



Index-Energy Estimates for Yang–Mills Connections and Einstein Metrics

Matthew J. Gursky¹, Casey Lynn Kelleher², Jeffrey Streets³

¹ Department of Mathematics, 255 Hurley Bldg, University of Notre Dame, Notre Dame, IN 46556, USA.

E-mail: mgursky@nd.edu

² Department of Mathematics, Princeton University, Princeton, NJ 08540, USA.

E-mail: ckelleher@princeton.edu

³ Department of Mathematics, Rowland Hall, University of California, Irvine, Irvine 92697, USA.

E-mail: jstreets@uci.edu

Received: 17 January 2019 / Accepted: 3 September 2019

Published online: 4 December 2019 – © Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract: We prove a conformally invariant estimate for the index of Schrödinger operators acting on vector bundles over four-manifolds, related to the classical Cwikel–Lieb–Rozenblum estimate. Applied to Yang–Mills connections we obtain a bound for the index in terms of its energy which is conformally invariant, and captures the sharp growth rate. Furthermore we derive an index estimate for Einstein metrics in terms of the topology and the Einstein–Hilbert energy. Lastly we derive conformally invariant estimates for the Betti numbers of an oriented four-manifold with positive scalar curvature.

1. Introduction

The classical Cwikel–Lieb–Rozenblum (CLR) estimate [Cwi77, Lie76, Ros72], related to the famous asymptotic formula of Weyl [Wey11] on the growth of eigenvalues, bounds the Morse index of a Schrödinger operator $L = -\Delta + V$ on a bounded domain in $\mathbb{R}^n$ in terms of the $L^{\frac{n}{2}}$ norm of the negative part of V . This central result has applications to mathematical physics, where it is referred to as an estimate of the number of bound states for the linear Schrödinger operator. From the point of view of both geometry and mathematical physics, it is important to find similar index/bound state estimates for nonlinear problems, specifically for Yang–Mills connections and Einstein metrics.

Let (X^n, g) be a smooth, compact Riemannian manifold, and suppose ∇ is a connection on a vector bundle E over X . The Yang–Mills energy associated to ∇ is given by

$$\mathcal{YM}[\nabla] := \int_{X^n} |F_\nabla|^2 dV_g.$$

Critical points for $\mathcal{YM}$ are called Yang–Mills connections, including the special class of instantons, which always minimize $\mathcal{YM}$ when they exist. While there are many existence results for instantons (eg. [Tau82]), it is also known that generically one expects non-instanton, non-minimizing Yang–Mills connections to exist even in the critical dimension $n = 4$ [SJU89, HM90, SS92, Bor92]. Furthermore, in dimension 4 every stable Yang–Mills connection with small gauge group is an instanton [BL81], so non-minimizing Yang–Mills connections in this setting will have positive index. Thus, to understand the Yang–Mills functional it becomes important to understand the structure of these non-minimizing Yang–Mills connections, in particular to understand their Morse index. This index is that of the relevant Jacobi operator, a Schrödinger operator acting on Lie algebra-valued 1-forms, with inhomogeneous term determined by the curvature of the underlying Riemannian metric as well as the bundle connection’s curvature. Taking a cue from the CLR estimate one may hope roughly that for a connection to have high Morse index it must also have high Yang–Mills energy. The first main result yields an estimate of this type.

Theorem 1.1. *Let (X^4, g) be a closed, oriented four-manifold, with Yamabe invariant $Y(X^4, [g]) > 0$. Suppose ∇ is a non-instanton Yang–Mills connection on a vector bundle E over X^4 with structure group $G \subset \mathrm{SO}(E)$, and curvature F_∇ . Let $\iota(\nabla)$ denote the index and $\nu(\nabla)$ the nullity of ∇ . Then*

$$\begin{aligned} \iota(\nabla) + \nu(\nabla) \leq & \frac{144e^2 \dim(\mathfrak{g}_E)}{Y(X^4, [g])^2} \left\{ -12\pi^2 \chi(X^4) + 12 \int_X |F_\nabla|^2 dV_g \right. \\ & \left. + 3\sqrt{2} \int_X |W_g| |F_\nabla| dV_g + 3 \int_X |W_g|^2 dV_g \right\}, \end{aligned}$$

where $\chi(X^4)$ is the Euler characteristic of X^4 and W_g is the Weyl tensor.

If ∇ is an instanton, then $\nu(\nabla) = 0$ and the Atiyah–Singer index formula gives an explicit formula for $\iota(\nabla)$ depending on topological data (see Chapter 4 of [DK90]). Our statement explicitly does not include this case, and we use the assumption of nonvanishing of F_∇^+ when constructing a metric conformal to the base, with respect to which we carry out the index estimate (see Proposition 3.7). When the base manifold is the round sphere we can simplify the statement to the following:

Corollary 1.2. *Let $E \rightarrow (\mathbb{S}^4, g_{\mathbb{S}^4})$ be a vector bundle over the round sphere with structure group $G \subset \mathrm{SO}(E)$, with ∇ a non-instanton Yang–Mills connection. Then*

$$\iota(\nabla) + \nu(\nabla) \leq 9e^2 \dim(\mathfrak{g}_E) \left\{ -1 + \frac{1}{4\pi^2} \int_{\mathbb{S}^4} |F_\nabla|^2 dV_g \right\}.$$

An index plus nullity estimate for Yang–Mills connections appeared in [Ura86], under the much stronger assumption that the base manifold has positive Ricci curvature and with a bound depending on the L^∞ -norm of the bundle curvature. Our result only assumes positive Yamabe invariant, and the bound depends on conformal invariants of the base manifold and the Yang–Mills energy. This is more natural, in view of the fact that the index and nullity are conformal invariants. Furthermore, although the constants in Theorem 1.1 are almost certainly not sharp (in fact, the sharp value is not known in the classical CLR inequality; cf. [HKRV18]), we can show by means of examples that the growth rate of the index as a function of the Yang–Mills energy is sharp. Specifically, combining an index estimate of Taubes [Tau83] as well as an explicit construction of

non-instanton Yang–Mills connections due to Sadun–Segert [SS92], we exhibit a family of connections whose index grows linearly in the Yang–Mills energy (Proposition 3.10 below). Lastly we point out that the estimate we give in Sect. 2 can be adapted to give an index estimate for Yang–Mills connections in any dimension in terms of the $L^{\frac{n}{2}}$ norms of F and the Ricci curvature, and the Sobolev constant, and in this case the proof is a very direct adaptation of the method of Li–Yau [LY83] (see Remark 2.6).

Our second main result is an index estimate for Einstein metrics in dimension four. Einstein metrics arise as critical points of the normalized total scalar curvature functional

$$\mathcal{S}[g] = \text{Vol}(g)^{-1/2} \int_{X^4} R_g \, dV_g. \quad (1.1)$$

It is well-known that Einstein metrics are never stable critical points, since $\mathcal{S}$ is minimized over conformal variations but is locally maximized over transverse-traceless variations, possibly up to a finite dimensional subspace. The index $\iota(g)$ of an Einstein metric, which we define to be the Morse index of $-\mathcal{S}$, is the dimension of the maximal subspace on which the second variation is negative when restricted to transverse traceless variations, while the nullity $\nu(g)$ is the dimension of the space of infinitesimal Einstein deformations. While there are some works characterizing the stability and space of deformations of Einstein metrics ([Koi79, Koi82, DWW05, DWW07]), it seems very little is known about the index in the case it is positive. Intuitively, one might expect an Einstein metric with large index to have small energy. We derive an estimate of this kind which relies on explicit universal constants and the Euler characteristic.

Theorem 1.3. *Let (X^4, g) be an Einstein four-manifold with positive scalar curvature. Then*

$$\mathcal{S}[g] \leq 24\pi \sqrt{\frac{\chi(X^4)}{3 + \delta[\iota(g) + \nu(g)]}},$$

where $\delta = \frac{1}{24}e^{-2}$.

Our final application is a bound on the Betti numbers of an oriented four-manifold X^4 of positive scalar curvature. Bounds for the Betti numbers in terms of the curvature, Sobolev constant, and diameter of the manifold were proved by P. Li in [Li80]. These estimates can be viewed as refined or quantitative versions of the classical vanishing theorems; see [B88] for a beautiful survey. To state our results we need to introduce two conformal invariants of four-manifolds with positive Yamabe invariant.

To define the first conformal invariant, we need some additional notation. Let $A = A_g$ denote the Schouten tensor of g :

$$A = \frac{1}{2} \left(\text{Ric} - \frac{1}{6} Rg \right),$$

where Ric is the Ricci tensor and R the scalar curvature of g . Let $\sigma_2(A)$ denote the second symmetric function of the eigenvalues of A (viewed as a symmetric bilinear form on the tangent space at each point). Then

$$\sigma_2(A) = -\frac{1}{8} |\text{Ric}|^2 + \frac{1}{24} R^2.$$

The integral of this expression is a scalar conformal invariant of a four-manifold. Using this we define the following two conformal invariants:

$$\begin{aligned}\rho_1(X^4, [g]) &:= \frac{4 \int_X \sigma_2(A) \, dV}{Y(X^4, [g])^2}, \\ \rho_+(X^4, [g]) &:= \frac{24 \int_X |W^+|^2 \, dV}{Y(X^4, [g])^2}.\end{aligned}\tag{1.2}$$

Let $b_1(X^4)$ denote the first Betti number of X^4 , and let $b^+(X^4)$ denote the maximal dimension of a subspace of $\Lambda^2(X^4)$ on which the intersection form is positive. It follows from ([Gur98] Theorem 2) that if $b_1(X^4) > 0$ then $\rho_1 \leq 0$, with equality only when conformal to a quotient of $S^3 \times \mathbb{R}$ with the product metric. Furthermore, it follows from ([Gur00] Theorem 3.3) that if $b^+ > 0$ then $\rho_+ \geq 1$, with equality only when conformal to a Kähler metric with positive scalar curvature. Using the general index estimate of Section 2, we can prove quantitative versions of these estimates:

Theorem 1.4. *Let (X^4, g) be an oriented four-manifold with $Y(X^4, [g]) > 0$. Then*

$$b_1(X^4) \leq 9e^2 (1 - 24\rho_1), \tag{1.3}$$

and

$$b^+(X^4) \leq 3e^2 (2\sqrt{\rho_+} - 1)^2. \tag{1.4}$$

Here, as in the Yang–Mills estimate, our constants are not sharp but the growth rate likely is. In particular, by taking connect sums with sufficiently long necks, we can produce locally conformally flat metrics on the manifold $k\#\mathbb{S}^3 \times \mathbb{S}^1$ whose Yamabe invariant is uniformly bounded below. Evidently this manifold has $b_1 = k$, while for these conformal classes we see that the right hand side of (1.3) grows linearly in k . To verify that the growth rate of b_+ is sharp the natural candidates to consider are the self-dual metrics on $k\#\mathbb{CP}^2$ constructed by LeBrun [LeB91]. However, we do not know if the Yamabe invariant of these metrics has a uniform lower bound.

The proofs of these theorems all rely on an extension of the CLR estimate to elliptic operators on vector bundles with certain geometric backgrounds (see Sect. 2). The case of dimension $n = 4$ especially requires careful analysis of the curvature terms in the relevant index operator in order to capture the conformal invariance. While many proofs of the classical CLR inequality by now exist, the proof of Li–Yau [LY83] gives explicit bounds in terms of the Sobolev constant. By adapting their ideas to operators modeled on the conformal Laplacian but acting on sections of a vector bundle, we are able to obtain estimates in terms of conformal invariants. An important technical step is to compare the L^2 -trace of the heat kernel of a Schrödinger-type operator acting on sections of a vector bundle to the heat trace of an associated scalar operator. Again, many results of this kind exist (see [HSU77, HSU80, Sim79]), but we adapt a proof of Donnely–Li [DL82] as it is closest in spirit to the other estimates. Combining these ideas together with a conformal gauge-fixing argument yields our main index estimates.

