The Yang–Mills flow for cylindrical end 4-manifolds

David L. Duncan

Abstract

We establish short-time existence and uniqueness results for the Yang–Mills flow on cylindrical end 4-manifolds. We also show long-time existence and infinite-time convergence under certain hypotheses on the underlying data.

Contents

1 Introduction

The Yang–Mills flow is the flow of a natural vector field on the space of connections on a Riemannian n-manifold Z. The critical dimension for the flow is $n = 4$, and this is the dimension we consider here. In the closed case (compact with no boundary), this 4-dimensional flow has been studied extensively by many authors; e.g., [24, 3, 5, 28, 16, 26, 25, 27, 4, 32, 11]. However, the Yang–Mills flow on *non-compact* 4-manifolds is not as well understood.

Here we study the case where Z has cylindrical ends and an asymptotically cylindrical metric. Our main results establish short-, long-, and infinite-time existence and uniqueness results of the flow under certain hypotheses; see Theorems 1.1 and 1.2. Asymptotically cylindrical manifolds arise naturally when considering neck-stretching limits on closed manifolds; e.g., see [20]. Such manifolds also arise in Floer theory [6], where the Yang–Mills minimizers on Z can carry deep and interesting topological information about the 3-manifold ends of Z. Applications of these types have motivated various aspects of our set-up; for an elaboration, see Section 1.1.

Now we describe our set-up in more detail. Unless otherwise stated, Z will always denote a cylindrical end 4-manifold that is oriented and connected. In particular, this means we can write

$$Z = Z_0 \cup_Y ([0, \infty) \times Y),$$

where Y is a closed 3-manifold, and Z_0 is a compact manifold with $\partial Z_0 = Y$. We allow the case where Y has multiple connected components. To simplify the exposition, we assume that Y is non-empty, though the results have extensions to the case where Z is closed. We will use the term *(cylindrical) ends* to refer to $[0, \infty) \times Y$; although, at times, we will abuse terminology and refer to Y as the 'ends' as well.

A metric g_{cyl} on Z is *cylindrical* if it restricts on the ends to have the form

$$g_{cyl}|_{[0,\infty)\times Y} = ds^2 + g^Y,$$

where s is the coordinate-variable on $[0, \infty)$, and g^Y is a fixed metric on Y. A metric g on Z is *asymptotically cylindrical* if there is a cylindrical metric g_{cyl} on Z and some $\beta > 0$ relative to which

$$|\nabla^k_{cyl}(g - g_{cyl})| = O(e^{-\beta s})$$

for all $k \geq 0$. Here ∇_{cyl} is the covariant derivative induced from the Levi-Civita connection of g_{cyl}. We fix an asymptotically cylindrical metric g on Z. Unless otherwise specified, all metric quantities on Z are defined in terms of g.

Let G be a compact, connected Lie group with Lie algebra $\mathfrak{g}$, and fix a principal G-bundle $P \to Z$. We assume that P restricts on the cylindrical ends to be a product

$$P|_{[0,\infty)\times Y} = [0, \infty) \times Q,$$

for some bundle $Q \to Y$. Every principal G-bundle on Z is isomorphic to a bundle of this form.

Fix a Lie group homomorphism $G \to \mathrm{U}(N)$ that is also an immersion. Linearizing allows us to view the Lie algebra $\mathfrak{g} \subseteq \mathfrak{u}(N)$ as a Lie subalgebra of $\mathfrak{u}(N)$. Then we will consider the Ad-invariant inner product

$$\langle \xi, \zeta \rangle := \frac{1}{2\pi^2} \mathrm{tr}(\xi \cdot \zeta^*) = -\frac{1}{2\pi^2} \mathrm{tr}(\xi \cdot \zeta), \tag{1}$$

on $\mathfrak{g}$, where the trace is the one induced from the inclusion $\mathfrak{g} \subseteq \mathfrak{u}(N) \subset \mathrm{End}(\mathbb{C}^N)$. The coefficient $(2\pi^2)^{-1}$ is to ensure we obtain integers for certain characteristic numbers appearing below (see Example 5.3 (b) and Lemma 5.4).

