a distribution and paired with any $L^{2,2p^*}$ symmetric bilinear form ψ, of compact support in Ω_ε; $2p^* = 1 - \frac{1}{2p}$ is the conjugate exponent to $2p$. Note that $h \in L^{2p/2p-1}$, since $Ric \in L^{2p}$. For the remainder of this section, we assume $p > 1$ with p close to 1.

In fact, the regularity can be improved a little for metrics g which are local minima of $\mathcal{R}ic_\varepsilon^{2p}$, in the sense that g minimizes $\mathcal{R}ic_\varepsilon^{2p}$ among nearby, compact perturbations of g.

Lemma 4.5. *Let g be an $L^{2,2p}$ metric which locally minimizes $\mathcal{R}ic_{\varepsilon}^{2p}$ on a domain U, with ε small. Then,*

$$Ric \in L^4_{loc}, \tag{4.24}$$

so that g is locally in $L^{2,4}$.

Proof: This is proved in [46] for metrics locally minimizing the L^{2p} norm of the curvature R, and the proof for $\mathcal{R}ic_{\varepsilon}^{2p}$ is essentially the same; thus we will be somewhat brief and refer to [46] for further details.

The local Ricci flow $\frac{d}{dt}g(t) = -2\chi g(t)$, where χ is a local cutoff function supported in U, is defined for metrics in $L^{2,2p}$. Using the local minimizing property of $g = g(0)$, it suffices to show that

$$\frac{d}{dt}\mathcal{R}ic_{\varepsilon}^{2p}(g(t))|_{t=0} + c(\int \chi^2|Ric|^4)^{1/2} \leq C[1 + \mathcal{R}ic_{\varepsilon}^{2p}(g)].$$

It follows from (4.21) and the Bianchi identity $\delta Ric = -\frac{1}{2}ds$, that

$$\frac{d}{dt}\mathcal{R}ic_{\varepsilon}^{2p}(g(t)) \leq -\int\langle D^*D\chi Ric, fRic\rangle + C[1 + \int \chi|Ric|^{2p}|R|],$$

where f is defined as following (4.19). It is important for the following to observe that $f \geq 1$. One has $-\int\langle D^*D\chi Ric, fRic\rangle = -\int\langle D\chi Ric, DfRic\rangle$. Expanding this out and using the Cauchy-Schwarz and Young inequalities, $(ab \leq \mu a^2 + \mu^{-1}b^2)$, together with the fact that ε is small, gives $-\int\langle D\chi Ric, DfRic\rangle \leq -\int \chi|DRic|^2 + C[1 + \int \chi|Ric|^{2p+1}]$, so that

$$\frac{d}{dt}\mathcal{R}ic_{\varepsilon}^{2p}(g(t))_{t=0} + c\int \chi|DRic|^2 \leq C[1 + \int \chi|Ric|^{2p}|R|].$$

Since $|d|Ric||^2 \leq c|DRic|^2$, the result then follows easily from the local Sobolev embedding, $L^4 \subset L^{1,2}$ in dimension 4, via use of Remark 4.3(ii) and the Hölder inequality on the term $|Ric|^{2p}|R| \leq |Ric|^2|R|^{2p-1}$. ∎

An important point at this stage is to see that weak solutions of (4.19) are smooth.

Proposition 4.6. *Any locally defined weak $L^{2,2p}$ solution g of (4.19), which is a local minimizer of $\mathcal{R}ic_{\varepsilon}^{2p}$ is C^{∞} smooth.*

Proof: The idea is of course to use elliptic regularity results to boost the regularity of g. By (4.24) and Remark 4.3(ii), the term $\mathcal{P}_R \in L^{4/2p}$, so that (4.19) has the form

$$L(h) = D^*Dh - 2\delta^*\delta h - (\delta\delta h)g = -\mathcal{P}_R \in L^{4/2p}. \tag{4.25}$$

However, the linear operator L is not elliptic, due to its invariance properties under the diffeomorphism group. To deal with this, any symmetric bilinear form $\psi \in T_g\mathbb{M}$ on M may be decomposed as

$$\psi = \delta^*X + \phi g + k, \tag{4.26}$$

where X is a vector field, ϕ a function and k is transverse-traceless, $\delta k = trk = 0$. This also holds locally, and so $h = fRic \in L^{p'}$, $p' = 4/(2p-1)$ may be written as a sum as in (4.26), with δ^*X, ϕ and k in $L^{p'}$.

The equation (4.25) thus decomposes as

$$L(k) + L(\delta^*X) + L(\phi g) \in L^{4/2p}. \tag{4.27}$$

One has $L(k) = D^*Dk$, which is of course an elliptic operator on k. To compute $L(\delta^*X)$, note that $\mathcal{L}_X(Ric) = Ric'(\delta^*(X))$, while in local coordinates, $\mathcal{L}_X(Ric) \sim X(Ric) + (\partial X)Ric$. The second term is in $L^{4/2p}$ while the first term is in $L^{-1,q}$, where $q = 4$. Thus $\mathcal{L}_X(Ric) \in L^{-1,q}$. On the other hand, one has $2(Ric')^*(\beta) = 2Ric'(\beta) + D^2 tr\beta - (\delta\delta\beta)g$, (see (4.21)-(4.22)). It follows that

$$L(\delta^* X) = -\delta\delta\delta^* Xg - D^2\delta X + L^{-1,q},$$

where have used $tr\delta^* X = -\delta X$, and $+L^{-1,q}$ denotes the addition of an element in $L^{-1,q}$. A simple direct computation gives

$$L(\phi g) = -2\Delta\phi g + 2D^2\phi.$$

Now by a standard Weitzenbock formula,

$$\delta\delta^* X = D^* DX = \delta dX + d\delta X - Ric(X),$$

so that $\delta\delta\delta^* X = \delta d\delta X - \delta(Ric(X))$. One may decompose X as

$$X = \nabla\omega + Y,$$

where $\delta Y = 0$. In sum, combining these computations gives

$$L(\delta^* X) + L(\phi g) = -(\Delta\Delta\omega + 2\Delta\phi)g + D^2(\Delta\omega + 2\phi) + L^{-1,q}. \tag{4.28}$$

Now consider first the trace of the equation (4.27). Since k is transverse-traceless, via (4.28) this gives

$$\Delta\Delta\omega + 2\Delta\phi \in L^{-1,q}.$$

The coefficients of the Laplacian Δ are in C^α in local harmonic coordinates. By elliptic regularity, cf. [31], it follows that $\Delta\omega + 2\phi \in L^{1,q}$, and hence

$$L(\delta^* X) + L(\phi g) \in L^{-1,q}.$$

Returning to (4.27), we then have $L(k) \in L^{-1,q}$, so that again by elliptic regularity,

$$k \in L^{1,q}, \tag{4.29}$$

giving the main initial regularity boost.

Since $\Delta\omega, \phi \in L^{p'}$, while $\Delta\omega + 2\phi \in L^{1,q}$, if $\Delta\omega$ and ϕ are linearly independent, it follows that $\Delta\omega \in L^{1,q}$ and $\phi \in L^{1,q}$. In fact, this is case unless $\Delta\omega + 2\phi \equiv 0$. Suppose first that $\Delta\omega + 2\phi \neq 0$, so that $\Delta\omega \in L^{1,q}$, $\phi \in L^{1,q}$, and hence $trh = \Delta\omega + 4\phi \in L^{1,q}$. Thus, one has

$$h = \delta^* Y + L^{1,q}. \tag{4.30}$$

To show that $h = fRic \in L^{1,q}$, take the exterior derivative of (4.30), giving

$$dh = fdRic + df \wedge Ric = d\delta^* Y + L^q = R(Y) + L^q \in L^{q'},$$

where the terms are 2-forms with values in the tangent bundle and $q' < 4$. Taking the trace on the last two indices and using the Bianchi identity $\delta Ric = -\frac{1}{2}ds$, one obtains

$$trdh = dtrh + \delta h = \frac{1}{2}d(trh) - E(df) \in L^{q'},$$

and so $\delta h \in L^{q'}$, since $dtrh \in L^q$.

On the other hand, from (4.26) with $\psi = h$ and the Weitzenbock formula, one has $\delta h = \delta dY - d\Delta\omega - d\phi - Ric(X) = \delta dY + L^{q'}$. It follows that $\delta dY \in L^{q'}$. Since $\delta Y = 0$, this gives $Y \in L^{2,q'}$, and hence $h \in L^{1,q'}$. In turn, this now implies $Ric \in L^{1,(2p-1)q'}$, so that $g \in L^{3,(2p-1)q'} \subset C^{2,\alpha}$ in local harmonic coordinates.

Suppose instead $\Delta\omega + 2\phi = 0$. Then in place of (4.30) one has

$$h = \delta^* X - \tfrac{1}{2}(\Delta\omega)g + L^{1,q}.$$

Via the Weitzenbock formula, this gives $\delta\delta h = -\Delta\Delta\omega + L^{-1,q}$. Also, as below (4.30), one has $dtrh + \delta h = -\frac{3}{2}d\Delta\omega + L^{q'}$, so that $-\Delta trh + \delta\delta h = \frac{3}{2}\Delta\Delta\omega + L^{-1,q'}$. Since $trh = -\Delta\omega$, this gives $\Delta\Delta\omega \in L^{-1,q'}$, and hence $\Delta\omega \in L^{q'}$. It follows then as before that $h \in L^{1,q'}$.