2. General Index Estimate

In this section we adapt the proof of the Cwikel–Lieb–Rosenblum inequality due to Li–Yau [LY83] to prove an index estimate for a certain class of elliptic operators acting on

sections of vector bundles. Given a vector bundle $\mathcal{E} \rightarrow (X^4, g)$ with a metric-compatible connection ∇ , let $\Delta = \Delta_g : \Gamma(\mathcal{E}) \rightarrow \Gamma(\mathcal{E})$ denote the rough Laplacian, where in local coordinates $\Delta = g^{ij} \nabla_i \nabla_j$ (note this convention differs from some references). Given a non-negative function $V \in C^0(X^4)$, consider the operator

$$S = -\Delta + \frac{1}{6}R - V, \quad (2.1)$$

where $R = R_g$ is the scalar curvature of g . We will assume throughout this section that $R \geq 0$, and the Yamabe invariant $Y(X^4, [g]) > 0$. Our main result is

Theorem 2.1. *If $N_0(S)$ denotes the number of non-positive eigenvalues of S , then*

$$N_0(S) \leq \frac{36e^2 \operatorname{rank}(\mathcal{E}) \|V\|_{L^2}^2}{Y(X^4, [g])^2}. \quad (2.2)$$

The proof is a consequence of a series of technical lemmas, and will appear at the end of the section. We begin with some notation. We need to distinguish between the Laplacian on functions and the rough Laplacian acting on sections of $\mathcal{E}$, so from now on we set

$$\begin{aligned} \Delta_0 : C^\infty(X^4) &\rightarrow C^\infty(X^4), \\ \Delta : C^\infty(\mathcal{E}) &\rightarrow C^\infty(\mathcal{E}). \end{aligned}$$

Fix some small $\epsilon > 0$ define

$$V_\epsilon := V + \epsilon. \quad (2.3)$$

Consider the two operators

$$\begin{aligned} \mathcal{P}_0 &:= \frac{1}{V_\epsilon} \left(\Delta_0 - \frac{1}{6}R \right), \\ \mathcal{P} &:= \frac{1}{V_\epsilon} \left(\Delta - \frac{1}{6}R \right). \end{aligned}$$

As a first step we give the following analogue of an estimate in Li–Yau:

Lemma 2.2. *Let $\mu_1^0 \leq \mu_2^0 \leq \dots$ denote the eigenvalues of $-\mathcal{P}_0$, counted with multiplicity. Then for all $t > 0$,*

$$\sum_{i=1}^{\infty} e^{-2\mu_i^0 t} \leq \frac{36 \|V_\epsilon\|_{L^2}^2}{Y(X^4, [g])^2} t^{-2}. \quad (2.4)$$

Proof. As in [LY83], we take $\{\psi_i\}$ to be an orthonormal basis of $L^2(V_\epsilon dV)$ consisting of eigenfunctions of $-\mathcal{P}_0$:

$$-\mathcal{P}_0 \psi_i = \mu_i \psi_i,$$

with

$$\int_X \psi_i(x) \psi_j(x) V_\epsilon(x) dV_x \equiv \delta_{ij}.$$

Let

$$H_0(x, y, t) := \sum_{i=1} e^{-t\mu_i} \psi_i(x) \psi_i(y).$$

Note that H_0 is the heat kernel associated to the operator $\mathcal{P}_0$ with respect to the weighted inner product $L^2(V_\epsilon dV)$. In particular,

$$\begin{aligned} \frac{\partial}{\partial t} [H_0(x, y, t)] &= \mathcal{P}_0 H_0(x, y, t) \\ &= \frac{1}{V_\epsilon} \left(\Delta_0 - \frac{1}{6} R \right) H_0(x, y, t). \end{aligned} \quad (2.5)$$

Moreover, since $R \geq 0$ we have

$$H_0(x, y, t) > 0,$$

and for any $f \in C^0(X^4)$,

$$\lim_{t \rightarrow 0} \int_X H_0(x, y, t) f(y) V_\epsilon(y) dV_y = f(x). \quad (2.6)$$

We also let

$$\begin{aligned} h(t) &:= \int_X \int_X H_0(x, y, t)^2 V_\epsilon(x) V_\epsilon(y) dV_x dV_y \\ &= \sum_{i=1}^{\infty} e^{-2\mu_i^0 t}. \end{aligned}$$

We now argue as in the proof of Theorem 2 of [LY83]: differentiating h , using (2.5) and integrating by parts, we have

$$\begin{aligned} \frac{dh}{dt} &= 2 \int_X V_\epsilon(x) \int_X H_0(x, y, t) (\mathcal{P}_0)_y H_0(x, y, t) V_\epsilon(y) dV_y dV_x \\ &= 2 \int_X V_\epsilon(x) \int_X H_0(x, y, t) \left(\Delta_0 - \frac{1}{6} R \right)_y H_0(x, y, t) dV_y dV_x \\ &= -2 \int_X V_\epsilon(x) \int_X \left[|\nabla_y H_0(x, y, t)|^2 + \frac{1}{6} R(y) H_0(x, y, t)^2 \right] dV_y dV_x, \\ &= -2 \int_X V_\epsilon(x) \int_X \left[|\nabla_y H_0(x, y, t)|^2 + \frac{1}{6} R(y) H_0(x, y, t)^2 \right] dV_y dV_x. \end{aligned} \quad (2.7)$$

By the definition of the Yamabe invariant,

$$\begin{aligned} Y(X^4, [g]) &\left(\int_X H_0(x, y, t)^4 dV_y \right)^{1/2} \\ &\leq 6 \int_X \left[|\nabla_y H_0(x, y, t)|^2 + \frac{1}{6} R_y H_0(x, y, t)^2 \right] dV_y. \end{aligned}$$

Using this, we can rewrite (2.7) as

$$\frac{dh}{dt} \leq -\frac{1}{3} Y(X^4, [g]) \int_X V_\epsilon(x) \left(\int_X H_0(x, y, t)^4 dV_y \right)^{1/2} dV_x. \quad (2.8)$$

To obtain a differential inequality for h we need a further a priori upper bound. Iterating Hölder's inequality twice and using the fact that $H_0(x, y, t) > 0$ we note

$$\begin{aligned}
 h(t) &= \int_X V_\epsilon(x) \int_X H_0(x, y, t)^2 V_\epsilon(y) dV_y dV_x \\
 &\leq \int_X V_\epsilon(x) \left[\left(\int_X H_0(x, y, t)^4 dV_y \right)^{1/3} \left(\int_X H_0(x, y, t) V_\epsilon^{3/2}(y) dV_y \right)^{2/3} \right] dV_x \\
 &\leq \left[\int_X V_\epsilon(x) \left(\int_X H_0(x, y, t)^4 dV_y \right)^{1/2} dV_x \right]^{2/3} \\
 &\quad \left[\int_X V_\epsilon(x) \left(\int_X H_0(x, y, t) V_\epsilon^{3/2}(y) dV_y \right)^2 dV_x \right]^{1/3}. \tag{2.9}
 \end{aligned}$$

It remains to estimate the second term on the right hand side above, which is done by treating it as an auxiliary solution to the heat equation. In particular set

$$Q(x, t) := \int_X H_0(x, y, t) V_\epsilon(y)^{3/2} dV_y. \tag{2.10}$$

Note that Q is a solution of the heat equation associated to $\mathcal{P}_0$:

$$\begin{aligned}
 \frac{\partial}{\partial t} [Q(x, t)] &= (\mathcal{P}_0 Q)(x, t) = \frac{1}{V_\epsilon(x)} (\Delta_0 - \frac{1}{6} R) Q(x, t), \\
 Q(x, 0) &= V_\epsilon^{1/2}(x). \tag{2.11}
 \end{aligned}$$

Note in particular the power of V_ϵ , which is a consequence of the weighted inner product. We first compute

$$\begin{aligned}
 \frac{d}{dt} \left[\int_X Q(x, t)^2 V_\epsilon(x) dV_x \right] &= 2 \int_X Q(x, t) \frac{\partial}{\partial t} [Q(x, t)] V_\epsilon(x) dV \\
 &= 2 \int_X Q(x, t) (\Delta_0 - \frac{1}{6} R) Q(x, t) dV \\
 &= -2 \int_X \left[|\nabla Q(x, t)|^2 + \frac{1}{6} R(x) Q(x, t)^2 \right] dV \\
 &\leq 0. \tag{2.12}
 \end{aligned}$$

Integrating this and applying (2.11),

$$\begin{aligned}
 \int_X Q(x, t)^2 V_\epsilon(x) dV &\leq \int_X Q(x, 0)^2 V_\epsilon(x) dV \\
 &= \int_X V_\epsilon(x)^2 dV.
 \end{aligned}$$

Now, using (2.10),

$$\int_X Q(x, t)^2 V_\epsilon(x) dV = \int_X V_\epsilon(x) \left(\int_X H_0(x, y, t) V_\epsilon(y)^{3/2} dV_y \right)^2 dV_x,$$

and so substituting into (2.12) we obtain

$$\|V_\epsilon\|_{L^2} \geq \left[\int_X V_\epsilon(x) \left(\int_X H_0(x, y, t) V_\epsilon(y)^{3/2} dV_y \right)^2 dV_x \right]^{1/2}.$$

Substituting this into (2.9), we have

$$h(t) \leq \left[\int_X V_\epsilon(x) \left(\int_X H_0(x, y, t)^4 dV_y \right)^{1/2} dV_x \right]^{2/3} \|V_\epsilon\|_{L^2}^{2/3}.$$

By (2.8), we conclude

$$\frac{dh}{dt} \leq -\frac{1}{3} \frac{Y(X^4, [g])}{\|V_\epsilon\|_{L^2}} h(t)^{3/2}. \quad (2.13)$$

Integrating and using the fact that $h(t) \rightarrow \infty$ as $t \rightarrow 0^+$ we conclude

$$h(t) \leq \frac{36 \|V_\epsilon\|_{L^2}^2}{Y(X^4, [g])^2} t^{-2},$$

which is equivalent to (2.4). $\square$

The key lemma that allows us to pass from Lemma 2.2 to Theorem 2.1 is the following:

Lemma 2.3. *We have*

$$\mathrm{tr}_{L^2} e^{t\mathcal{P}} \leq \mathrm{rank}(\mathcal{E}) \mathrm{tr}_{L^2} e^{t\mathcal{P}_0}. \quad (2.14)$$

Proof. This is based on argument in [DL82], Theorem 4.3 and Corollary 4.4. Let $H(x, y, t)$ denote the heat kernel associated to $\mathcal{P}$ with respect to the weighted inner product of Lemma 2.2. More precisely, let $\mu_1 \leq \mu_2 \leq \dots$ denote the eigenvalues of $-\mathcal{P}$, counted with multiplicity, and let $\{\phi_i\}$ be an orthonormal basis of sections of $L^2(\mathcal{E}, V_\epsilon dV)$ consisting of eigenfunctions of $-\mathcal{P}$:

$$-\mathcal{P}\phi_i = \mu_i\phi_i,$$

with

$$\int_X \langle \phi_i(x), \phi_j(x) \rangle V_\epsilon(x) dV_x = \delta_{ij}.$$