To obtain a good analytic problem, we want to consider only those connections on P that have fixed asymptotics down the cylindrical ends of Z. For this purpose, fix a flat connection a on Q. Let

$$\mathcal{A}(P; a)$$

denote the space of smooth connections on P that, together with all of their derivatives, decay rapidly down the cylindrical end to the fixed connection a. We denote by

$$d_A : \Omega^k(Z, \mathfrak{g}_P) \to \Omega^{k+1}(Z, \mathfrak{g}_P), \qquad \text{and} \qquad F_A \in \Omega^2(Z, \mathfrak{g}_P)$$

the covariant derivative and curvature of a connection $A \in \mathcal{A}(P; a)$. Here $\mathfrak{g}_P \to Z$ is the adjoint bundle associated to P, and $\Omega^k(Z, \mathfrak{g}_P)$ is the space of smooth $\mathfrak{g}_P$-valued k-forms on Z.

Given an initial connection $A_0 \in \mathcal{A}(P; a)$, the *Yang–Mills flow* is given by

$$\partial_\tau A = -d_A^* F_A, \qquad A(0) = A_0, \tag{2}$$

where $A = A(\tau)$ is a path of connections in $\mathcal{A}(P; a)$. This is the formal negative gradient flow of the *Yang–Mills functional*

$$\mathcal{YM} : \mathcal{A}(P; a) \to \mathbb{R}, \qquad A \mapsto \frac{1}{2} \|F_A\|^2_{L^2(Z)}.$$

The critical points of $\mathcal{YM}$ are the *Yang–Mills connections*. The Yang–Mills functional is invariant under the group $\mathcal{G}(P; e)$ of gauge transformations that, together with all of their derivatives, decay rapidly down the ends to the identity gauge transformation e on Q.

Our first main result establishes short-time existence and uniqueness for the flow (2).

Theorem 1.1 (Short-time existence and uniqueness)**.** *Fix a flat connection a on Q, and a smooth initial condition $A_0 \in \mathcal{A}(P; a)$. Then there is some $\tau_0 \in (0, \infty]$, and a smooth solution*

$$A : [0, \tau_0) \longrightarrow \mathcal{A}(P; a)$$

to the Yang–Mills flow (2). If a is irreducible, then the solution A is unique.

This is a special case of Theorem 3.1 in Section 3. The proof we give follows Feehan's monograph [11], which summarizes and expands upon the original short-time existence proofs by Struwe [28] and Kozono–Maeda–Naito [16]. These authors work in the setting of closed 4-manifolds, and the main point we wish to emphasize is that, by fixing the flat connection a at infinity, there is no essential analytic difference in passing from the closed case to the cylindrical end case considered here; see Remark 3.2 (c). This somewhat surprising fact comes down to the observation that, if one is careful with the estimates involved, all relevant norms can be taken to be defined on the full 4-manifold Z.

It would be interesting to see the extent to which this observation extends to other non-compact settings.

Let τ_0 be the *maximal* existence time for which Theorem 1.1 holds. When τ_0 is finite, our analysis produces a characterization of τ_0 in terms of concentration of energy. This characterization is familiar from the work of Struwe [28] and Schlatter [26], and is discussed further in Section 4.

Now we turn to long-time existence and convergence at infinite time. To obtain good long-time behavior, we will work under certain general position and low-energy hypotheses. To state these, fix a flat connection a on Q, and assume this is *acyclic* in the sense that a is irreducible and non-degenerate as a critical point of the Chern–Simons functional; see Section 2.1.1. There is an identity of the form