One may now repeat the arguments above inductively, improving the regularity of h and g at each step, to obtain $g \in C^{\infty}$. ■

Conjecture 4.7. The minimizers of $\mathcal{R}ic_{\varepsilon}^{2p}$ are complete, i.e. g_{ε} is complete on each component of Ω_{ε}, and Ω_{ε} has only finitely many components.

Results of this type are proved for minimizers of analogous functionals in dimension 3 in [5]. However, it seems difficult to extend the proof in the 3-dimensional case to 4-dimensions.

Consider now the analogous procedure for either $\mathcal{R}^2$ or $\mathcal{W}^2$, i.e. the perturbations

$$\int |R|^2 + \varepsilon \int (1+|R|^2)^p, \quad \text{or} \quad \int |W|^2 + \varepsilon \int (1+|W|^2)^p. \tag{4.31}$$

Corollary 4.4 obviously holds for the first of these perturbed functionals, (via Theorem 4.1), but it is unknown if it holds for the second. Even if it did, it is not clear if Proposition 4.6 holds for either of the functionals in (4.31). For example, the Euler-Lagrange equation for the perturbation of $\mathcal{R}^2$ has a term of the form

$$(1+\varepsilon p(1+|R|^2)^{p-1})\langle R', R\rangle = (R')^*[(1+\varepsilon p(1+|R|^2)^{p-1}R].$$

This has a much more complicated form than (4.19), and the proof of Proposition 4.6 will not apply directly. The same remarks apply to the functional $\mathcal{W}^2$. A similar difficulty remains if one perturbs either functional by $(1+|Ric|^2)^p$ in place of $(1+|R|^2)^p$ or $(1+|W|^2)^p$. Instead, we use a slightly different path.

By the Chern-Gauss-Bonnet theorem (3.3), minimizing $\mathcal{R}^2$ is the same as minimizing $\mathcal{Z}^2$; these two functionals also have the same Euler-Lagrange equations or critical points. Similarly, minimizers or critical points of $\mathcal{W}^2$ are the same as those of $\mathcal{Z}^2$ - $\frac{1}{12}\mathcal{S}^2 = \mathcal{R}ic^2$ - $\frac{1}{3}\mathcal{S}^2$, cf. (3.4). Thus, consider first the functional

$$\mathcal{Z}_{\varepsilon}^{2p}(g) = \int (|Ric|^2 - \frac{1}{4}s^2) + \varepsilon \int (1+|Ric|^2)^p. \tag{4.32}$$

For $\varepsilon > 0$ small, this is a small perturbation of $\mathcal{Z}^2$, with the property that a bound on $\mathcal{Z}_{\varepsilon}^{2p}$ implies a bound on $\mathcal{R}ic^{2p}$, as in (4.1).

The Euler-Lagrange equation for $\mathcal{Z}_{\varepsilon}^{2p}$ is

$$D^*Dh - 2\delta^*\delta h - (\delta\delta h)g - \tfrac{1}{2}(D^2 s - (\Delta s)g) + \mathcal{P}_Z = 0, \tag{4.33}$$

where $h = fRic$, $f = (1+\varepsilon p(1+|Ric|^2)^{p-1})$, with $\mathcal{P}_Z$ given by

$$\mathcal{P}_Z = -2R(h) + \tfrac{1}{2}sRic + \frac{1}{2}(\zeta + c_{\varepsilon})g, \tag{4.34}$$

where $\zeta = |Ric|^2 + \varepsilon(1+|Ric|^2)^p$.

To derive (4.33), for f and ζ as above, the variation of the integrand of $\mathcal{Z}_{\varepsilon}^{2p}$ is

$$2f[\langle Ric', Ric\rangle - \langle Ric^2, g'\rangle] - \tfrac{1}{2}ss' + \tfrac{1}{2}(\zeta + c_{\varepsilon})\langle g', g\rangle.$$

One has $2f\langle Ric', Ric\rangle = 2(Ric')^*(fRic)$, and similarly $ss' = (s')^*(s)$. The first term is given by (4.22) while the second term is $D^2 s - (\Delta s)g - sRic$, and combining these expressions gives (4.33).

As before, when $\varepsilon = 0$, $h = Ric$ and (4.33) becomes the Euler-Lagrange equation $\nabla\mathcal{Z}^2 = 0$, i.e.

$$D^*Dz - 2\delta^*\delta z - (\delta\delta z)g - 2R(z) + \tfrac{1}{2}|z|^2 g = 0. \tag{4.35}$$

Corollary 4.8. *Corollary 4.4 and Proposition 4.6 hold for the functional* $\mathcal{Z}_{\varepsilon}^{2p}$.

Proof: The proof of Corollary 4.4 via Theorem 4.1 is the same as before, using the fact noted above that a bound on $\mathcal{Z}_\varepsilon^{2p}$ implies a bound on Ric^{2p}. It is straightforward to show that Lemma 4.5 also holds for $\mathcal{Z}_\varepsilon^{2p}$ in place of $\mathcal{R}ic_\varepsilon^{2p}$. The only difference is that the term $|DRic|^2$ is replaced by $|DRic|^2 - \frac{1}{4}|ds|^2$. However, $Ric = z + \frac{s}{4}g$ and $|Dz|^2 \geq \frac{1}{4}(\delta z)^2 = \frac{1}{64}|ds|^2$, so that $|DRic|^2 - \frac{1}{4}|ds|^2 \geq \frac{1}{64}|ds|^2$. Given this adjustment, the rest of the proof carries over as before.

Moreover, the proof of Proposition 4.6 is also essentially the same as before. The addition of the s terms in (4.33) implies that (4.28) is changed to

$$\begin{aligned} L(\delta^* X) + L(\phi g) - \tfrac{1}{2}(D^2 s - \Delta s g) \\ = D^2(\Delta\omega + 2\phi - \tfrac{1}{2}s) - (\Delta\Delta\omega + 2\Delta\phi - \tfrac{1}{2}\Delta s)g. \end{aligned} \tag{4.36}$$

Given this, the rest of the proof remains valid, with only minor changes. ■

Next, we treat the case of $\mathcal{W}^2$ in a similar way. Among various possible perturbations, consider

$$\mathcal{W}_{\varepsilon,\lambda}^{2p}(g) = \int (|Ric|^2 - \frac{1}{3}s^2) + \varepsilon \int (1 + |Ric|^2)^p + \lambda \int s^2 = \mathcal{R}ic_\varepsilon^{2p} - (\frac{1}{3} - \lambda)\mathcal{S}^2, \tag{4.37}$$

which, modulo $\chi(M)$, gives $2\mathcal{W}^2$ when $\varepsilon = \lambda = 0$. As will be seen later, the relations between $\varepsilon > 0$, $p > 1$ and $\lambda > 0$ determine the choice of conformal gauge as $\varepsilon \to 0$. It is possible in the following to dispense with λ, i.e. set $\lambda = 0$, but we do not do so since a suitable choice of $\lambda > 0$ leads to a Yamabe-type gauge in the limit $\varepsilon = 0$. As discussed in Example 3.1, a good choice of conformal gauge is of some importance. It is clear that a bound on $\mathcal{W}_{\varepsilon,\lambda}^{2p}$ implies a bound on $\mathcal{R}ic^{2p}$.

The Euler-Lagrange equation for $\mathcal{W}_{\varepsilon,\lambda}^{2p}$ is

$$D^*Dh - 2\delta^*\delta h - (\delta\delta h)g - 2(\tfrac{1}{3} - \lambda)(D^2 s - (\Delta s)g) + \mathcal{P}_W = 0, \tag{4.38}$$

where $h = fRic$ is as in (4.19). The curvature term $\mathcal{P}_W$ is

$$\mathcal{P}_W = -2R(h) + 2(\tfrac{1}{3} - \lambda)sz + \tfrac{1}{2}[|Ric|^2 + \varepsilon(1 + |Ric|^2)^p + c_\varepsilon]g. \tag{4.39}$$

The derivation of (4.38) is straightforward, given the derivations of (4.19) and (4.33). A little computation shows that the trace of (4.38) is given by

$$-\Delta trh - 2\delta\delta h + (2 - 6\lambda)\Delta s - 2\varepsilon p(1 + |Ric|^2)^{p-1}|Ric|^2 + 2[\varepsilon(1 + |Ric|^2)^p + c_\varepsilon]. \tag{4.40}$$

When $\varepsilon = \lambda = 0$, $h = Ric$ and $f = 1$, and so (4.38) becomes

$$D^*DRic + \tfrac{1}{3}D^2 s + \tfrac{1}{6}\Delta s g - 2R(Ric) + \tfrac{2}{3}sRic + \tfrac{1}{2}[|Ric|^2 - \tfrac{1}{3}s^2]g = 0. \tag{4.41}$$

The equation (4.41) is of course the conformally invariant Bach equation, i.e. the Euler-Lagrange equations $\nabla\mathcal{W}^2 = 0$ for $\mathcal{W}^2$, cf. [12].

We are now in position to prove:

Corollary 4.9. *Corollary 4.4 and Proposition 4.6 hold for the functional* $\mathcal{W}_{\varepsilon,\lambda}^{2p}$.

Proof: As noted above, a bound on $\mathcal{W}_{\varepsilon,\lambda}^{2p}$ implies a bound on $\mathcal{R}ic^{2p}$, so Theorem 4.1 and Corollary 4.4 hold for $\mathcal{W}_{\varepsilon,\lambda}^{2p}$.