Then the associated heat kernel is given by

$$H(x, y, t) = \sum_{i=1} e^{-t\mu_i} \phi_i(x) \otimes \phi_i(y).$$

If $|H|$ denotes the norm of H as an endomorphism $H(\cdot, x, y) : \mathcal{E}_x \rightarrow \mathcal{E}_y$, then $|H|$ is a subsolution of (2.5) (in the sense of distributions):

$$\begin{aligned} \frac{\partial}{\partial t} [|H|(x, y, t)] &\leq \mathcal{P}_0 |H|(x, y, t) \\ &= \frac{1}{V_\epsilon} (\Delta_0 |H| - \frac{1}{6} R) |H|(x, y, t), \end{aligned} \quad (2.15)$$

see Lemma 4.1 of [DL82]. Also, in analogy with (2.6), for any $f \in C^0(X^4)$ we have

$$\lim_{t \rightarrow 0} \int_X |H(x, y, t)| f(y) V_\epsilon(y) dV_y = f(x). \quad (2.16)$$

By (2.6) and (2.16),

$$\begin{aligned} & |H|(x, y, t) - H_0(x, y, t) \\ &= \lim_{\tau \rightarrow 0} \left\{ \int_X |H|(x, z, t) H_0(z, y, \tau) V_\epsilon(z) dV_z \right. \\ &\quad \left. - \int_X |H|(x, z, \tau) H_0(z, y, t) V_\epsilon(z) dV_z \right\} \\ &= \int_0^t \frac{d}{ds} \left[\int_X |H|(x, z, s) H_0(z, y, t-s) V_\epsilon(z) dV_z \right] ds \\ &= \left[\int_0^t \left[\int_X \frac{d}{ds} [|H|(x, z, s)] H_0(z, y, t-s) V_\epsilon(z) dV_z \right] ds \right]_{T_1} \\ &\quad + \left[\int_0^t \left[\int_X |H|(x, z, s) \frac{d}{ds} [H_0(z, y, t-s)] V_\epsilon(z) dV_z \right] ds \right]_{T_2}. \end{aligned} \quad (2.17)$$

We manipulate the second term T_2 using (2.5),

$$\begin{aligned} T_2 &= \int_0^t \int_X |H|(x, z, s) \frac{\partial}{\partial s} H_0(z, y, t-s) V_\epsilon(z) dV_z ds \\ &= - \int_0^t \int_X |H|(x, z, s) \left(\frac{1}{V_\epsilon(z)} (\Delta_0 - \frac{1}{6} R) H_0(z, y, t-s) \right) V_\epsilon(z) dV_z ds \\ &= - \int_0^t \int_X |H|(x, z, s) \Delta_0 H_0(z, y, t-s) dV_z ds \\ &\quad + \int_0^t \int_X |H|(x, z, s) \frac{1}{6} R(z) H_0(z, y, t-s) dV_z ds. \end{aligned} \quad (2.18)$$

Integrating by parts in the term involving Δ_0 and using (2.15), reincorporating T_2 into (2.17),

$$\begin{aligned} & |H|(x, y, t) - H_0(x, y, t) \\ &= \int_0^t \int_X \left(\frac{\partial}{\partial s} - \mathcal{P}_0 \right) |H|(x, z, s) H_0(z, y, t-s) q_\epsilon(z) dV_z \leq 0. \end{aligned} \quad (2.19)$$

Therefore, if $\text{tr}_g H$ denotes the pointwise trace of $H(\cdot, x, x) : \mathcal{E}_x \rightarrow \mathcal{E}_x$,

$$\begin{aligned} \text{tr}_{L^2} e^{tP} &= \int_X \text{tr}_g H(x, x, t) V_\epsilon(x) dV_x \\ &\leq \text{rank}(\mathcal{E}) \int_X |H|(x, x, t) V_\epsilon(x) dV_x \\ &\leq \text{rank}(\mathcal{E}) \int_X H_0(x, x, t) V_\epsilon(x) dV_x \\ &= \text{rank}(\mathcal{E}) \text{tr}_{L^2} e^{tP_0}. \end{aligned}$$

The result follows. $\square$

Combining Lemma 2.2 with Lemma 2.3 we have

Proposition 2.4. *Let $\mu_1 \leq \mu_2 \leq \dots$ denote the eigenvalues of $-\mathcal{P}$, counted with multiplicity. Then for all $t > 0$,*

$$\sum_{i=1}^{\infty} e^{-2\mu_i t} \leq \frac{36 \operatorname{rank}(\mathcal{E}) \|V_\epsilon\|_{L^2}^2}{Y(X^4, [g])^2} t^{-2}. \quad (2.20)$$

Proof. Observe that

$$\sum_{i=1}^{\infty} e^{-2\mu_i t} = \left(\operatorname{tr}_{L^2} e^{(2t)\mathcal{P}} \right). \quad (2.21)$$

But by Lemma 2.3,

$$\left(\operatorname{tr}_{L^2} e^{(2t)\mathcal{P}} \right) \leq \operatorname{rank}(\mathcal{E}) \left(\operatorname{tr}_{L^2} e^{(2t)\mathcal{P}_0} \right) = \operatorname{rank}(\mathcal{E}) \sum_{i=1}^{\infty} e^{-2\mu_i^0 t}. \quad (2.22)$$

Thus the result follows from Lemma 2.2.

Corollary 2.5. *Let μ_k denote the k^{th} -eigenvalue of $-\mathcal{P}$. Then*

$$\frac{36e^2 \operatorname{rank}(\mathcal{E}) \|V_\epsilon\|_{L^2}^2}{Y(X^4, [g])^2} \mu_k^2 \geq k. \quad (2.23)$$

Proof. As in [LY83], take $t = \frac{1}{\mu_k}$ in (2.20), then

$$\begin{aligned} \frac{36 \operatorname{rank}(\mathcal{E}) \|V_\epsilon\|_{L^2}^2}{Y(X^4, [g])^2} \mu_k^2 &\geq \sum_{i=1}^{\infty} \exp\left(-2 \frac{\mu_i}{\mu_k}\right) \\ &\geq \sum_{i=1}^k \exp\left(-2 \frac{\mu_i}{\mu_k}\right) \\ &\geq k e^{-2}. \end{aligned}$$

The result follows. $\square$

Proof of Theorem 2.1. By the argument of Birman–Schwinger, the number of non-positive eigenvalues of the operator $-\Delta + \frac{1}{6}R + V_\epsilon$ is less than or equal to the number of eigenvalues of the operator $-\mathcal{P} = \frac{1}{V_\epsilon}(-\Delta + \frac{1}{6}R)$ that are less than or equal to 1 (for an overview of the argument, see (iv) in the proof of [LY83] Corollary 2). But, by (2.23), if μ_k the greatest eigenvalue of $-\mathcal{P}$ that is less than or equal to 1, then

$$k \leq \frac{36e^2 \operatorname{rank}(\mathcal{E}) \|V_\epsilon\|_{L^2}^2}{Y(X^4, [g])^2}.$$

Therefore, taking $\epsilon \rightarrow 0$ we conclude

$$N_0(\mathcal{S}) \leq \frac{36e^2 \operatorname{rank}(\mathcal{E}) \|V\|_{L^2}^2}{Y(X^4, [g])^2},$$

which completes the proof. $\square$

Remark 2.6. If $\mathcal{E} \rightarrow (X^n, g)$ is a vector bundle, $n \geq 3$, and $\mathcal{S} = -\Delta + V$ is a linear operator acting on sections of E with $V \geq 0$, then the preceding arguments can easily be adapted to give an estimate for the number of non-positive eigenvalues of $\mathcal{S}$. If $C_S(g)$ denotes the Sobolev constant,

$$C_S(g) \left(\int_X |f|^{\frac{2n}{n-2}} dV \right)^{\frac{n-2}{n}} \leq \int_X [|\nabla f|^2 + f^2] dV,$$

then

$$N_0(\mathcal{S}) \leq c_n \frac{\text{rank}(\mathcal{E})}{C_S(g)^{\frac{n}{2}}} \|(1 + V)\|_{L^{n/2}}^{n/2}.$$

3. Index Estimate for Yang–Mills Connections

3.1. Background. Let $(E, h) \rightarrow (X^n, g)$ be a vector bundle with metric over a closed Riemannian manifold with structure group $G \subset \text{SO}(E)$. Let $\Gamma(E)$ denote the smooth sections of E , and $\mathfrak{g}_E$ denote the associated Lie algebra of E . For each point $x \in X^n$ choose a local orthonormal basis of TX^n given by $\{e_i\}$ with dual basis $\{e^i\}$ and a local basis for E given by $\{\mu_\alpha\}$ with dual basis $\{(\mu^*)^\alpha\}$ of the dual E^* . Let Λ^p denote the space of smooth p -forms over X and set $\Lambda^p(E) := \Lambda^p \otimes \Gamma(E)$. Given an element in $\Lambda^p(E)$ its components are understood be with respect to the forgoing bases. We will also use the fact that when $p = 1$, we can take tensor products of the basis elements $\{e^i\}, \{\mu_\alpha\}, \{(\mu^*)^\alpha\}$ to obtain a (local) basis of $\Lambda^1(E)$.

We will use the following conventions for the various inner products that appear:

$$\begin{aligned} \langle \eta, \omega \rangle_{\Lambda^2} &= \frac{1}{2} \sum_{i,j} \eta_{ij} \omega_{ij}, & \langle \nu, \mu \rangle_{S_0^2(X)} &= \sum_{i,j} \nu_{ij} \mu_{ij}, \\ \langle A, B \rangle_{\mathfrak{g}_E} &= -\frac{1}{2} \text{tr}_E(AB) = -\frac{1}{2} \sum_{\alpha,\beta} A_\alpha^\beta B_\beta^\alpha \\ \langle P, Q \rangle_{\Lambda^1(\mathfrak{g}_E)} &= -\frac{1}{2} P_{i\beta}^\alpha Q_{i\alpha}^\beta, & \langle R, S \rangle_{\Lambda^2(\mathfrak{g}_E)} &= -\frac{1}{4} R_{ij\beta}^\alpha S_{i\alpha}^\beta. \end{aligned}$$

Here, repeated Latin indices indicate contractions by the metric g on X^n , and the components are with respect to the orthonormal basis above. Unless specified otherwise, we will use Einstein summation notation for both bundle and base components.

We need certain algebraic actions as well. First there is the bracket operation $[\cdot, \cdot] : \Lambda^1(\mathfrak{g}_E) \times \Lambda^1(\mathfrak{g}_E) \rightarrow \Lambda^2(\mathfrak{g}_E)$ defined by

$$[A, B]_{jk\alpha}^\beta := A_{j\delta}^\beta B_{k\alpha}^\delta - B_{k\delta}^\beta A_{j\alpha}^\delta, \quad A, B \in \Lambda^1(\mathfrak{g}_E).$$

Also, given $\eta \in S^2(TX)$ and $\Phi \in \Lambda^2(\mathfrak{g}_E)$, we may view both as elements of $\text{End}(\Lambda^1(\mathfrak{g}_E))$ via the formulas

$$\begin{aligned} (\eta(A))_{i\alpha}^\beta &= \eta_{ij} A_{j\alpha}^\beta, \\ ([\Phi, A])_{i\alpha}^\beta &= [\Phi_{ji}, A_j]_\alpha^\beta = \Phi_{ji\mu}^\beta A_{j\alpha}^\mu - A_{j\mu}^\beta \Phi_{ji\alpha}^\mu. \end{aligned}$$

We next recall the definition of the Jacobi operator of $\mathcal{YM}$ (see Theorem (6.8) [BL81]).