To verify that (4.24) holds, write $\mathcal{W}_{\varepsilon,\lambda}^{2p}$ as

$$\mathcal{W}_{\varepsilon,\lambda}^{2p} = 2\mathcal{W}^2 - 16\pi^2\chi(M) + \varepsilon \int (1 + |Ric|^2)^p + \lambda \int s^2.$$

The three leading order terms in the Euler-Lagrange equation (4.41) for $\mathcal{W}^2$ may be written in the form $\delta d(Ric - \frac{s}{6}g)$. In place of the local Ricci flow, deform the metric locally in the direction

$-2(Ric - \frac{s}{6}g)$, i.e. $\frac{d}{dt}g = -2\chi(Ric - \frac{s}{6}g)$. Using the results in [18], the proof that this local flow exists is the same as that for the local Ricci flow. As in the proof of Lemma 4.5, one then has

$$\frac{d}{dt}\mathcal{W}^2(g(t)) \leq C(1 + \int \chi|R|^3).$$

On the other hand, the same estimates as before in the proof of Lemma 4.5 hold for the variation of functional $\varepsilon \int (1+|Ric|^2)^p$ in the direction of the modified local Ricci flow, $(-2\chi(Ric - \frac{s}{6}g))$, in place of the variation of $\mathcal{R}ic_\varepsilon^{2p}$ or $\mathcal{Z}_\varepsilon^{2p}$ along the local Ricci flow; (this uses the fact that $\frac{1}{6} < \frac{1}{4}$ from the proof of Corollary 4.8). The variation of $\lambda\mathcal{S}^2$ contributes only lower order terms, and so the estimates are unaffected. One then obtains (4.24) in the same way as before.

It is also straightforward to see that the proof of Proposition 4.6 for $\mathcal{W}_{\varepsilon,\lambda}^{2p}$ proceeds exactly as before in the case of $\mathcal{R}ic_\varepsilon^{2p}$ or $\mathcal{Z}_\varepsilon^{2p}$. ■

The same results hold for the functionals $\mathcal{W}_\pm^2$, which differ from $\mathcal{W}^2$ just by a topological term, (as with $\mathcal{R}^2$ and $\mathcal{Z}^2$).

Summarizing the work above, we have now produced suitable perturbations of the functionals $\mathcal{F}$ in (3.1), and proved the existence of minimizing configurations for each of them. As mentioned above, the idea now is to take a sequence $\varepsilon = \varepsilon_i \to 0$, and consider the behavior of (subsequences of) a sequence $(\Omega_\varepsilon, g_\varepsilon)$ of minimizing pairs in the limit $\varepsilon \to 0$. Although the metrics g_ε are C^∞ smooth, one no longer has uniform control of the L^{2p} norm of the Ricci curvature of g_ε, so that Theorem 4.1 does not apply. Nevertheless, using the Chern-Gauss-Bonnet theorem, together with the smoothness of g_ε and the fact that g_ε satisfies an (essentially) elliptic Euler-Lagrange equation, we show that the conclusions of Theorem 4.1 do in fact hold. As seen above and especially in Example 3.1, the cases $\mathcal{R}ic^2$, $\mathcal{R}^2$ and the cases $\mathcal{W}^2$, $\mathcal{W}_\pm^2$ are somewhat different, and so these two situations are treated separately.

Theorem 4.10. *(Geometric Decomposition with respect to $\mathcal{R}ic^2, \mathcal{R}^2$.)*

Let M be a closed, oriented 4-manifold and let $\mathcal{F}$ be one of the functionals $\mathcal{R}ic^2$, $\mathcal{R}ic^2$ on $\mathbb{M}_1$. Then minimizers of $\mathcal{F}$ on $\mathbb{M}_1$ are realized in the idealized sense that there exist minimizing sequences $\{g_i\}$ converging in the pointed Gromov-Hausdorff topology to one of the following configurations:

(I). *A compact, oriented, possibly reducible orbifold (V, g_0) associated to M, with C^∞ metric g_0 on the regular set V_0. The metric g_0 extends C^0 across the orbifold singularities and*

$$vol_{g_0}V = 1. \tag{4.42}$$

Further

$$\mathcal{F}(g_0) + \sum_k \mathcal{F}(N_k, g_\infty^k) = \inf_{g\in\mathbb{M}_1} \mathcal{F}(g), \tag{4.43}$$

where the sum is over the collection of ALE spaces (N_k, g_∞^k) associated with the singularities of V, cf. also (4.65).

(II). *A maximal orbifold domain $\Omega \subset\subset M$, possibly reducible and possibly empty, with C^∞ smooth metric g_0 on the regular set Ω_0. The metric g_0 extends C^0 across the orbifold singularities, and satisfies*

$$vol_{g_0}V \leq 1, \tag{4.44}$$

together with

$$\mathcal{F}(g_0) + \sum_k \mathcal{F}(N_k, g_\infty^k) \leq \inf_{g\in\mathbb{M}_1} \mathcal{F}(g), \tag{4.45}$$

where the sum is as in (4.43).

Outside a sufficiently large compact set $K \subset \Omega$, Ω carries an F-structure along which g_0 collapses with locally bounded curvature on approach to $\partial\Omega$. The complement $M \setminus K$ also carries an F-structure outside a finite number of balls $B_{z_j}(\varepsilon_i)$ with respect to g_i, where $\varepsilon_i \to 0$ as $i \to \infty$. If K_i is an exhaustion of Ω, then $M \setminus K_i$ collapses everywhere, and collapses with locally bounded curvature away from the singular points $\{z_j\}$.

In all cases, the metric g_0 satisfies the Euler-Lagrange equation

$$\nabla \mathcal{F} = 0, \tag{4.46}$$

and minimizes $\mathcal{F}$ among compact perturbations.

In comparing this result with Theorem 1.1, one should note that it is possible that $\Omega = \emptyset$. Thus, (II) above includes both cases (II) and (III) of Theorem 1.1.

Proof: By Proposition 4.6 and Corollaries 4.4 and 4.8, there exist minimizing pairs $(\Omega_\varepsilon, g_\varepsilon)$ for $\mathcal{F}$, for any given $\varepsilon > 0$. As in the proof of Theorem 4.1, consider a thick-thin decomposition of Ω_ε. As before, the discussion below applies to a sequence $(\Omega_{\varepsilon_i}, g_{\varepsilon_i})$ with $\varepsilon_i \to 0$, but we will usually drop the subscript from the notation. For any given $\nu_0 > 0$, let

$$\Omega_\varepsilon^{\nu_0} = \{x \in \Omega_\varepsilon : \nu(x) > \nu_0\}. \tag{4.47}$$

For any given $\rho_0 > 0$, consider the subdomain

$$\Omega_\varepsilon^{\nu_0,\rho_0} = \{x \in \Omega_\varepsilon^{\nu_0} : \rho(x) > \rho_0\}, \tag{4.48}$$

where ρ is the L^2 curvature radius.

We claim that on $\Omega_\varepsilon^{\nu_0,\rho_0}$, one has C^∞ smooth convergence (in a subsequence) to a limit $\Omega_0^{\nu_0,\rho_0}$, away from the boundary. The Euler-Lagrange equations for g_ε, (4.19) or (4.33), are (essentially) uniformly elliptic as $\varepsilon \to 0$. More precisely, although (4.19) or (4.25) is not elliptic, the proof of Proposition 4.6 shows that uniform elliptic estimates hold on each summand in (4.26), independent of ε. The same applies to the Euler-Lagrange equation (4.33) for $\mathcal{Z}_\varepsilon^{2p}$.

However, the elliptic regularity estimates require, to get started, control on the C^α norm of the leading order coefficients, for some fixed $\alpha > 0$. Thus, uniform estimates require uniform control of the metric locally in C^α, (in harmonic coordinates). Since there is no longer a uniform bound on $\mathcal{R}ic^{2p}$, but only a uniform bound on $\mathcal{R}^2$, one thus needs some stronger initial control on g_ε in order to proceed. We note that Lemma 4.5 is not uniform in ε, since, for instance, it requires uniform control on the local Sobolev constant.

To obtain this initial control, let $\rho^{(q)}$ be the L^q curvature radius, $q > 2$, with fixed parameter $c_0 = c_0(q)$ in (4.4). We first claim that if $c_0 = c_0(2)$ is sufficiently small, then there is a constant $\delta_0 > 0$, depending only on $c_0(2)$, $c_0(q)$ and ν_0, such that

$$\rho^{(q)}(x) \geq \delta_0 \rho(x), \tag{4.49}$$

for all $x \in \Omega_\varepsilon^{\nu_0,\rho_0}$. This gives uniform local $L^{2,q}$ and so $C^\alpha \cap L^{1,k}$ control of the metric in harmonic coordinates on $\Omega_\varepsilon^{\nu_0,\rho_0}$, with $\alpha > 0$, $k > 4$. Thus, the proof of Proposition 4.6 applies uniformly as $\varepsilon \to 0$, which gives the claim above of smooth convergence.

The proof of (4.49) is by contradiction. If (4.49) is not true, then there exists $x \in \Omega_\varepsilon^{\nu_0,\rho_0}$ such that $\rho^{(q)}(x) \leq \delta\rho(x)$, where δ is arbitrarily small. Without loss of generality, we may assume that x realizes, or almost realizes, the minimal value of the ratio $\rho(x)/dist(x, \partial\Omega_\varepsilon^{\nu_0,\rho_0})$. Work in the scale $\hat{g}_\varepsilon = \rho^{(q)}(x)^{-2} g_\varepsilon$ where $\hat{\rho}^{(q)}(x) = 1$, so that $\hat{\rho}(x) >> 1$ and $\nu(x) >> 1$. Now in this scale, one does have uniform local control of $\hat{g}_\varepsilon$ in $L^{2,q}$, (independent of ε), and so in C^α in harmonic coordinates. Hence, as noted above following (4.48), $\hat{g}_\varepsilon$ is thus uniformly controlled in C^k, for any k.