Theorem 3.1. *Suppose ∇ is a Yang–Mills connection on a vector bundle E over X^n with structure group $G \subset \mathrm{SO}(E)$, and $\{\nabla_s\}$ is a one parameter family of connections with $\nabla \equiv \nabla_s|_{s=0}$. Furthermore, suppose $B := \frac{\partial}{\partial s} [\nabla_s]|_{s=0} \in \Lambda^1(\mathfrak{g}_E)$. Then*

$$\frac{d^2}{ds^2} [\mathcal{YM}(\nabla_s)]|_{s=0} = 2 \int_X \left\langle \mathcal{J}^\nabla(B), B \right\rangle_{\Lambda^1(\mathfrak{g}_E)} dV,$$

where

$$\mathcal{J}^\nabla(B)_i = -\Delta B_i - \nabla_i \nabla^j B_j + 2 \left[F_{ji}, B^j \right] + \mathrm{Ric}_i^j B_j,$$

where $\Delta = \nabla^a \nabla_a$ denotes the rough Laplacian.

The operator $\mathcal{J}^\nabla$ is degenerate elliptic, due to the action of the infinite dimensional gauge group. Questions of index and nullity always refer to the operator restricted to divergence-free sections B , one which the operator takes the simpler form:

$$\mathcal{J}^\nabla(B)_i = -\Delta B_i + 2 \left[F_{ji}, B^j \right] + \mathrm{Ric}_i^j B_j. \quad (3.1)$$

The index and nullity of a Yang–Mills connection are understood to be those quantities associated to this operator. It follows from the conformal invariance of the Yang–Mills energy that both the index and nullity are conformally invariant.

3.2. Linear algebraic estimates. In this subsection we obtain linear algebraic estimates which enter into estimating the Jacobi operator. The key point is Proposition 3.5, which provides a sharp inequality between the operator and Hilbert–Schmidt norms of the bilinear form appearing in the Jacobi operator. Let $Z \in S_0^2(TX)$ and $\Phi \in \Lambda^2(\mathfrak{g}_E)$; in the following we can view both as elements of $\mathrm{End}(\Lambda^1(\mathfrak{g}_E))$.

Lemma 3.2. *Suppose $E \rightarrow (X^n, g)$ is a vector bundle. Then Z and Φ , viewed as endomorphisms of $\Lambda^1(\mathfrak{g}_E)$, are symmetric. Moreover, Z is trace-free as an endomorphism of $\Lambda^1(\mathfrak{g}_E)$.*

Proof. Take $A, B \in \Lambda^1(\mathfrak{g}_E)$. Using the symmetry of both Z and the inner product on E ,

$$\begin{aligned} \langle Z(A), B \rangle_{\Lambda^1(\mathfrak{g}_E)} &= -\frac{1}{2} Z_{ij} A_{j\alpha}^\beta B_{i\beta}^\alpha \\ &= -\frac{1}{2} Z_{ji} B_{i\alpha}^\beta A_{j\beta}^\alpha \\ &= \langle Z(B), A \rangle_{\Lambda^1(\mathfrak{g}_E)}. \end{aligned}$$

The symmetry of Z follows. Next, using the cyclicity of inner products over $\mathfrak{g}_E$, reindexing and skew symmetry of the bracket operation and Φ ,

$$\begin{aligned} \langle [\Phi, A], B \rangle_{\Lambda^1(\mathfrak{g}_E)} &= -\frac{1}{2} [\Phi_{ij}, A_i]_\alpha^\beta B_{j\beta}^\alpha \\ &= -\frac{1}{2} \Phi_{ij\delta}^\beta A_{i\alpha}^\delta B_{j\beta}^\alpha + \frac{1}{2} A_{i\delta}^\beta \Phi_{ij\alpha}^\delta B_{j\beta}^\alpha \\ &= -\frac{1}{2} A_{i\alpha}^\delta B_{j\beta}^\alpha \Phi_{ij\delta}^\beta + \frac{1}{2} A_{i\delta}^\beta \Phi_{ij\alpha}^\delta B_{j\beta}^\alpha \\ &= -\frac{1}{2} A_{i\delta}^\beta B_{j\alpha}^\delta \Phi_{ij\beta}^\alpha + \frac{1}{2} A_{i\delta}^\beta \Phi_{ij\alpha}^\delta B_{j\beta}^\alpha \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{2} A_{i\delta}^\beta [B_j, \Phi_{ij}]_\beta^\delta \\
&= -\frac{1}{2} A_{i\delta}^\beta [\Phi_{ji}, B_j]_\beta^\delta \\
&= \langle [\Phi, B], A \rangle_{\Lambda^1(\mathfrak{g}_E)},
\end{aligned}$$

hence Φ is symmetric as an endomorphism.

To show that Z is trace-free as an operator on $\Lambda^1(\mathfrak{g}_E)$, we construct an orthonormal basis for $\Lambda^1(\mathfrak{g}_E)$ as described at the beginning of Sect. 3.1: for fixed (k, α, β) , let

$$A_{(k, \alpha, \beta)} := e^k \otimes (\mu^*)^\alpha \otimes \mu_\beta, \quad (3.2)$$

where $\{e_i\}$ is a basis of TM that diagonalizes Z . Note that the components of these basis elements are given by

$$(A_{(k, \alpha, \beta)})_{\ell\mu}^v = \delta_{k\ell} \delta_\alpha^v \delta_\mu^\beta, \quad \alpha \neq \beta,$$

so the only nonzero entry is the (k, α, β) -component. Computing the trace of Z with respect to this basis yields

$$\begin{aligned}
\langle Z(A_{(k, \alpha, \beta)}), A_{(k, \alpha, \beta)} \rangle_{\Lambda^1(\mathfrak{g}_E)} &= -\frac{1}{2} Z_{ij} (A_{(k, \alpha, \beta)})_{i\mu}^v (A_{(k, \alpha, \beta)})_{jv}^\mu \\
&= -\frac{1}{2} Z_{ij} \delta_{ki} \delta_\alpha^v \delta_\mu^\beta \delta_{kj} \delta_\alpha^\mu \delta_v^\beta \\
&= -\frac{1}{2} Z_{ii} \delta_\alpha^\beta \delta_\alpha^\beta \\
&= 0,
\end{aligned}$$

since Z is traceless on TM . The result follows. $\square$

Lemma 3.3. *As operators on $\Lambda^1(\mathfrak{g}_E)$, the ranges of Z and Φ are orthogonal subspaces.*

Proof. The orthogonality of Z and $[\Phi, \cdot]$ will follow since Z preserves the bundle components while Φ is skew symmetric with respect to the bundle components. Using the basis (3.2) as above, for fixed (k, α, β) , then

$$\begin{aligned}
(Z(A_{(k, \alpha, \beta)}))_{i\mu}^v &= Z_{\ell i} (A_{(k, \alpha, \beta)})_{\ell\mu}^v \\
&= Z_{\ell i} \delta_{k\ell} \delta_\alpha^v \delta_\mu^\beta \\
&= Z_{ki} \delta_\alpha^v \delta_\mu^\beta \\
&= \begin{cases} Z_{ki} & \text{if } \mu = \alpha, \beta = v, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Similarly,

$$\begin{aligned}
[\Phi, A_{(k, \alpha, \beta)}]_{i\mu}^v &= \Phi_{\ell i\delta}^v (A_{(k, \alpha, \beta)})_{\ell\mu}^\delta - (A_{(k, \alpha, \beta)})_{\ell\delta}^v \Phi_{\ell i\mu}^\delta \\
&= \Phi_{\ell i\delta}^v \delta_{k\ell} \delta_\alpha^\delta \delta_\mu^\beta - \delta_{k\ell} \delta_\alpha^v \delta_\delta^\beta \Phi_{\ell i\mu}^\delta \\
&= \Phi_{ki\alpha}^v \delta_\mu^\beta - \delta_\alpha^v \Phi_{ki\mu}^\beta \\
&= \begin{cases} 0 & \alpha = v \text{ and } \beta = \mu \\ -\Phi_{ki\mu}^\beta & \alpha = v \text{ and } \beta \neq \mu \\ \Phi_{ki\alpha}^v & \alpha \neq v \text{ and } \beta = \mu \\ 0 & \alpha \neq v \text{ and } \beta \neq \mu \end{cases}.
\end{aligned}$$

Where here, we are noting that since $\Phi \in \Lambda^2(\mathfrak{g}_E)$, its endomorphism indices cannot coincide. $\square$

To state our next result, we need to introduce an algebraic invariant defined by Bourguignon–Lawson. Let

$$\gamma_0 := \sup_{A, B \in \Gamma(\mathfrak{g}_E) \setminus \{0\}} \frac{|[A, B]|}{|A||B|}.$$

Lemma 2.30 of [BL81] gives the universal upper bound

$$\gamma_0 \leq \sqrt{2}, \quad (3.3)$$

and characterizes the case of equality.

Lemma 3.4. *If $A \in \Lambda^1(\mathfrak{g}_E)$, then*

$$|[A, A]|_{\Lambda^2(\mathfrak{g}_E)} \leq \gamma_0 \sqrt{\frac{n-1}{2n}} |A|_{\Lambda^1(\mathfrak{g}_E)}^2. \quad (3.4)$$

Since $\gamma_0 \leq \sqrt{2}$, in general we have

$$|[A, A]|_{\Lambda^2(\mathfrak{g}_E)} \leq \sqrt{\frac{n-1}{n}} |A|_{\Lambda^1(\mathfrak{g}_E)}^2. \quad (3.5)$$

Proof. Fix a point $p \in X^n$ and let $\{e^i\}$ to be an orthonormal basis of Λ^1 . If $A \in \Lambda^1(\mathfrak{g}_E)$, then we can express $A = A_i e^i$ for $A_i \in \Gamma(\mathfrak{g}_E)$. Then

$$\begin{aligned} |[A, A]|_{\Lambda^2(\mathfrak{g}_E)}^2 &= -\frac{1}{4} [A, A]_{ij\alpha}^\beta [A, A]_{ij\beta}^\alpha \\ &= \frac{1}{2} \sum_{i,j} |[A, A]_{ij}|_{\mathfrak{g}_E}^2 \\ &= \frac{1}{2} \sum_{i,j} |[A_i, A_j]|_{\mathfrak{g}_E}^2 \\ &= \sum_{i < j} |[A_i, A_j]|_{\mathfrak{g}_E}^2. \end{aligned}$$

By the definition of γ_0 , this gives

$$\begin{aligned} |[A, A]|_{\Lambda^2(\mathfrak{g}_E)}^2 &= \left(\sum_{i < j} |[A_i, A_j]|_{\mathfrak{g}_E}^2 \right) \\ &\leq \gamma_0^2 \left(\sum_{i < j} |A_i|_{\mathfrak{g}_E}^2 |A_j|_{\mathfrak{g}_E}^2 \right). \end{aligned} \quad (3.6)$$