Now since $\hat{\rho}^{(q)}(x) = 1$, the metric $\hat{g}_\varepsilon$ has a definite amount of curvature in L^q on $B_x(1)$, depending only on the choice of $c_0(q)$. However, $\rho(x) >> 1$, and $c_0(2)$ is very small, so that the curvature is very small in L^2 on $B_x(1)$. Since the metric $\hat{g}_\varepsilon$ is uniformly controlled in C^k, for any k, this is impossible, and so proves (4.49). Clearly, if $c_0(q)$ is fixed, the size of $c_0(2)$ may be explicitly estimated in terms of $c_0(q)$.

Next consider the complementary region of $\Omega_\varepsilon^{\nu_0}$ where $\rho \leq \rho_0 \leq \nu_0$, where ρ_0 is (arbitrarily) small, i.e.

$$(\Omega_\varepsilon)_{\rho_0}^{\nu_0} = \{x \in \Omega_\varepsilon^{\nu_0} : \rho(x) \leq \rho_0\}. \tag{4.50}$$

Then Lemma 4.2, in the scale g_ε, (not $\hat{g}_\varepsilon$), shows that there is at most a bounded number of points $q_j \in \Omega_\varepsilon^{\nu_0}$ such that $(\Omega_\varepsilon)_{\rho_0}^{\nu_0} \subset B_{q_j}(\rho_0)$. Thus, there is bounded number of points, independent of ε, where the curvature R can concentrate in L^2. In fact, the bound is independent of ν_0, and depends only on c_0 and μ.

It follows that if ρ_ε is any sequence such that $\rho_\varepsilon \to 0$, as $\varepsilon \to 0$, $(\Omega_\varepsilon)_{\rho_\varepsilon}^{\nu_0}$ converges metrically to a finite number of points. We claim that in the limit, each of these point singularities is an orbifold singularity, possibly reducible, and possibly with $\Gamma = \{e\}$. To prove this, we use the orbifold compactness theorem of [8], which, given the L^2 bound on the curvature R from the Chern-Gauss-Bonnet theorem, and the lower bound ν_0 on the volume radius, shows that it suffices to have the following small curvature estimate on $(\Omega_\varepsilon^{\nu_0}, g_\varepsilon)$: if $\int_{B(r)} |R|^2 \leq \delta$, for some fixed δ small, then there is a constant C, independent of δ, such that

$$\sup_{B(\frac{1}{2}r)} |R|^2 \leq \frac{C}{volB(r)} \int_{B(r)} |R|^2. \tag{4.51}$$

Since we are working in Ω^{ν_0} for a fixed (but arbitrary) ν_0, this is equivalent to showing that the curvature is bounded in L^∞ in balls of radius ρ, when such ρ balls are scaled to radius $\rho = 1$; (the estimate in (4.51) is scale-invariant). However, this has already been done; (4.49) gives a lower bound on the L^q curvature radius, and higher order control follows as before via elliptic regularity as in Proposition 4.6.

Suppose there exists $\nu_0 > 0$ such that $\nu(x) \geq \nu_0$, for all $x \in \Omega_\varepsilon$ as $\varepsilon = \varepsilon_i \to 0$. Then Ω_ε is a compact orbifold V_ε associated to M and the analysis above shows that $(V_\varepsilon, g_\varepsilon)$ converges, in a subsequence, to a limit orbifold (V, g_0), smoothly away from orbifold singularities. The equation (4.42) is obvious from the smooth convergence while the equation (4.43) follows from the proof of the orbifold compactness theorem [8] above, together with the work in [9] and [11]; this issue is discussed further below, (following Theorem 4.15). This completes the proof in Case (I).

Next, suppose there exist $x_i \in \Omega_\varepsilon$ such that $\nu(x_i) \to 0$ as $\varepsilon \to 0$. One may then choose a sequence $\nu_j \to 0$ and consider the domain $(\Omega_\varepsilon^{\nu_j}, g_\varepsilon)$. As in the proof of Theorem 4.1, a diagonal subsequence of (i, j) converges in the pointed Gromov-Hausdorff topology to a limit maximal orbifold domain (Ω, g_0). For the same reasons as above, the limit satisfies the properties claimed in Theorem 4.10. The equalities in (4.42) and (4.43) are replaced however by inequalities, since part of the volume and part of the value of $\mathcal{F}(g_\varepsilon)$ may be contained in the collapsing region.

To complete the proof, consider the structure of the complementary, collapsing region

$$(\Omega_\varepsilon)_{\nu_0} = \{x \in \Omega_\varepsilon : \nu(x) \leq \nu_0\}, \tag{4.52}$$

where ν_0 is small, and may be assumed to be arbitrarily small for i sufficiently large. As in (4.12), form the two subdomains

$$U_\varepsilon = \{x \in (\Omega_\varepsilon)_{\nu_0} : \rho(x) > K\nu(x)\}, \text{ and } W_\varepsilon = \{x \in (\Omega_\varepsilon)_{\nu_0} : \rho(x) \leq K\nu(x)\}. \tag{4.53}$$

The proof that W tends to finitely many points in Lemma 4.2 holds without any changes here also. Regarding U_ε, although (4.15) does not hold in the present situation, in the scale $\hat{g}_\varepsilon = \nu(x_i)^{-2} g_\varepsilon$ where $\hat{\nu}(x_i) = 1$, one has $\hat{\rho}(x_i) \geq K$. Hence, as described above in and following (4.49), the metric $\hat{g}_\varepsilon$ is smoothly close to a flat metric on the ball $B_{x_i}(K-1)$. For the same reasons as in the proof of Theorem 4.1 following (4.16), such a ball is close to a ball in a non-trivial flat manifold $\mathbb{R}^4/\Gamma$, giving rise to an elementary F-structure. Again, as previously, these elementary F-structures patch together to give a global F-structure on U_ε. The same arguments based on Theorem 4.1 also show that there exist minimizing sequences $\{g_i\}$ for $\mathcal{F}$ on M such that $M \setminus K$ carries an F-structure outside a bounded number of ε_j-balls, with $\varepsilon_j \to 0$ as $\varepsilon \to 0$, for $K \subset \Omega$ sufficiently large.

It is clear that the metric g_0 on V or Ω satisfies the Euler-Lagrange equations

$$\nabla \mathcal{F} = 0$$

given by (4.19) or (4.33). Further, it is also clear from the construction that the configurations $(\Omega_\varepsilon, g_\varepsilon)$ are pointed Gromov-Hausdorff limits of sequences of unit volume metrics on M. ■

Next consider the functionals $\mathcal{W}^2$, or $\mathcal{W}^2_\pm$. Although we conjecture that Theorem 4.10 also holds for these functionals, there is a basic obstacle to proving this. Namely, given an (arbitrary) compact 4-manifold, it is unknown if there exists a minimizing sequence $\{g_i\}$ for $\mathcal{W}^2$ such that

$$\int_M s^2_{g_i} \leq \Lambda, \tag{4.54}$$

for some (arbitrary) $\Lambda < \infty$. Since $\mathcal{W}^2$ is conformally invariant, one may choose g_i to be Yamabe metrics, in which case (4.54) is equivalent to

$$s_{g_i} \geq -\Lambda' > -\infty. \tag{4.55}$$

If M has no such minimizing sequence, there seems little hope (at present) of proving the existence of a generalized metric realizing $\inf \mathcal{W}^2$. On the other hand, we conjecture that (4.54) always holds. Of course the bound (4.54) or (4.55) is equivalent to a bound on $\mathcal{R}^2$, given a bound on $\mathcal{W}^2$, via the Chern-Gauss-Bonnet theorem.

Thus, in the following, we essentially assume (4.54). More precisely, since we are working with special minimizing sequences $(\Omega_{\varepsilon,\lambda}, g_{\varepsilon,\lambda})$ obtained by minimizing $\mathcal{W}^{2p}_{\varepsilon,\lambda}$, we impose (4.54) on $(\Omega_{\varepsilon,\lambda}, g_{\varepsilon,\lambda})$, for a suitable choice of ε, λ. Due to the conformal invariance of the $\varepsilon = 0$ limit, one should impose a gauge condition which minimizes $\mathcal{S}^2$ in its conformal class, (see again Example 3.1). As will seen below, this is implied by the condition that as $i \to \infty$,

$$\varepsilon_i \to 0, \quad \lambda_i \to 0 \quad \text{and} \quad \varepsilon_i/\lambda_i \to 0. \tag{4.56}$$

One may either keep p fixed, or let $p_i \to 1$. A resulting sequence $(\Omega_i, g_i) = (\Omega_{\varepsilon_i,\lambda_i}, g_{\varepsilon_i,\lambda_i})$ is then called *preferred* if there exists a constant $\Lambda < \infty$ such that

$$\int_{\Omega_i} s^2_{g_i} \leq \Lambda. \tag{4.57}$$

The existence of a preferred minimizing sequence clearly depends only on the diffeomorphism type of M. In fact, it is easily seen to be equivalent to the following condition: for each i, there exists δ_i, with $\delta_i \to 0$ as $i \to \infty$, and a metric g_i on M such that (4.54) holds and

$$\mathcal{F}(g_i) \leq \inf_{\mathbb{M}_1} \mathcal{F}_{\varepsilon_i,\lambda_i} + \delta_i, \tag{4.58}$$

where $\mathcal{F} = \mathcal{W}^2$ or $\mathcal{W}^2_\pm$.