Now

$$|A|_{\Lambda^1(\mathfrak{g}_E)}^4 = \sum_{i,j} |A_i|_{\mathfrak{g}_E}^2 |A_j|_{\mathfrak{g}_E}^2 = 2 \sum_{i < j} |A_i|_{\mathfrak{g}_E}^2 |A_j|_{\mathfrak{g}_E}^2 + \sum_i |A_i|_{\mathfrak{g}_E}^4,$$

while the arithmetic-geometric mean implies

$$\sum_i |A_i|_{\mathfrak{g}_E}^4 \geq \frac{1}{n} \left(\sum_i |A_i|_{\mathfrak{g}_E}^2 \right)^2 = \frac{1}{n} |A|_{\Lambda^1(\mathfrak{g}_E)}^4.$$

Therefore,

$$\sum_{1 \leq i < j \leq n} |A_i|_{\mathfrak{g}_E}^2 |A_j|_{\mathfrak{g}_E}^2 \leq \frac{(n-1)}{2n} |A|_{\Lambda^1(\mathfrak{g}_E)}^4.$$

Substituting this into (3.6) gives

$$|[A, A]|_{\Lambda^2(\mathfrak{g}_E)}^2 \leq \gamma_0^2 \left(\frac{n-1}{2n} \right) |A|_{\Lambda^1(\mathfrak{g}_E)}^4,$$

and taking the square root yields (3.4). $\square$

Proposition 3.5. *Suppose $E \rightarrow (X^n, g)$ is a vector bundle and let*

$$\mathcal{B} = Z + [\Phi, \cdot] : \Lambda^1(\mathfrak{g}_E) \rightarrow \Lambda^1(\mathfrak{g}_E).$$

Then

$$|\mathcal{B}(A, A)| \leq \sqrt{\frac{n-1}{n}} \cdot \left(\sqrt{|Z|_{S_0^2(T^*M)}^2 + 2\gamma_0^2 |\Phi|_{\Lambda^2(\mathfrak{g}_E)}^2} \right) |A|_{\Lambda^1(\mathfrak{g}_E)}^2.$$

Proof. Since $\mathcal{B}$ is symmetric by Lemma 3.2, there exists an orthonormal basis of $\Lambda^1(\mathfrak{g}_E)$ with respect to which the matrix of $\mathcal{B}$ is diagonalized. Since the ranges of Z and Φ are orthogonal by Lemma 3.3, we can express the matrix of $\mathcal{B}$ as

$$[\mathcal{B}] = \begin{pmatrix} \vec{z} & 0 \\ 0 & \vec{\phi} \end{pmatrix},$$

where

$$[Z] = \begin{pmatrix} \vec{z} & 0 \\ 0 & 0 \end{pmatrix}, \quad [\Phi] = \begin{pmatrix} 0 & 0 \\ 0 & \vec{\phi} \end{pmatrix},$$

are the matrices of Z and Φ with respect to this basis, $\vec{z} = (z_1, \dots, z_n)$, $\vec{\phi} = (\phi_1, \dots, \phi_N)$ are the eigenvalues of Z and Φ respectively. If $A \in \Lambda^1(\mathfrak{g}_E)$, then we can write $A = A_1 + A_2$, where

$$A_1 = \begin{pmatrix} \vec{a} \\ 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 \\ \vec{b} \end{pmatrix},$$

with $\vec{a} = (a_1, \dots, a_n)$, $\vec{b} = (b_1, \dots, b_N)$. Therefore, as a bilinear form

$$\begin{aligned} \mathcal{B}(A, A) &= Z(A_1, A_1) + \Phi(A_2, A_2) \\ &= \begin{pmatrix} \vec{z} & 0 \\ 0 & \vec{\phi} \end{pmatrix} \begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix} \cdot (\vec{a} \ \vec{b}) \\ &= \sum_i z_i a_i^2 + \sum_j \phi_j b_j^2. \end{aligned}$$

Since Z is trace-free via Lemma 3.2,

$$\begin{aligned} |Z(A_1, A_1)| &= \left| \sum_i z_i a_i^2 \right| \\ &\leq \sqrt{\frac{n-1}{n}} |\vec{z}| |\vec{a}|^2 \\ &= \sqrt{\frac{n-1}{n}} |Z|_{S_0^2(T^*M)} |A_1|_{\Lambda^1(\mathfrak{g}_E)}^2. \end{aligned}$$

Also, by Lemma 3.4,

$$\begin{aligned} |\Phi(A_2, A_2)| &= \left| \sum_j \phi_j b_j^2 \right| \\ &= |\langle [\Phi, A_2], A_2 \rangle|_{\Lambda^1(\mathfrak{g}_E)} \\ &= 2 |\langle \Phi, [A_2, A_2] \rangle|_{\Lambda^2(\mathfrak{g}_E)} \\ &\leq 2 |\Phi|_{\Lambda^2(\mathfrak{g}_E)} |[A_2, A_2]|_{\Lambda^2(\mathfrak{g}_E)} \\ &\leq 2\gamma_0 |\Phi|_{\Lambda^2(\mathfrak{g}_E)} \sqrt{\frac{n-1}{2n}} |A_2|_{\Lambda^1(\mathfrak{g}_E)}^2. \end{aligned}$$

Therefore,

$$|\mathcal{B}(A, A)| \leq \sqrt{\frac{n-1}{n}} \left(|Z|_{S_0^2(T^*M)} |A_1|_{\Lambda^1(\mathfrak{g}_E)}^2 + \sqrt{2}\gamma_0 |\Phi|_{\Lambda^2(\mathfrak{g}_E)} |A_2|_{\Lambda^1(\mathfrak{g}_E)}^2 \right),$$

where we have dropped the subscripts designating the norms in order to simplify notation. By the Cauchy-Schwartz inequality,

$$\begin{aligned} |\mathcal{B}(A, A)| &\leq \sqrt{\frac{n-1}{n}} \cdot \sqrt{|Z|_{S_0^2(T^*M)}^2 + 2\gamma_0^2 |\Phi|_{\Lambda^2(\mathfrak{g}_E)}^2} \cdot \sqrt{|A_1|_{\Lambda^1(\mathfrak{g}_E)}^4 + |A_2|_{\Lambda^1(\mathfrak{g}_E)}^4} \\ &\leq \sqrt{\frac{n-1}{n}} \cdot \left(\sqrt{|Z|_{S_0^2(T^*M)}^2 + 2\gamma_0^2 |\Phi|_{\Lambda^2(\mathfrak{g}_E)}^2} \right) |A|_{\Lambda^1(\mathfrak{g}_E)}^2. \end{aligned}$$

The result follows. $\square$

3.3. A canonical conformal representative. Since the index and nullity of a Yang–Mills connection in four dimensions are conformally invariant, we may estimate them with respect to any metric conformal to the base metric g . In this subsection, we specify a choice of conformal metric based on our work in [GKS18]. To this end, suppose ∇ is a Yang–Mills connection on a vector bundle E over (X^4, g) with structure group $G \subset \mathrm{SO}(E)$, and denote the curvature by $F = F_\nabla$. For $t \geq 0$, define

$$\Phi_g^t = R_g - t[2\sqrt{6}|W|_g + 3\gamma_1|F|_g],$$

where R_g is the scalar curvature of g , W_g is the Weyl tensor, and $\gamma_1(E)$ is the constant given by

$$\gamma_1(E) := \sup_{\omega \in \Lambda_+^2(\mathfrak{g}_E) \setminus \{0\}} \frac{\langle \omega, [\omega, \omega] \rangle}{|\omega|^3}. \quad (3.7)$$

Remark 3.6. The definition of the inner product on $\Lambda_+^2(\mathfrak{g}_E)$ given in [GKS18] differs from the definition of this paper. In particular, the estimate for $\gamma_1(E)$ in Section 2 of [GKS18] needs to be adjusted. With respect to our current conventions, we have the estimate

$$\begin{aligned}\gamma_1(E) &\leq \frac{2\sqrt{6}}{3}\gamma_0(E) \\ &\leq \frac{4\sqrt{3}}{3}.\end{aligned}\tag{3.8}$$

We also define the associated operator

$$L_g^t = -6\Delta_g + \Phi_g^t.$$

In [GKS18], based on the ideas of [Gur00], we defined the related curvature and operator

$$\begin{aligned}\Phi_g &= R_g - 2\sqrt{6}|W^+|_g - 3\gamma_1|F^+|_g, \\ L_g &= -6\Delta_g + \Phi_g.\end{aligned}\tag{3.9}$$

It is easy to see that the expression $\gamma_1(E)|F|$ is independent of the choice of norms. Therefore, despite the difference of conventions pointed out in Remark 3.6, the definition of Φ_g in (3.9) agrees with the corresponding formula (3.5) in [GKS18].

Observe that

$$\begin{aligned}\Phi_g^0 &= R_g, \\ \Phi_g^1 &\leq \Phi_g.\end{aligned}$$

Note that the latter inequality implies

$$\lambda_1(L_g^1) \leq \lambda_1(L_g).\tag{3.10}$$

In addition, Φ^t satisfies the same kind of conformal transformation formula as Φ : given $\hat{g} = u^2g$,

$$\Phi_{\hat{g}}^t = u^{-3}L_g^t u.$$

If $\lambda_1(L^t)$ denotes the first eigenvalue of L^t ,

$$\lambda_1(L_g^t) = \inf_{\phi \in C^\infty(X) \setminus \{0\}} \frac{\int_X \phi L_g^t \phi \, dV_g}{\int_X \phi^2 \, dV_g},\tag{3.11}$$

then the sign of $\lambda_1(L^t)$ is a conformal invariant (see [Gur00], Proposition 3.2). In particular, by using an eigenfunction associated with $\lambda_1(L^t)$ as a conformal factor, it follows that $[g]$ admits a metric $\hat{g}$ with $\Phi_{\hat{g}}^t > 0$ (*resp.*, $= 0$, < 0) if and only if $\lambda_1(L_g^t) > 0$ (*resp.*, $= 0$, < 0).

Proposition 3.7. *Assume $(X^4, [g])$ has $Y(X^4, [g]) > 0$. Given ∇ a Yang–Mills connection which is not an instanton, there exists $t_0 \in (0, 1]$ such that $\lambda_1(L_g^{t_0}) = 0$. In particular, we can choose a conformal metric $\hat{g} \in [g]$ with respect to which $\Phi_{\hat{g}}^{t_0} \equiv 0$, hence*

$$R_{\hat{g}} = 2\sqrt{6}t_0|W_{\hat{g}}| + 3\gamma_1 t_0|F|_{\hat{g}}.\tag{3.12}$$

Moreover,

$$\frac{Y(X^4, [g])}{2\sqrt{6}\|W\|_{L^2} + 3\gamma_1\|F\|_{L^2}} \leq t_0 \leq 1. \quad (3.13)$$

Proof. Using the Bochner formula for Yang-Mills connections, in [GKS18] we showed that either $F^+ \equiv 0$, or else $\lambda_1(L_g) \leq 0$ (see [GKS18] following (3.8) of the proof of Theorem 1.1). Since we are ruling out the former by assumption, the latter condition must hold. Consequently, by (3.10), $\lambda_1(L_g^1) \leq 0$. In fact, we can assume $\lambda_1(L_g^1) < 0$, since otherwise we could take $t_0 = 1$.

Clearly, $\lambda_1(L_g^t)$ depends continuously on the parameter t . Since $\Phi_g^0 = R_g$ and the Yamabe invariant of $(X^4, [g])$ is positive, we know that $\lambda_1(L_g^0) > 0$. By the intermediate value theorem, it follows there is $t_0 \in (0, 1]$ with $\lambda_1(L_g^{t_0}) = 0$. Also, integrating (3.12) and using the Cauchy-Schwarz inequality it is easy to see that t_0 satisfies (3.13). $\square$

3.4. The proof of Theorem 1.1. In this subsection use Theorem 2.1 to give the proof of Theorem 1.1. As remarked above, since the index and nullity are conformal invariants we are free to make a conformal modification of the base metric and we choose the conformal gauge guaranteed by Proposition 3.7. To begin we obtain an algebraic estimate for the Jacobi operator. Specifically, let Z now denote the trace-free Ricci tensor, i.e.

$$Z := \text{Ric} - \frac{1}{4}Rg.$$

We express $\mathcal{J}^\nabla$ as

$$\begin{aligned} \mathcal{J}^\nabla &= -\Delta + \frac{1}{4}R + Z + 2[F_\nabla, \cdot] \\ &= -\Delta + \frac{1}{6}R + \left\{ \frac{1}{12}R + \frac{\sqrt{3}}{12}\gamma_1 t_0 [F, \cdot] \right\}_{\mathcal{A}} + \left\{ Z + \left(2 - \frac{\sqrt{3}}{12}\gamma_1 t_0 \right) [F, \cdot] \right\}_{\mathcal{B}}, \end{aligned} \quad (3.14)$$

and proceed to estimate the zeroth-order operators $\mathcal{A}$ and $\mathcal{B}$ labeled above.

Lemma 3.8. *As a bilinear form, $\mathcal{A} \geq 0$.*

Proof. If we take $Z = 0$ and $\Phi = F_\nabla$ in Proposition 3.5, then

$$\begin{aligned} |F(A, A)| &= |\langle [F, A], A \rangle| \\ &\leq \frac{\sqrt{3}}{2} \sqrt{2\gamma_0^2 |F|^2 |A|^2} \\ &= \frac{\sqrt{6}}{2} \gamma_0 |F| |A|^2. \end{aligned}$$

Since $\gamma_0 \leq \sqrt{2}$, it follows that

$$|\langle [F, A], A \rangle| \leq \sqrt{3} |F| |A|^2.$$

Therefore,

$$\begin{aligned} \mathcal{A}(A, A) &= \frac{1}{12}R |A|^2 + \frac{\sqrt{3}}{12}\gamma_1 t_0 \langle [F, A], A \rangle \\ &\geq \frac{1}{12}R |A|^2 - \frac{\sqrt{3}}{12}\gamma_1 t_0 \left(\sqrt{3} |F| \right) |A|^2 \\ &= \frac{1}{12} (R - 3\gamma_1 t_0 |F|) |A|^2. \end{aligned}$$

Using the formula for the scalar curvature in (3.12), we conclude

$$\begin{aligned}\mathcal{A}(A, A) &\geq \frac{1}{12} \left(R - 3\gamma_1 t_0 |F| |A|^2 \right) \\ &= \frac{\sqrt{6}}{6} t_0 |W| |A|^2 \\ &\geq 0.\end{aligned}$$

□

Lemma 3.9. *Let*

$$\alpha = 2 - \frac{\sqrt{3}}{12} \gamma_1 t_0 > 0. \quad (3.15)$$

Then

$$\mathcal{B}(A, A) \geq - \left[\frac{3}{4} |Z|^2 + 3\alpha^2 |F|^2 \right]^{1/2} |A|^2. \quad (3.16)$$

Proof. Note that $\mathcal{B} = Z + \alpha[F, \cdot]$. If we take $\Phi = \alpha F$ in Proposition 3.5 and use the fact that $\gamma_0 \leq \sqrt{2}$, then

$$\begin{aligned}\mathcal{B}(A, A) &\geq -\frac{\sqrt{3}}{2} \left[|Z|^2 + 2\gamma_0^2 \alpha^2 |F|^2 \right]^{1/2} |A|^2 \\ &\geq \left[\frac{3}{4} |Z|^2 + 3\alpha^2 |F|^2 \right]^{1/2} |A|^2,\end{aligned}$$

as claimed. □

In view of (3.14) and Lemmas 3.8 and 3.9, we have

$$\begin{aligned}\langle \mathcal{J}^\nabla A, A \rangle_{L^2} &\geq \langle \left(-\Delta + \frac{1}{6} R - \left[\frac{3}{4} |Z|^2 + 3\alpha^2 |F|^2 \right]^{1/2} \right) A, A \rangle_{L^2} \\ &= \langle (-\Delta + \frac{1}{6} R - V) A, A \rangle_{L^2},\end{aligned} \quad (3.17)$$

where

$$V = \left[\frac{3}{4} |Z|^2 + 3\alpha^2 |F|^2 \right]^{1/2}. \quad (3.18)$$

We therefore define

$$\mathcal{S} = -\Delta + \frac{1}{6} R - V. \quad (3.19)$$

To estimate the index and nullity of $\mathcal{J}^\nabla$ it suffices to obtain the estimate for $\mathcal{S}$, since by (3.17) whenever $\mathcal{J}^\nabla$ is nonpositive on a subspace, then so is $\mathcal{S}$. Applying Theorem 2.1 to the operator $\mathcal{S}$ on the bundle $\Lambda^1(\mathfrak{g}_E)$, which has rank $4d$, where $d = \dim(\mathfrak{g}_E)$, we obtain

$$\begin{aligned}N_0(\mathcal{S}) &\leq \frac{144e^2 d}{Y(X^4, [g])^2} \int_X V^2 \, dV \\ &\leq \frac{144e^2 d}{Y(X^4, [g])^2} \left\{ \frac{3}{4} \int_X |Z|^2 \, dV + 3\alpha^2 \int_X |F|^2 \, dV \right\}.\end{aligned} \quad (3.20)$$

By the Chern–Gauss–Bonnet formula

$$\frac{3}{4} \int_X |Z|^2 \, dV = -12\pi^2 \chi(X^4) + \frac{3}{2} \int_X |W|^2 \, dV + \frac{1}{16} \int_X R^2 \, dV. \quad (3.21)$$

Using the conformal gauge fixing of Proposition 3.7, we can estimate the scalar curvature term above as

$$\begin{aligned} \frac{1}{16} \int_X R^2 \, dV &= \frac{t_0^2}{16} \int_X \left(2\sqrt{6}|W| + 3\gamma_1|F| \right)^2 \, dV \\ &= \frac{3}{2}t_0^2 \int_X |W|^2 \, dV + \frac{3\sqrt{6}}{4}\gamma_1 t_0^2 \int_X |W||F| \, dV + \frac{9}{16}\gamma_1^2 t_0^2 \int_X |F|^2 \, dV. \end{aligned}$$

Substituting this into (3.21) gives

$$\begin{aligned} \frac{3}{4} \int_X |Z|^2 \, dV &= -12\pi^2 \chi(X^4) + \frac{3}{2} (1 + t_0^2) \int_X |W|^2 \, dV \\ &\quad + \frac{3\sqrt{6}}{4}\gamma_1 t_0^2 \int_X |W||F| \, dV + \frac{9}{16}\gamma_1^2 t_0^2 \int_X |F|^2 \, dV. \end{aligned}$$

We now substitute this into (3.20) to get

$$\begin{aligned} N_0(S) &\leq \frac{144e^2 d}{Y(X^4, [g])^2} \left\{ -12\pi^2 \chi(X^4) + \frac{3}{2} (1 + t_0^2) \int_X |W|^2 \, dV \right. \\ &\quad \left. + \frac{3\sqrt{6}}{4}\gamma_1 t_0^2 \int_X |W||F| \, dV + \left(3\alpha^2 + \frac{9}{16}\gamma_1^2 t_0^2 \right) \int_X |F|^2 \, dV \right\}. \end{aligned} \quad (3.22)$$

We estimate the coefficients of each of terms above as follows: For the first coefficient, since $t_0 \leq 1$ we have

$$\frac{3}{2} (1 + t_0^2) \leq 3.$$

Since $0 \leq t_0 \leq 1$ and by (3.8) $\gamma_1 \leq \frac{4\sqrt{3}}{3}$, we can bound the second coefficient by

$$\begin{aligned} \frac{3\sqrt{6}}{4}\gamma_1 t_0^2 &\leq \frac{3\sqrt{6}}{4}\gamma_1 \\ &\leq 3\sqrt{2}. \end{aligned} \quad (3.23)$$

For the third coefficient we use the formula for α in (3.15) to write

$$\left(3\alpha^2 + \frac{9}{16}\gamma_1^2 t_0^2 \right) = \frac{5}{8}(\gamma_1 t_0)^2 - \sqrt{3}(\gamma_1 t_0) + 12. \quad (3.24)$$

Now $\gamma_1 t_0 \leq \frac{4\sqrt{3}}{3}$, and the quadratic polynomial $q(x) = \frac{5}{8}x^2 - \sqrt{3}x + 12$ attains its maximum at $x = 0$ on the interval $\left[0, \frac{4\sqrt{3}}{3} \right]$. Consequently,

$$\left(3\alpha^2 + \frac{9}{16}\gamma_1^2 t_0^2 \right) \leq 12.$$

With these estimates on the coefficients, we can rewrite (3.22) as

$$\begin{aligned} N_0(S) &\leq \frac{144e^2 d}{Y(X^4, [g])^2} \left\{ -12\pi^2 \chi(X^4) + 3 \int_X |W|^2 \, dV \right. \\ &\quad \left. + 3\sqrt{2} \int_X |W||F| \, dV + 12 \int_X |F|^2 \, dV \right\}, \end{aligned}$$

finishing the proof. $\square$

3.5. Linear growth rate in four dimensions. Theorem 1.1 exhibits that the index can grow at worst linearly in the Yang–Mills energy of the connection. In this section we show that this growth rate is sharp through an explicit family of examples. Various authors [SJU89, HM90, SS92, Bor92] have shown the existence of families of noninstanton Yang–Mills connection for a given $SU(2)$ bundle over S^4 provided that the charge κ satisfies $\kappa(E) \neq \pm 1$. We will use the work of Sadun–Segert [SS92], who constructed non-instanton Yang–Mills connections on the so-called ‘quadrupole bundles.’ The proposition below analyzes this construction in conjunction with an index estimate of Taubes ([Tau83] Theorem 1.1) to exhibit the required index growth.

Proposition 3.10. *Given $l = 4k - 1 > 1$, let ∇^l denote the Sadun–Segert connection on the quadrupole bundle $P_{(l,3)} \rightarrow S^4$. There exists a constant $\delta > 0$ so that*

$$\iota(\nabla^l) \geq \delta \|F_{\nabla^l}\|_{L^2}^2.$$

Proof. We assume familiarity with the results and notation of [SS92]. The quadrupole bundles are defined by different lifts of the unique irreducible representation of $SU(2)$ on $\mathbb{R}^5$, and are classified by a pair of odd positive integers (n_+, n_-) , with the bundle denoted $P_{(n_+, n_-)}$. The construction of [SS92] further restricts to the case $n_{\pm} \neq 1$. We will choose $n_+ = l = 4k - 1 > 1$, $n_- = 3$, and let ∇^l denote the Sadun–Segert connection on $P_{(l,3)}$. As computed in [SS89, ASSS89] one has

$$\kappa(P_{(n_+, n_-)}) = \frac{1}{8}(n_+^2 - n_-^2) = \frac{1}{8}(l^2 - 9). \quad (3.25)$$

Furthermore, as the connection ∇^l is not self-dual, [Tau83] Theorem 1.1 yields

$$\iota(\nabla^l) \geq 2(|\kappa(P_{(l,3)})| + 1). \quad (3.26)$$

We claim that there exists a constant $C > 0$ so that ∇^l satisfies

$$\|F_{\nabla^l}\|_{L^2}^2 \leq Cl^2. \quad (3.27)$$

Assuming this for the moment, putting together (3.25) - (3.27) yields

$$\begin{aligned} \iota(\nabla^l) &\geq 2(|\kappa(P_{(l,3)})| + 1) \\ &= 2\left(\frac{1}{8}(l^2 - 9) + 1\right) \\ &\geq \frac{1}{4}l^2 \\ &\geq \frac{1}{4C}\|F_{\nabla^l}\|_{L^2}^2, \end{aligned}$$

as required.