Theorem 4.11. *Let M be as in Theorem 4.10 and let $\mathcal{F}$ be one of the functionals $\mathcal{W}^2$, $\mathcal{W}^2_\pm$. Then any preferred sequence (Ω_i, g_i) has a subsequence converging in the pointed Gromov-Hausdorff topology to a minimizing configuration (Ω, g_0) for $\mathcal{F}$ which satisfies all the properties of Theorem 4.10. Within the conformal class $[g_0]$, the metric g_0 minimizes $\mathcal{S}^2$ among compact perturbations. The limit (V, g_0) or (Ω, g_0) satisfies the Bach equations (4.41).*

Proof: Since the sequence (Ω_i, g_i) is preferred, it has a uniform bound on the L^2 norm of curvature R, cf. (4.57)-(4.58) above. The proof that a subsequence converges to a limit (V, g_0) or

(Ω, g_0) satisfying the properties in Theorem 4.10 is then exactly the same as the proof of Theorem 4.10.

The conformal gauge condition satisfied by g_0 is determined by the form of the perturbation $\mathcal{W}^{2p}_{\varepsilon,\lambda}$; different conformal gauges are obtained by choosing different perturbations. To obtain the conformal gauge condition for g_0, divide the trace equation (4.40) by ε. A small computation then gives

$$-\frac{\lambda}{\varepsilon}\Delta s - 2p\Delta\{(1+|Ric|^2)^{p-1}s\} - p\langle E, D^2(1+|Ric|^2)^{p-1}\rangle \\ +2(1+|Ric|^2)^{p-1}[1+(1-p)|Ric|^2+\frac{c_\varepsilon}{\varepsilon}] = 0. \tag{4.59}$$

It then follows from (4.56) and the smooth convergence of g_ε to g_0 on the regular part Ω_0 of Ω, that the metric g_0 satisfies the limit equation

$$\Delta s = 0. \tag{4.60}$$

In fact, we claim that g_0 minimizes the functional $\mathcal{S}^2$ in the conformal class $[g_0]$ among comparison metrics which agree with g_0 outside some compact set. To see this, fix any $\varepsilon > 0$, $\lambda > 0$, (and $p > 1$), and let $g_k = g_k(\varepsilon, \lambda, p)$ be a sequence of metrics on M minimizing $\mathcal{W}^{2p}_{\varepsilon,\lambda}$ and converging to $(\Omega_{\varepsilon,\lambda}, g_{\varepsilon,\lambda})$. Now let $[g_k]$ be the conformal class of $[g_k]$ and let γ_k be metrics in $[g_k]$ minimizing $\mathcal{W}^{2p}_{\varepsilon,\lambda}$ in the conformal class $[g_k]$. (If γ_k is not unique, one may choose a γ_k closest to g_k in a given smooth topology). By a compactness result of Gursky [21], for each k, the metrics γ_k exist on M, and are in $[g_k]$. Thus, the metrics γ_k minimize the functional

$$\frac{2}{\lambda}\mathcal{W}^2 + \frac{\varepsilon}{\lambda}\int(1+|Ric|^2)^p + \int s^2,$$

in $[g_k] \subset \mathbb{M}_1$. Since $\mathcal{W}^2$ is conformally invariant, γ_k minimizes the functional $\frac{\varepsilon}{\lambda}\int(1+|Ric|^2)^p + \int s^2$ in the conformal class $[g_k] \subset \mathbb{M}_1$.

Now choose a sequence $\varepsilon_j \to 0$, $\lambda_j \to 0$ satisfying (4.56), with $g_{\varepsilon_j} \to g_0$. Then a suitable diagonal subsequence of the double sequence (k, j) gives the convergence of γ_{k_j} to g_0, which proves the claim. Of course, (4.60) is the Euler-Lagrange equation for $\mathcal{S}^2$ in a given conformal class. This completes the proof of Theorem 4.11. ∎

Remark 4.12. The following remarks pertain to both Theorems 4.10 and 4.11. It is not known if the metric g_0 is complete on Ω, or if Ω has finitely many components. As in Conjecture 4.7, we conjecture both of these to be the case, and that equality holds in (4.44) as well as in (4.45) when the analogous contributions of the singularities in the collapsing region of $M \setminus K$ are added to the left hand side. Moreover, we conjecture that Ω has finite topological type, and the regular set Ω_0 embeds in M, $\Omega_0 \subset M$.

We point out that it follows from Theorem 4.18 below that all but finitely many components of Ω have an F-structure.

It is clear that if the curvature $|R|$ of g_0 is pointwise bounded, then (Ω, g_0) is complete. In fact, as in Remark 4.3(i), completeness follows from just a uniform lower bound on the Ricci curvature, $Ric_{g_0} \geq -\lambda$, for some $\lambda < \infty$, by the Bishop-Gromov volume comparison theorem.

Again, the functionals $\mathcal{W}^2$ and $\mathcal{W}^2_{\pm}$ are conformally invariant, and so completeness is understood to be with respect to a metric g_0 minimizing $\mathcal{S}^2$ in its conformal class.

Theorems 4.10 and 4.11 give an existence result for idealized minimizers of $\mathcal{F}$ on a given 4-manifold M. As is to be expected, it does not assert the existence of metrics on the given (arbitrary) 4-manifold M which minimize $\mathcal{F}$; instead the objects are "generalized metrics", as described by Cases (I) and (II). Thus, these results define a generalized moduli space $\widetilde{\mathcal{P}}_{\mathcal{F}}$ of minimizers, associated to each functional $\mathcal{F}$ in (3.1) and 4-manifold M.

By their construction in Theorem 4.10, the moduli spaces $\widetilde{\mathcal{P}}_{\mathcal{F}}$ are compact, in that given any sequence $\{g_i\} \in \widetilde{\mathcal{P}}_{\mathcal{F}}$, a subsequence converges in the pointed Gromov-Hausdorff topology to a limit $g \in \widetilde{\mathcal{P}}_{\mathcal{F}}$. The same result holds for the conformally invariant functionals in Theorem 4.11, for the portion $\widetilde{\mathcal{P}}_{\mathcal{F}}(\Lambda)$ of the moduli space $\widetilde{\mathcal{P}}_{\mathcal{F}}$ for which

$$\int_{\Omega} s^2 \leq \Lambda,$$

for any given $\Lambda < \infty$. (As seen in Example 3.1, $\widetilde{\mathcal{P}}_{\mathcal{F}}(\Lambda)$ may be a strict subset of $\widetilde{\mathcal{P}}_{\mathcal{F}}$, for all Λ).

On the other hand, let $\mathcal{M}_{\mathcal{F}}$ be the "actual" moduli space of minimizers of $\mathcal{F}$ on M, i.e. the space of smooth metrics on M, modulo diffeomorphisms, which minimize $\mathcal{F}$. By the elliptic estimates used in the proof of Proposition 4.6, it is straightforward to see that any weak $L^{2,q}$ solution of the Euler-Lagrange equations $\nabla\mathcal{F} = 0$, with $q > 4$, is C^{∞} smooth, (away from the orbifold singularities).

Of course $\mathcal{M}_{\mathcal{F}}$ may well often be empty. At best, the space $\widetilde{\mathcal{P}}_{\mathcal{F}}$ can only be considered empty if $\Omega = \emptyset$, for all minimizers (Ω, g_0) of $\mathcal{F}$, or if M has no preferred minimizing sequences in the case of the conformally invariant functionals. Even in the case $\Omega = \emptyset$, the space $\widetilde{\mathcal{P}}_{\mathcal{F}}$ is better thought of as consisting of fully degenerate metrics on M instead of empty; one expects that the fact that $\Omega = \emptyset$ leads to strong topological restrictions on M.

However, perhaps surprisingly, it is not clear that one always necessarily has

$$\mathcal{M}_{\mathcal{F}} \subset \widetilde{\mathcal{P}}_{\mathcal{F}}. \tag{4.61}$$

Namely, the configurations $(\Omega, g_0) \in \widetilde{\mathcal{P}}_{\mathcal{F}}$ are constructed by a very specific process of first passing to minimizers for a perturbed functional, and then examining their behavior as the perturbation parameter ε is taken to 0. From this construction then, it is not clear that one has obtained the full moduli space of minimizers of $\mathcal{F}$. In particular, if g is a smooth metric on M which minimizes $\mathcal{F}$, then one has to consider the issue whether g may be approximated arbitarily closely by minimizers $(\Omega_\varepsilon, g_\varepsilon)$ of the ε-perturbed functional as $\varepsilon \to 0$. We conjecture that this is always the case, so that in fact (4.61) does hold.

In any case, suppose then that

$$\mathcal{M}_{\mathcal{F}} \neq \emptyset,$$

so that (4.61) is a non-trivial statement, and let g be any smooth metric on M in $\mathcal{M}_{\mathcal{F}}$. Then it is obvious that g can be approximated by smooth metrics on M with a uniform bound on $\mathcal{R}ic^{2p}$. In this context, Theorem 4.1 and Proposition 4.6 hold, and imply that the completion of $\mathcal{M}_{\mathcal{F}}$ with respect to this norm consists of $\mathcal{M}_{\mathcal{F}}$ together with limits which have exactly the same structure as described in Cases (I) and (II) of Theorems 4.10 and 4.11.

The following Lemma is useful and of some interest in this setting.

Lemma 4.13. *Given M and $\mathcal{F} = \mathcal{R}ic^2$ or $\mathcal{R}^2$, suppose that*

$$\mathcal{M}_{\mathcal{F}} \neq \emptyset. \tag{4.62}$$

Then minimizers $(\Omega_\varepsilon, g_\varepsilon)$ of the perturbed functional $\mathcal{F}_\varepsilon$ satisfy

$$\mathcal{R}ic^{2p}(\Omega_\varepsilon, g_\varepsilon) \leq \Lambda, \tag{4.63}$$

for some fixed $\Lambda < \infty$, independent of ε or the choice of minimizer $(\Omega_\varepsilon, g_\varepsilon)$.

Proof: We give the proof in the case of $\mathcal{F} = \mathcal{R}ic^2$; the proof in the case of $\mathcal{R}^2$ is the same. Let $\gamma \in \mathcal{M}_{\mathcal{F}}$ be a smooth metric on M minimizing $\mathcal{F}$ and let $(\Omega_\varepsilon, g_\varepsilon)$ be any minimizing configuration for $\mathcal{F}_\varepsilon = \mathcal{R}ic_\varepsilon^{2p}$. Then $(\Omega_\varepsilon, g_\varepsilon)$ also minimizes the functional

$$I(g) = \frac{1}{\varepsilon}\int |Ric|^2 + \int (1 + |Ric|^2)^p - \inf_{\mathbb{M}_1} \frac{1}{\varepsilon}\int |Ric|^2,$$

which differs from the functional $\frac{1}{\varepsilon}\mathcal{R}ic^{2p}$ by a constant. Let g_j be a sequence of metrics on M with $\mathcal{R}ic^{2p}(g_j)$ uniformly bounded, and converging to $(\Omega_\varepsilon, g_\varepsilon)$ in the pointed Gromov-Hausdorff topology. Since the metric γ realizes $\inf_{\mathbb{M}_1} \frac{1}{\varepsilon}\int |Ric|^2$, one has for any j sufficiently large,

$$I(g_j) \leq I(\gamma) + \delta_j = \int (1 + |Ric|^2(\gamma))^p + \delta_j,$$

where $\delta_j \to 0$ as $j \to \infty$. The result then follows from the obvious inequality $\int (1 + |Ric|^2(g_j))^p \leq I(g_j)$, and taking the limit $j \to \infty$. ■

Thus, the hypothesis (4.62) implies that all elements in $\widetilde{\mathcal{P}}_{\mathcal{F}}$, for $\mathcal{F} = \mathcal{R}ic^2$ or $\mathcal{R}^2$, are also limits of sequences of smooth metrics on M which have a uniform bound on $\mathcal{R}ic^{2p}$. This leads naturally to the following definition.

Definition 4.14. For $\mathcal{F}$ any of the functionals in (3.1), the generalized moduli space $\widetilde{\mathcal{M}}_{\mathcal{F}}$ is defined to be the space of minimizing configurations of $\mathcal{F}$, satisfying the conclusions of Theorems 4.10 or 4.11, endowed with the pointed Gromov-Hausdorff topology.

Thus, by definition, one has

$$\mathcal{M}_{\mathcal{F}} \subset \widetilde{\mathcal{M}}_{\mathcal{F}}. \tag{4.64}$$

Further, Theorem 4.10 implies that for either of the functionals $\mathcal{R}ic^2$ or $\mathcal{R}^2$, the generalized moduli space $\widetilde{\mathcal{M}}_{\mathcal{F}}$ always exists, while for $\mathcal{W}^2$ or $\mathcal{W}^2_{\pm}$, Theorem 4.11 implies the space $\widetilde{\mathcal{M}}_{\mathcal{F}}$ exists provided (4.57) or (4.58) holds for the manifold M.

This discussion leads now quite easily to the following structure theorem on the moduli space $\mathcal{M}_{\mathcal{F}}$. This result generalizes the results of [8], [45] discussed previously in §3(II), cf. also Remark 4.17 below.

Theorem 4.15. *Let $\mathcal{F} = \mathcal{R}ic^2$ or $\mathcal{R}^2$. Then the completion $\overline{\mathcal{M}}_{\mathcal{F}}$ of $\mathcal{M}_{\mathcal{F}}$ with respect to the pointed Gromov-Hausdorff topology is contained in $\widetilde{\mathcal{M}}_{\mathcal{F}}$, so that the metrics are of the form described by (I), (II) in Theorem 4.10. Moreover, the completion $\overline{\mathcal{M}}_{\mathcal{F}}$ is compact.*

The same result holds for $\mathcal{F} = \mathcal{W}^2$ or $\mathcal{W}^2_{\pm}$ for the portion $\mathcal{M}_{\mathcal{F}}(\Lambda)$ given by (4.54), for any given $\Lambda < \infty$.

Proof: For the cases $\mathcal{F} = \mathcal{R}ic^2$ or $\mathcal{R}^2$, Lemma 4.13 gives a uniform bound on $\mathcal{R}ic^{2p}$ on $\widetilde{\mathcal{M}}_{\mathcal{F}}$, and so the result follows from Theorem 4.1.

For the case $\mathcal{F} = \mathcal{W}^2$ or $\mathcal{W}^2_{\pm}$, Lemma 4.13 does not hold. However, one can apply, essentially word for word, the proof of Theorem 4.10 to sequences $g_i \in \mathcal{M}_{\mathcal{F}}(\Lambda)$, using the elliptic regularity estimates from Proposition 4.6, and the result follows from this proof. ■

There are course many questions one can now begin to ask regarding the structure of metrics in $\widetilde{\mathcal{M}}_{\mathcal{F}}$ and the structure of $\widetilde{\mathcal{M}}_{\mathcal{F}}$ as a whole. Obviously, the main question is how the geometric decomposition with respect to $\mathcal{F}$ is related to the topology of M, as in the Thurston decomposition in dimension 3 discussed in §2. We will discuss the case of orbifold and cusp singularities separately.

As in the Einstein case discussed in §3 the geometry and topology of the orbifold singularities is quite well understood. Thus, with each orbifold singularity $q \in V$ or $q \in \Omega$, one has associated a finite number of complete ALE spaces (N_j, γ_j) which arise as blow-up limits at different scales. Each space (N_j, γ_j) is a solution of the Euler-Lagrange equations $\nabla\mathcal{F} = 0$, (in fact a minimizer for $\mathcal{F}$ among compact perturbations). The orbifold singularities at q arise by crushing all of these spaces to points, at different scales, as $\varepsilon \to 0$.

As a simple example, consider the Schwarzschild metric g_S on $S^3 \times \mathbb{R} \sim \mathbb{R}^4 \setminus \{0\}$, given by

$$g_S = (1 + \frac{2m}{r^2})^2 g_{Eucl},$$

where g_{Eucl} is the Euclidean metric on $\mathbb{R}^4 \setminus \{0\}$. This metric is conformally flat, with zero scalar curvature, and so is a critical point (in fact a minimizer) of all the functionals $\mathcal{F}$ in (3.1). The metric g_S has two ALE (in fact AE) ends and the blow-down of g_S is the orbifold consisting of two copies of $\mathbb{R}^4$ glued together at the origin $\{0\}$. The same process can be carried out for instance for conformally flat metrics with any finite number of ALE ends. Similarly, one may have a pair of Schwarzschild metrics joined at a point, resulting from blowing-down a Schwarzschild neck joining them at a higher scale.

This picture holds in general, and the orbifold singularities correspond to a generalized connected sum decomposition of M. In more detail, suppose one is in Case (I) of Theorem 4.10/4.11. Then M is the union of the regular set V_0 with a finite collection of ALE spaces:

$$M = V_0 \cup \{N_k(m)\},$$

where V_0 has a finite number of connected components, and a finite number of singular points q_k, $1 \leq k \leq Q$. The union takes place along spherical space forms S^3/Γ. We recall, as in the example above that the singular points may be topologically regular, i.e. the local fundamental groups attached to q_k may be trivial. For each fixed singular point q_k, one has a finite tree of orbifold ALE spaces, which are smooth manifolds at the top level. Thus, with any q_k itself is associated a finite collection of ALE orbifolds $N_k^{j_1}(1)$, whose blow-downs give rise to the singularity at q_k. Each orbifold singularity $q_{k,\ell}^{j_1}(1)$ in any $N_k^{j_1}(1)$ arises in the same manner, so that a next level collection of ALE orbifolds $N_{k,\ell}^{j_1,j_2}(2)$ is associated with each $q_{k,\ell}^{j_1}(1)$ by blowing down. This process repeats itself a finite number of times, until reaching the top scale, where the ALE spaces are all smooth manifolds; cf. [8], [9] and [11] for further details.

A minimizing sequence g_i for $\mathcal{F}$ converging to (V, g_0) collapses the full collection of ALE spaces $N_k(m)$ to the singular points of V, (at varying scales). Thus, one may write, (somewhat loosely),

$$M = V \#_\Gamma \{N_k(m)\}, \tag{4.65}$$

where the connected sum is along spherical space-forms, (not just along 3-spheres). In most cases, the decomposition (4.65) is topologically non-trivial. However, its an open question whether some component in (4.65) could be a 4-sphere; (especially for minimizers of $\mathcal{F}$, this seems unlikely). The same description as above holds for the orbifold singularities in Ω in the situation of Case (II) of Theorem 4.10/4.11.