We now prove line (3.27). Connections with quadrupole symmetry on these bundles are described in terms of a triple of functions $a_i : (0, \frac{\pi}{3}) \rightarrow \mathbb{R}$, $i = 1, 2, 3$. The bundle on which the connection is defined is determined by the boundary data. In particular, as per ([SS92] Definition 2.5, Lemma 2.6), we require that $a = (a_1, a_2, a_3)$ satisfies

$$\lim_{\theta \rightarrow 0} a(\theta) = (0, 0, l), \quad \lim_{\theta \rightarrow \frac{\pi}{3}} a(\theta) = (0, 3, 0), \quad (3.28)$$

and moreover each a_i extends to $(-\epsilon, \frac{\pi}{3} + \epsilon)$ such that for all $\theta \in (-\epsilon, \epsilon)$,

$$\begin{aligned} a_1(\theta) &= a_2(-\theta), & a_3(\theta) &= a_3(-\theta) \\ a_1\left(\frac{\pi}{3} + \theta\right) &= a_3\left(\frac{\pi}{3} - \theta\right), & a_2\left(\frac{\pi}{3} + \theta\right) &= a_2\left(\frac{\pi}{3} - \theta\right). \end{aligned} \quad (3.29)$$

We can construct a test connection which satisfies these conditions as follows. First set $a_1 \equiv 0$. Fix some small $\delta > 0$ and define a_2 via

$$\begin{aligned} a_2(\theta) &\equiv 0 && \text{for } \theta \in (-\delta, \delta) \\ a_2(\theta) &\equiv 3 && \text{for } \theta \in \left(\frac{\pi}{3} - \delta, \frac{\pi}{3} + \delta\right) \\ 0 \leq a_2(\theta) &\leq 3 && \text{for } \theta \in [0, \frac{\pi}{3}] \\ 0 \leq a_2'(\theta) &\leq 5 && \text{for } \theta \in [0, \frac{\pi}{3}] \end{aligned}$$

and we define a_3 via

$$\begin{aligned} a_3(\theta) &\equiv 3 && \text{for } \theta \in (-\delta, \delta) \\ a_3(\theta) &\equiv 0 && \text{for } \theta \in \left(\frac{\pi}{3} - \delta, \frac{\pi}{3} + \delta\right) \\ 0 \leq a_3(\theta) &\leq 3 && \text{for } \theta \in [0, \frac{\pi}{3}] \\ 0 \geq a_3'(\theta) &\geq -5 && \text{for } \theta \in [0, \frac{\pi}{3}]. \end{aligned}$$

One easily checks that this satisfies conditions (3.28) and (3.29) for $l = 3$. Furthermore, if we set, for $l > 0$,

$$a_l := (a_1, a_2, \frac{l}{3}a_3)$$

then a_l satisfies the conditions of (3.28) and (3.29) for the $(l, 3)$ bundle, and furthermore satisfies

$$0 \leq a_3(\theta) \leq l, \quad 0 \geq a_3'(\theta) \geq -5l.$$

In ([SS92] Proposition 2.7) the Yang-Mills energy of these connections is computed, and takes the form

$$\begin{aligned} \|F_{\nabla(a)}\|_{L^2}^2 &= \pi^2 \int_0^{\frac{\pi}{3}} \left[(a_1')^2 G_1 + (a_1 + a_2 a_3)^2 / G_1 + (a_2')^2 G_2 + (a_2 + a_1 a_3) / G_2 \right. \\ &\quad \left. + (a_3')^2 G_3 + (a_3 + a_1 a_2)^2 / G_3 \right] d\theta, \end{aligned} \quad (3.30)$$

where

$$\begin{aligned} G_1 &= \frac{f_2 f_3}{f_1}, & G_2 &= \frac{f_3 f_1}{f_2}, & G_3 &= \frac{f_1 f_2}{f_3} \\ f_1(\theta) &= 2 \sin\left(\frac{\pi}{3} + \theta\right), & f_2(\theta) &= 2 \sin\left(\frac{\pi}{3} - \theta\right), & f_3(\theta) &= 2 \sin(\theta). \end{aligned}$$

Note that some terms in the energy formula involve factors of the G_i which can blowup at one endpoint or the other, but the boundary conditions for a ensure that these are finite integrals. In particular, for our initial choice of $a = a_3$, we obtain some value for the Yang-Mills energy, call it C . We furthermore observe that every term in (3.30) is at worst quadratic in a_3 and a_3' , which both grow linearly with l , and hence it follows that there is a different constant C such that

$$\|F_{\nabla(a_l)}\|_{L^2}^2 \leq Cl^2.$$

As the Sadun–Segert connection is constructed by energy minimization within this symmetry class ([SS92] Proposition 3.4, Theorem 3.10), its energy must lie below that of this test connection, finishing the proof of (3.27). $\square$

4. The Index of a Positive Einstein Metric

Let X^4 be a smooth, closed, four-dimensional manifold. Furthermore suppose g is a critical point for the normalized total scalar curvature functional given in (1.1):

$$\mathcal{S}[g] = \text{Vol}(g)^{-1/2} \int_{X^4} R_g \, dV_g,$$

where R_g is the scalar curvature of g . It follows that g is an Einstein metric, whose Ricci tensor is given by

$$\text{Ric}(g) = \frac{1}{4} Rg$$

(see [Bes87], Chapter 4C).

To study the second variation of $\mathcal{S}$ at g , one uses the splitting of the space of sections of the bundle of symmetric two-tensors (see [Sch06] for details). The stability operator, corresponding to transverse-traceless variations of g , is given by

$$\begin{aligned} (\mathcal{L}(h))_{ij} &= \Delta h_{ij} + 2R_{ikj\ell} h_{k\ell} \\ &= \Delta h_{ij} + 2W_{ikj\ell} h_{k\ell} - \frac{1}{6} R h_{ij}. \end{aligned} \quad (4.1)$$

This defines an index form

$$\begin{aligned} I(h, h) &= \int_X \langle h, \mathcal{L}(h) \rangle \, dV \\ &= \int_X \left[-|\nabla h|^2 + 2W(h, h) - \frac{1}{6} R|h|^2 \right] \, dV, \end{aligned} \quad (4.2)$$

where

$$W(h, h) = W_{ikj\ell} h_{k\ell} h_{ij}.$$

The index $\iota(g)$ of an Einstein metric is the number of positive eigenvalues of $\mathcal{L}$ (equivalently, the number of negative eigenvalues of $-\mathcal{L}$). The nullity $\nu(g)$ of an Einstein metric is the dimension of the kernel of $\mathcal{L}$, i.e., the dimension of the space of infinitesimal Einstein deformations (see Chapter 12 of [Bes87]). With this background we can give the proof of Theorem 1.3.

Proof of Theorem 1.3. Note that $\mathcal{L} : S_0^2(T^*X^4) \rightarrow S_0^2(T^*X^4)$, where $S_0^2(T^*X^4)$ is the bundle of trace-free symmetric two-tensors. It follows from ([Hui85], Lemma 3.4), that¹

$$-W(h, h) \geq -\frac{2}{\sqrt{3}} |W||h|^2.$$

Therefore,

$$\begin{aligned} \int_X \langle h, -\mathcal{L}(h) \rangle \, dV &\geq \int_X \left[|\nabla h|^2 - \frac{4}{\sqrt{3}} |W||h|^2 + \frac{1}{6} R|h|^2 \right] \, dV \\ &= \int_X \langle h, (-\Delta + \frac{1}{6} R - V)h \rangle \, dV, \end{aligned}$$

¹ Note that in [Hui85], the norm of Weyl is the one induced by the metric on covariant 4-tensors, while we are using the norm of Weyl viewed as a section of $\text{End}(\Lambda^2)$.

where

$$V = \frac{4}{\sqrt{3}}|W|.$$

Since $\dim(S_0^2(T^*X^4)) = 9$, applying Theorem 2.1 to the operator $\mathcal{N} = -\Delta + \frac{1}{6}R - V$ gives

$$\iota(g) + \nu(g) \leq 1728e^2 \frac{\int_X |W|^2 \, dV}{Y(X^4, [g])^2}. \quad (4.3)$$

Since g is Einstein,

$$Y(X^4, [g]) = \mathcal{S}[g]. \quad (4.4)$$

Also, by the Chern–Gauss–Bonnet formula,

$$\begin{aligned} 8\pi^2 \chi(X^4) &= \int_X (|W|^2 + \frac{1}{24}R^2) \, dV \\ &= \int_X |W|^2 \, dV + \frac{1}{24}\mathcal{S}[g]^2. \end{aligned}$$

Substituting this into (4.3), using (4.4), and rearranging the inequality gives

$$\mathcal{S}[g] \leq 24\pi \sqrt{\frac{\chi(X^4)}{3 + \delta [\iota(g) + \nu(g)]}},$$

where $\delta = (24e^2)^{-1}$, as required. $\square$

5. The Proof of Theorem 1.4

Proof of Theorem 1.4. Let (X^4, g) be an oriented four-manifold with positive scalar curvature. To obtain the estimate for the first Betti number we only need to make minor changes to the index estimate for Yang–Mills connections, since the Jacobi operator in the case of the trivial bundle is the Hodge Laplacian acting on Λ^1 . The only difference is the choice of conformal representative: in the trivial case, we use a Yamabe metric in the conformal class of g instead of the metric specified in Proposition 3.7.

Let $\mathcal{H}_1 : \Lambda^1 \rightarrow \Lambda^1$ denote the Hodge Laplacian. Then by the Hodge-de Rham theorem, $H^1(X^4, \mathbb{R}) = \ker \mathcal{H}_1$, and $\dim \ker \mathcal{H}_1 = b_1(X^4)$. Let $\omega \in H^1(X^4, \mathbb{R})$ be a harmonic one-form; by the classical Bochner formula,

$$\begin{aligned} \langle -\mathcal{H}_1 \omega, \omega \rangle_{L^2} &= \int_X (|\nabla \omega|^2 + \text{Ric}(\omega, \omega)) \, dV \\ &= \int_X (|\nabla \omega|^2 + \frac{1}{4}R|\omega|^2 + Z(\omega, \omega)) \, dV \\ &\geq \int_X (|\nabla \omega|^2 + \frac{1}{6}R|\omega|^2 + Z(\omega, \omega)) \, dV. \end{aligned}$$

Since Z is trace-free,

$$Z(\omega, \omega) \geq -\frac{\sqrt{3}}{2}|\omega|^2.$$

Therefore,

$$\begin{aligned} \langle -\mathcal{H}_1 \omega, \omega \rangle_{L^2} &\geq \int_X \left(|\nabla \omega|^2 + \frac{1}{12} R |\omega|^2 - \frac{\sqrt{3}}{2} |Z| |\omega|^2 \right) dV \\ &= \langle (-\Delta + \frac{1}{6} R - V) \omega, \omega \rangle_{L^2}, \end{aligned}$$

where

$$V = \frac{\sqrt{3}}{2} |Z|.$$

Applying Theorem 2.1 to the operator $-\Delta + \frac{1}{6} R - V$ with $\mathcal{E} = \Lambda_1$, we get

$$b_1(X^4) \leq \frac{108e^2}{Y(X^4, [g])^2} \int_X |Z|^2 dV. \quad (5.1)$$

Recall

$$\rho_1(X^4, [g]) = \frac{4 \int \sigma_2(A_g) dV}{Y(X^4, [g])^2} = \frac{\int_X \left(-\frac{1}{2} |Z|^2 + \frac{1}{24} R^2 \right) dV}{Y(X^4, [g])^2}.$$

Since g is a Yamabe metric,

$$\int_X R^2 dV = Y(X^4, [g])^2.$$

Consequently,

$$\begin{aligned} \int_X |Z|^2 dV &= -2\rho_1(X^4, [g]) Y(X^4, [g])^2 + \frac{1}{12} \int_X R^2 dV \\ &= \frac{1}{12} \left(1 - 24\rho_1(X^4, [g]) \right) Y(X^4, [g])^2. \end{aligned}$$

Substituting this into (5.1) gives (1.3).