On the other hand, essentially nothing is currently known about the relation of the topology of Ω to M when $\Omega \neq V$. In dimension 3, Ω corresponds to the hyperbolic piece H, and so is topologically essential in M^3, in that $\pi_1(H)$ injects in $\pi_1(M^3)$. It would be very interesting to see any progress on this question.

Continuing with the discussion on orbifolds, let $\widetilde{M}_{\mathcal{F}}^o$ be the part of the generalized moduli space $\widetilde{\mathcal{M}}_{\mathcal{F}}$ consisting of the orbifold metrics (V, g_0), i.e. Case (I) of Theorem 4.10/4.11. Suppose

$$\widetilde{\mathcal{M}}_{\mathcal{F}}^o \neq \emptyset.$$

It is natural to ask if $\widetilde{M}_{\mathcal{F}}^o$ arises as the orbifold frontier $\partial_o \mathcal{M}_{\mathcal{F}}$ in the completion $\overline{\mathcal{M}}_{\mathcal{F}}$ of the moduli space $\mathcal{M}_{\mathcal{F}}$ of smooth minimizers (M, g_0) on M. In other words, does there exist a sequence of smooth metrics $g_i \in \mathcal{M}_{\mathcal{F}}$ on M, for which $(M, g_i) \to (V, g_0)$ in the Gromov-Hausdorff (or L^2) topology? Thus, the metrics g_i on M resolve the orbifold singularities on (V, g_0). This of course implies in particular that

$$\mathcal{M}_{\mathcal{F}} \neq \emptyset.$$

This issue is essentially equivalent to the problem of glueing together smooth or orbifold singular metrics in $\widetilde{\mathcal{M}}_{\mathcal{F}}^o$. Namely, one would like to reverse the generalized connected sum decomposition

(4.65) by a glueing process, as follows. Suppose M is a manifold of the form

$$M = V_0 \cup_\Gamma \{N_k\},$$

where the regular set V_0 of V may have several, although finitely many components and $\{N_k\}$ is a collection of ALE manifolds. Suppose V carries an orbifold metric g_0 in $\widetilde{\mathcal{M}_{\mathcal{F}}}$ and each N_k carries a complete smooth metric g_k minimizing $\mathcal{F}$ on compact sets. Can one find a smooth metric g on M, minimizing $\mathcal{F}$, which is close to the metric g_0 on V_0 and to blow-down rescalings of each g_k on N_k?

While this has been a long-standing open question for general Einstein metrics on 4-manifolds, quite a lot is now understood on this question in the case of self-dual metrics. As mentioned in the beginning of §4, Taubes [40] proved the first such geometric glueing theorem in the context of self-dual solutions of the Yang-Mills equations. General glueing techniques and results for self-dual metrics were developed by Floer [20], Donaldson-Friedman [19], culminating in the general result of Taubes [41] that given any compact oriented 4-manifold M, the manifold $\hat{M} = M \# n\mathbb{CP}^2$ for n sufficiently large, (depending on M) admits a smooth self-dual metric. Explicit self-dual metrics on $n\mathbb{CP}^2$ were first constructed by LeBrun [27] and later by Joyce [24]. A very interesting and quite general glueing result for self-dual orbifolds has been obtained by Kovalev-Singer [26]; this answers the question above in the self-dual case affirmatively in many cases.

Remark 4.16. It is of interest to understand whether the results above hold for the much weaker functionals $\mathcal{S}^2$ or $-\mathcal{S}|_{\mathcal{Y}}$. Thus, consider for instance the perturbation

$$\mathcal{S}^{2p}_\varepsilon = \int s^2 + \varepsilon \int (1 + |Ric|^2)^p. \tag{4.66}$$

Theorem 4.1 holds of course for $\mathcal{S}^{2p}_\varepsilon$.

The Euler-Lagrange equations of $\mathcal{S}^{2p}_\varepsilon$ on $\mathbb{M}_1$ are:

$$D^*Dh - 2\delta^*\delta h - \delta\delta h g - 4\delta^*\delta Ric - 4(\delta\delta Ric)g + \mathcal{P}_S = 0, \tag{4.67}$$

where $h = \varepsilon p(1 + |Ric|^2)^{p-1} Ric$ and $\mathcal{P}_S$ is a curvature term, given by

$$\mathcal{P}_S = -2sRic + \tfrac{1}{2}s^2 g - 2R(h) + \tfrac{1}{2}[\varepsilon(1 + |Ric|^2)^p + c]g.$$

Using the same methods as above, it is not difficult to show, although we will not carry out the details, that Proposition 4.6 holds for $\mathcal{S}^{2p}_\varepsilon$, i.e. weak solutions of the Euler-Lagrange equations are C^∞ smooth.

However the proof of Theorem 4.10 does not hold for $\mathcal{S}^2$ (or $-\mathcal{S}|_{\mathcal{Y}}$), and one does not expect this result to hold for such weak functionals. First, a bound on $\mathcal{S}^2$ does not imply a bound on $\mathcal{R}^2$, and so one immediately loses the statements on finiteness of the number of singular points in Theorem 4.10. Moreover, the dominant terms in the Euler-Lagrange equation (4.67), i.e. D^*Dh and $\delta^*\delta h$ tend to 0 weakly as $\varepsilon \to 0$, and so the argument in the proof of Theorem 4.10 concerning smooth convergence to the limit no longer holds. (The equation (4.67) is no longer uniformly elliptic as $\varepsilon \to 0$). Using the methods in [7], it may be possible to obtain such smooth convergence in the regions where $s \leq -\lambda$, for some $\lambda > 0$, but this is likely to fail in regions where $s \geq 0$.

Remark 4.17. We point out that most of the results above do not actually require the limit metrics to be minimizers of $\mathcal{F}$. For example, with the exception of the statements regarding the infimum in (4.43) and (4.45), Theorem 4.10 holds for the components of the moduli space of critical points which are local minimizers of $\mathcal{F}$ for which

$$\mathcal{F} \leq \Lambda, \tag{4.68}$$

for some $\Lambda < \infty$. (Recall that $\mathcal{F}$ is constant on connected components of $\widetilde{\mathcal{M}_{\mathcal{F}}}$). Similarly, Theorem 4.11 holds for the part $\widetilde{\mathcal{M}_{\mathcal{F}}}(\Lambda)$ of $\widetilde{\mathcal{M}_{\mathcal{F}}}$ satisfying (4.68) and (4.54).

For example, all of Theorem 4.15 holds for these parts of the moduli space of critical points of $\mathcal{F}$. We also conjecture that the restriction that the critical metrics locally minimize $\mathcal{F}$ is not necessary.

We conclude with a simple application of the methods discussed above. A well-known question of Gromov asks if there is an $\varepsilon_0 = \varepsilon_0(n) > 0$ such that M is a closed n-manifold admitting a metric such that

$$\mathcal{R}^{n/2} = \int_M |R|^{n/2} \leq \varepsilon_0, \tag{4.69}$$

then is $\inf \mathcal{R}^{n/2} = 0$? In other words, is there a gap for the values of $\inf \mathcal{R}^{n/2}$ about 0. The L^∞ version of this question, i.e.

$$\inf_{\mathbb{M}}(vol^{2/n}|R|_{L^\infty}) \leq \varepsilon_0 \Rightarrow \inf_{\mathbb{M}}(vol^{2/n}|R|_{L^\infty}) = 0, \tag{4.70}$$

was proved in dimension 4 by Rong [37], by showing that the hypothesis in (4.70) implies the existence of a polarized F-structure; the conclusion in (4.70) then follows from the work of Cheeger-Gromov [15].

The following result gives a partial answer to Gromov's question in dimension 4.

Theorem 4.18. *There is an $\varepsilon_0 > 0$, such that if M is a 4-manifold admitting a metric with*

$$\int_M |R|^2 \leq \varepsilon_0, \tag{4.71}$$

then M has an F-structure.

Proof: This is a simple consequence of the proof of Theorem 4.10, with which we thus assume some familiarity. Let (Ω, g_0) be a minimizing configuration for $\mathcal{R}^2$, given by Theorem 4.10. Suppose first $\Omega = \emptyset$. This may happen in two ways. First the minimizing configurations $(\Omega_\varepsilon, g_\varepsilon)$ for $\mathcal{R}_\varepsilon^{2p}$ may be non-empty, but collapse everywhere as $\varepsilon \to 0$. Second, $(\Omega_\varepsilon, g_\varepsilon)$ may be empty for any ε sufficiently small. In either case, it follows from Theorem 4.10 that M has an F-structure metrically on the complement of finitely many singular points. The singular points arise from a concentration of curvature in L^2. However, Lemma 4.2 shows that each singular point contributes a definite amount, c_1, to the integral $\int |R|^2$. Hence, (4.71) implies there are no singular points, and so M itself has an F-structure.

Next suppose $\Omega \neq \emptyset$. Exactly the same argument rules out any orbifold singular points in (Ω, g_0), (or (V, g_0)). To prove then that M has an F-structure, it suffices to prove that Ω has an F-structure, since the complementary region $M \setminus K$ already has an F-structure for K sufficiently large. In turn, the statement that Ω has an F-structure follows from the claim that there exists $\varepsilon_1 = \varepsilon_1(\varepsilon_0)$, such that

$$(\nu^2|R|)(x) \leq \varepsilon_1, \tag{4.72}$$

for all $x \in (\Omega, g_0)$; here ν is the volume radius and $|R|$ is the pointwise norm of the curvature. The estimate (4.72) is scale-invariant and in the scale $\nu = \nu(x) = 1$ requires $|R|(x) \leq \varepsilon_1$. Now the metric g_0 satisfies the Euler-Lagrange equations (4.35) and so by the elliptic estimates in Proposition 4.6 or Corollary 4.8, one has in this scale,

$$\sup_{B_x(1/2)} |R|^2 \leq c \int_{B_x(1)} |R|^2,$$

for a fixed constant $c < \infty$, independent of x and g_0. Since the bound (4.71) is also scale-invariant, this proves (4.72), which proves the result. ■

The proof of Theorem 4.18 above shows in fact that if M has a metric satisfying $0 < \int_M |R|^2 \leq \varepsilon_0$, then M has a (possibly distinct) metric which is ε_1-volume collapsed, i.e. $vol \leq \varepsilon_1$, with locally bounded curvature, i.e. $(inj^2|R|)(x) \leq \varepsilon_1$, where $\varepsilon_1 = \varepsilon_1(\varepsilon_0)$. Thus, the question (4.69) follows if Rong's result can be generalized from a global L^∞ bound on curvature to a local bound.

Via the Chern-Gauss-Bonnet theorem, the condition (4.71) can be reexpressed as

$$\chi(M) + \frac{1}{8\pi^2}\int_M |z|^2 \leq \varepsilon_0,$$

or

$$-4\chi(M) + \frac{1}{\pi^2}\int_M |W|^2 + \frac{1}{24}s^2 \leq \varepsilon_0.$$

Of course, the existence of an F-structure implies that $\chi(M) = 0$.

We point out that the proof of equality in (4.44), (and in (4.45) when the contribution of the singularities in the collapsed part is added), is likely to require a positive solution of the question (4.69); in fact the two questions are probably equivalent.

References

[1] K. Akutagawa and A. Neves, Classification of all 3-manifolds with Yamabe invariant greater than $\mathbb{RP}^3$, math.DG/0502122.

[2] M. Anderson, Ricci curvature bounds and Einstein metrics on compact manifolds, Jour. Amer. Math. Soc. **2**, (1989), 455-490.

[3] M. Anderson, Convergence and rigidity of manifolds under Ricci curvature bounds, Inventiones Math.. **102**, (1990), 429-445.

[4] M. Anderson, The L^2 structure of moduli spaces of Einstein metrics, Geom. Funct. Anal. **2**, (1992), 29-89.

[5] M. Anderson, Extrema of curvature functionals on the space of metrics on 3-manifolds, Calc. Var. and PDE, **5**, (1997), 199-269.

[6] M. Anderson, Scalar curvature and geometrization conjectures for 3-manifolds, in Comparison Geometry, (Ed. K. Grove and P. Petersen), MSRI Publications, vol. 30, (1997), 49-82.

[7] M. Anderson, Scalar curvature and the existence of geometric structures on 3-manifolds, I, II, Jour. Reine Angew. Math. **553**, (2002), 125-182, and **563**, (2003), 115-195.

[8] M. Anderson, Orbifold compactness for spaces of Riemannian metrics and applications, Math. Annalen, **331**, (2005), 739-778.

[9] M. Anderson and J. Cheeger, Diffeomorphism finiteness for manifolds with Ricci curvature and $L^{n/2}$-norm of curvature bounded, Geom. Funct. Anal. **1**, (1991), 231-252.

[10] S. Bando, A. Kasue and H. Nakajima, On a construction of coordinates at infinity on manifolds with fast curvature decay and maximal volume growth, Inventiones Math. **97**, (1989), 313-349.

[11] S. Bando, Bubbling out of Einstein manifolds, Tohoku Math. Jour., **42**, (1990), 205-216 and 587-588.

[12] A. Besse, Einstein Manifolds, Springer Verlag, New York, (1987).

[13] G. Besson, G. Courtois and S. Gallot, Entropies et rigidites des espaces localement symmetriques de courbure strictement negativ, Geom. Funct. Anal., **5**, (1995), 731-799.

[14] H. Bray and A. Neves, Classification of prime 3-manifolds with Yamabe invariant greater than $\mathbb{RP}^3$, Annals of Math. **159**, (2004), 407-424.

[15] J. Cheeger and M. Gromov, Collapsing Riemannian manifolds while keeping their curvature bounded, I, II, Jour. Diff. Geom. **23**, (1986), 309-346 and **32**, (1990), 269-298.

[16] J. Cheeger and G. Tian, Curvature and injectivity radius estimates for 4-Einstein manifolds, (preprint).

[17] S.S. Chern, On the curvatura integra of a Riemannian manifold, Annals of Math. **46**, (1945), 674-684.

[18] D. DeTurck, Local solvability of systems with "extra symbol equivalence", Rend. Sem. Math. Unv. Pol. Torino, Fasc. Spec. (1989), 85-94.

[19] S. Donaldson and R. Friedman, Connect sums of self-dual manifolds and deformations of singular spaces, Nonlinearity, **2**, (1989), 197-239.

[20] A. Floer, Self-dual conformal structures on $l\mathbb{CP}^2$, Jour. Diff. Geom., **33**, (1991), 551-574.

[21] M. Gursky, Compactness of conformal metrics with integral bounds on curvature, Duke Math. Jour. **72**, (1993), 339-367.

[22] M. Gromov, Spaces and questions, GAFA 2000 (Tel Aviv 1999), Geom. Funct. Anal., Special Volume, Part 1, (2000), 118-161.

[23] M. Gross and P.M.H. Wilson, Large complex structure limits of $K3$ surfaces, Jour. Diff. Geom., **55**, (2000), 475-546.

[24] D. Joyce, Explicit construction of self-dual 4-manifolds, Duke Math. Jour., **77**, (1995), 519-552.

[25] B. Kleiner and J. Lott, Perelman website, http://www.math.lsa.umich.edu/research/ricciflow/perelman.html.

[26] A. Kovalev and M. Singer, Gluing theorems for complete anti-self-dual spaces, Geom. Funct. Anal., **11**, (2001), 1229-1281.

[27] C. LeBrun, Explicit self-dual metrics on $\mathbb{CP}_2\#\cdots\#\mathbb{CP}_2$, Jour. Diff. Geom., **34**, (1991), 223-253.

[28] C. LeBrun, Four-dimensional Einstein manifolds, and beyond, in: Surveys in Differential Geometry, vol. VI: Essays on Einstein Manifolds, C. LeBrun and M. Wang, (Eds.), International Press, Boston, (1999), 247-285.

[29] C. LeBrun, Einstein metrics, four-manifolds and differential topology, in Surveys in Differential Geometry, vol. VIII: Papers in honor of Calabi, Lawson, Siu and Uhlenbeck, S.T. Yau (Ed.), International Press, Boston, (2003), 235-255.

[30] C. LeBrun, Curvature functionals, optimal metrics and the differential topology of 4-manifolds, in: Different Faces of Geometry, (S. Donaldson, Y. Eliashberg and M. Gromov, Eds.), Kluwer Academic/Plenum, (2004).

[31] C. Morrey, Jr. Multiple Integrals in the Calculus of Variations, Grundlehren Series 130, Springer Verlag, New York, (1966).

[32] H. Nakajima, Hausdorff convergence of Einstein 4-manifolds, Jour. Fac. Sci. Univ. Tokyo, **35**, (1988), 411-424.

[33] G. Perelman, The entropy formula for the Ricci flow and its geometric applications, math.DG/0211159.

[34] G. Perelman, Ricci flow with surgery on three-manifolds, math.DG/0303109.

[35] G. Perelman, Finite extinction time for the solutions to the Ricci flow on certain three-manifolds, math.DG/0307245.

[36] P. Petersen and G. Wei, Relative volume comparison with integral curvature bounds, Geom. and Funct. Analysis, 7, (1997), 1031-1045.

[37] X. Rong, The existence of polarized F-structures on volume collapsed 4-manifolds, Geom. and Funct. Anal. **3**, (1993), 474-501.

[38] J. Sachs and K. Uhlenbeck, The existence of minimal immersions of 2-spheres, Annals of Math. **113**, (1981), 1-24.

[39] R. Schoen, Variational theory for the total scalar curvature functional for Riemannian metrics and related topics, Springer Lecture Notes in Math, vol. 1265, (1987), 120-154.

[40] C. Taubes, Self-dual Yang-Mills connections on non-self-dual 4-manifolds, Jour. Diff. Geom., **17**, (1982), 139-170.

[41] C. Taubes, The existence of anti-self-dual conformal structures, Jour. Diff. Geom. **36**, (1992), 163-253.

[42] C. Taubes, Metrics, Connections and Gluing Theorems, Conf. Board Math. Sci. **89**, Amer. Math. Soc., (1996).

[43] W. Thurston, Three dimensional manifolds, Kleinian groups and hyperbolic geometry, Bull. Amer. Math. Soc. **6**, (1982), 357-381.

[44] G. Tian, On Calabi's conjecture for complex surfaces with positive first Chern class, Inventiones Math., **101**, (1990), 101-172.

[45] G. Tian and J. Viaclovsky, Moduli spaces of critical Riemannian metrics in dimension four, Adv. in Math., **196**, (2005), 346-372.

[46] D. Yang, Existence and regularity of energy-minimizing Riemannian metrics, Internat. Math. Res. Notices, (1991), 7-13.

[47] S.-T. Yau, On the Ricci curvature of a compact Kähler manifold and the Monge-Ampere equation, Comm. Pure Appl. Math., **31**, (1978), 339-411.

August, 2005

Deptartment of Mathematics
S.U.N.Y. at Stony Brook
Stony Brook, N.Y. 11794-3651
E-mail: anderson@math.sunysb.edu