To estimate $b^+(X^4)$, let $\mathcal{H}_2 : H^2(X^4) \rightarrow H^2(X^4)$ denote the Hodge Laplacian. Then $b^+(X^4) = \dim \ker \mathcal{H}_2^+$, where $\mathcal{H}_2^+$ is the restriction of $\mathcal{H}_2$ to Λ_+^2 , the bundle of self-dual two-forms. The space of self-dual harmonic two-forms is conformally invariant since the Hodge $\star$ operator is. Therefore, in estimating $b^+(X^4)$ we are free to choose a conformal metric. If we take the bundle E to be the trivial bundle in Proposition 3.7, then there is a conformal metric $\hat{g} \in [g]$ and a $t_0 \in (0, 1]$ such that

$$R_{\hat{g}} = 2\sqrt{6}t_0 |W_{\hat{g}}^+|. \quad (5.2)$$

From now on we assume $g = \hat{g}$.

The operator $\mathcal{H}_2^+$ satisfies the Weitzenböck formula

$$\mathcal{H}_2^+ = \Delta + 2W^+ - \frac{1}{3}R,$$

where Δ is the rough Laplacian. Since $W^+ : \Lambda_+^2 \rightarrow \Lambda_+^2$ is trace-free and $\dim \Lambda_+^2 = 3$, we have the sharp inequality

$$|W^+(\omega, \omega)| \leq \frac{2}{\sqrt{6}} |W^+| |\omega|^2.$$

Therefore,

$$\begin{aligned} \langle -\mathcal{H}_2^+ \omega, \omega \rangle_{L^2} &= \int_X \left(|\nabla \omega|^2 - 2W^+(\omega, \omega) + \frac{1}{3}R|\omega|^2 \right) dV \\ &\geq \int_X \left(|\nabla \omega|^2 - \frac{4}{\sqrt{6}}|W^+||\omega|^2 + \frac{1}{3}R|\omega|^2 \right) dV \\ &= \int_X \left(|\nabla \omega|^2 + \frac{1}{6}R|\omega|^2 + \left(\frac{1}{6}R - \frac{4}{\sqrt{6}}|W^+| \right) |\omega|^2 \right) dV. \end{aligned}$$

Using (5.2),

$$\begin{aligned} \langle -\mathcal{H}_2^+ \omega, \omega \rangle_{L^2} &\geq \int_X \left(|\nabla \omega|^2 + \frac{1}{6}R|\omega|^2 - \frac{\sqrt{6}}{3}(2 - t_0)|W^+||\omega|^2 \right) dV \\ &\geq \langle (-\Delta + \frac{1}{6}R - V) \omega, \omega \rangle_{L^2}, \end{aligned}$$

where

$$V = \frac{\sqrt{6}}{3}(2 - t_0)|W^+|.$$

Applying Theorem 2.1 to the operator $-\Delta + \frac{1}{6}R - V$ with $\mathcal{E} = \Lambda_2^+$, we get

$$\begin{aligned} b^+(X^4) &\leq \frac{72e^2}{Y(X^4, [g])^2} (2 - t_0)^2 \|W^+\|_{L^2}^2 \\ &= 3e^2 (2 - t_0)^2 \rho_+(X^4, [g]), \end{aligned} \tag{5.3}$$

where ρ_+ is given by (1.2). By (3.13) of Proposition 3.7,

$$\rho_+^{-1/2} \leq t_0 \leq 1,$$

hence

$$(2 - t_0)^2 \rho_+ \leq (2 - \rho_+^{-1/2})^2 \rho_+ \leq (2\rho_+^{1/2} - 1)^2.$$

Substituting this into (5.3) gives (1.4). $\square$

Acknowledgements The authors thank Elliott Lieb, Francesco Lin, Zhiqin Lu, and Richard Schoen for informative discussions.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- [ASSS89] Avron, J.E., Sadun, L., Segert, J., Simon, B.: Chern numbers, quaternions, and Berry's phases in Fermi systems. *Commun. Math. Phys.* **124**(4), 595–627 (1989)
- [B88] Bérard, P.H.: From vanishing theorems to estimating theorems: the Bochner technique revisited. *Bull. Am. Math. Soc. (N.S.)* **19**(2), 371–406 (1988)
- [Bes87] Besse, A.L.: Einstein manifolds. *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) Results in Mathematics and Related Areas (3)*, 10. Springer-Verlag, Berlin, (1987). xii+510 pp. ISBN: 3-540-15279-2
- [BL81] Bourguignon, J.-P., Lawson Jr., H.B.: Stability and isolation phenomena for Yang–Mills fields. *Commun. Math. Phys.* **79**(2), 189–230 (1981)

- [Bor92] Bor, G.: Yang–Mills fields which are not self-dual. *Commun. Math. Phys.* **145**(2), 393–410 (1992)
- [HM90] Bor, G., Montgomery, R.: *SO(3) Invariant Yang–Mills Fields Which are not Self-dual. Hamiltonian Systems, Transformation Groups and Spectral Transform Methods* (Montreal, PQ, 1989), pp. 191–198. University of Montreal, Montreal (1990)
- [Cwi77] Cwikel, M.: Weak type estimates for singular values and the number of bound states of Schrödinger operators. *Ann. Math. (2)* **106**(1), 93–100 (1977). 35P20 (81.35)
- [DK90] Donaldson, S.K., Kronheimer, P.B.: *The Geometry of Four-Manifolds*. Oxford Mathematical Monographs, p. x+440. Oxford Science Publications, New York (1990). ISBN: 0-19-853553-8
- [DL82] Donnelly, H., Li, P.: Lower bounds for the eigenvalues of Riemannian manifolds. *Mich. Math. J.* **29**(2), 149–161 (1982)
- [DWW05] Dai, X., Wang, X., Wei, G.: On the stability of Riemannian manifold with parallel spinors. *Invent. Math.* **161**(1), 151–176 (2005)
- [DWW07] Dai, X., Wang, X., Wei, G.: On the variational stability of Kähler–Einstein metrics. *Comm. Anal. Geom.* **15**(4), 669–693 (2007)
- [GKS18] Gursky, M.J., Kelleher, C.L., Streets, J.D.: A conformally invariant gap theorem in Yang–Mills theory. *Commun. Math. Phys.* **361**(3), 1155–1167 (2018)
- [Gur98] Gursky, M.: The Weyl functional, de Rham cohomology, and Kähler–Einstein metrics. *Ann. Math. (2)* **148**(1), 315–337 (1998)
- [Gur00] Gursky, M.J.: Four-manifolds with $\delta W^+ = 0$ and Einstein constants of the sphere. *Math. Ann.* **318**(3), 417–431 (2000)
- [HKRV18] Hundertmark, D., Kunstmann, P., Ried, T., Vugalter, S.: Cwikel’s bound reloaded. *arXiv* (2018)
- [HSU77] Hess, H., Schrader, R., Uhlenbrock, D.A.: Domination of semigroups and generalization of Kato’s inequality. *Duke Math. J.* **44**(4), 893–904 (1977)
- [HSU80] Hess, H., Schrader, R., Uhlenbrock, D.A.: Kato’s inequality and the spectral distribution of Laplacians on compact Riemannian manifolds. *J. Differ. Geom.* **15**(1), 27–37 (1980). (1981)
- [Hui85] Huisken, G.: Ricci deformation of the metric on a Riemannian manifold. *J. Differ. Geom.* **21**(1), 47–62 (1985)
- [Koi79] Koiso, N.: A decomposition of the space $\mathcal{M}$ of Riemannian metrics on a manifold. *Osaka Math. J.* **16**(2), 423–429 (1979)
- [Koi82] Koiso, N.: Rigidity and infinitesimal deformability of Einstein metrics. *Osaka Math. J.* **19**(3), 643–668 (1982)
- [LeB91] LeBrun, C.: Explicit self-dual metrics on $C P_2 \# \cdots \# C P_2$. *J. Differ. Geom.* **34**(1), 223–253 (1991)
- [Li80] Li, P.: On the Sobolev constant and the p -spectrum of a compact Riemannian manifold. *Ann. Sci. École Norm. Super. (4)* **13**(4), 451–468 (1980). 13(4):451–468, 1980
- [LY83] Li, P., Yau, S.-T.: On the Schrödinger equation and the eigenvalue problem. *Commun. Math. Phys.* **88**(3), 309–318 (1983)
- [Lie76] Lieb, E.H.: Bounds on the eigenvalues of the Laplace and Schrödinger operators. *Bull. Am. Math. Soc.* **82**(5), 751–753 (1976)
- [Ros72] Rozenbljum, G.V.: Distribution of the discrete spectrum of singular differential operators. *Dokl. Akad. Nauk SSSR* **202**, 1012–1015 (1972). (Russian)
- [Sch06] Schoen, R.M.: Variational theory for the total scalar curvature functional for Riemannian metrics and related topics. In: *Topics in Calculus of Variations* (Montecatini Terme, 1987). *Lecture Notes in Mathematics*, vol. 1365, pp. 120–154. Springer, Berlin (1989)
- [Sim79] Simon, B.: Kato’s inequality and the comparison of semigroups. *J. Funct. Anal.* **32**(1), 97–101 (1979)
- [SJu89] Sibner, L.M., Sibner, R.J., Uhlenbeck, K.: Solutions to Yang–Mills equations that are not self-dual. *Proc. Natl. Acad. Sci. U.S.A.* **86**(22), 8610–8613 (1989)
- [SS89] Sadun, L., Segert, J.: Chern numbers for fermionic quadrupole systems. *J. Phys. A* **124**(4), 595–627 (1989)
- [SS92] Sadun, L., Segert, J.: Non-self-dual Yang–Mills connections with quadrupole symmetry. *Commun. Math. Phys.* **145**(2), 363–391 (1992)
- [Tau82] Taubes, C.: Self-dual Yang–Mills connections on non-self-dual 4-manifolds. *J. Differ. Geom.* **17**(1), 139–170 (1982)
- [Tau83] Taubes, C.: Stability in Yang–Mills theories. *Commun. Math. Phys.* **91**(2), 235–263 (1983)
- [Ura86] Urakawa, H.: Indices and nullities of Yang–Mills fields. *Proc. Am. Math. Soc.* **98**(3), 475–479 (1986)
- [Wey11] Weyl, H.: Über die Asymptotische Verteilung der Eigenwerte. *Nachr. Königl. Ges. Wiss. Göttingen* **1911**, 110–117 (1911)