$$\mathcal{YM}(A) = \mathcal{CS}_P(a) + \|F_A^+\|^2_{L^2(Z)} \tag{3}$$

for all $A \in \mathcal{A}(P; a)$. Here F_A^+ is the self-dual part of the curvature, and $\mathcal{CS}_P(a)$ is the Chern–Simons value, a quantity depending only on a and the topological type of the bundle P. We will say that a connection A is *anti-self dual (ASD)* if $F_A^+ = 0$ and if $\mathcal{YM}(A)$ is finite (this latter condition is not generally required, but we include it for convenience). The identity (3) shows that, when they exist, the ASD connections are the absolute minimizers of $\mathcal{YM}$ on $\mathcal{A}(P; a)$. A related quantity is the *index* $\mathrm{Ind}_P(a) \in \mathbb{Z}$, which we define to be the formal dimension of the ASD moduli space for connections in $\mathcal{A}(P; a)$; see (8). We will say that *all ASD connections on P are regular* if all flat connections on Q are acyclic, and all ASD moduli spaces on $P \to Z$ (with metric g) and on $\mathbb{R} \times Q \to \mathbb{R} \times Y$ (with metric $ds^2 + g_Y$) are cut out transversely.

Theorem 1.2 (Long-time existence and infinite-time convergence)**.** *Assume all ASD connections on P are regular. There is an integer $\mathcal{I}_G > 0$ so that if a is any flat connection on Q with* $\mathrm{Ind}_P(a) < \mathcal{I}_G$*, then there is a positive constant $\eta(a)$ so the following holds.*

Assume the initial condition $A_0 \in \mathcal{A}(P; a)$ satisfies

$$\|F^+_{A_0}\|^2_{L^2(Z)} < \eta(a).$$

Then the flow (2) has a unique solution $A : [0, \infty) \to \mathcal{A}(P; a)$. Moreover, as τ approaches ∞, the connections $A(\tau)$ converge exponentially in $\mathcal{C}^\infty(Z)$ to a unique ASD connection $A_\infty \in \mathcal{A}(P; a)$.

This result follows immediately from Theorems 5.1 and 6.1, below. The constant $\mathcal{I}_G$ is essentially the lowest index $\mathrm{Ind}_P(a)$ for which bubbling can occur in the associated ASD moduli space; see Section 5.1. The constant $\eta(a)$ is defined in Section 5.3, and reflects an energy gap for the Yang–Mills functional on Z, $\mathbb{R} \times Y$, and S^4; our proof of the existence of this energy gap relies on the regularity hypothesis. Having established an energy gap, long-time existence follows from analysis familiar in the closed setting; e.g., see [16]. Indeed, bubbling is the only obstruction to long-time existence, but bubbling is ruled out for initial connections with Yang–Mills value within the energy gap.

The infinite-time convergence claimed in Theorem 1.2 relies on several ingredients. First, since we have excluded bubbling, it follows immediately from Uhlenbeck's weak compactness theorem that we have weak subsequential convergence at infinite time to a Yang–Mills connection A_∞, where the convergence is modulo gauge and on compact sets. A priori, this limiting connection may depend on the subsequence chosen, and it may be the case that the asymptotic limits of A_∞ are not a (i.e., A_∞ may belong to $\mathcal{A}(P; a')$ for some other flat connection a'). This latter phenomenon is due to the possibility of energy escaping down the cylindrical ends—a phenomenon not present in the closed setting. To exclude this possibility, we use the ASD-regularity and small energy assumptions again to show that, for each $k \geq 0$, the path $A(\tau)$ is $W^{k,2}$-Cauchy on the full 4-manifold Z. In particular, this implies A_∞ does in fact belong to $\mathcal{A}(P; a)$, as desired. Moreover, the positive energy gap forces A_∞ to be ASD, as opposed to just Yang–Mills.

In practice, the regularity hypothesis of Theorem 1.2 is frequently *not* satisfied. However, in many cases, it can be achieved by adding a perturbation to the curvature. This perturbation scheme also fits in nicely with various standard applications of gauge theory to low-dimensional topology; see Section 1.1. Consequently, we consider a suitably *perturbed* version of the flow (2). We discuss perturbations in Section 2, laying out axioms we require our perturbations to satisfy. Some of these axioms are analytic in nature, being used to ensure various linear operators are well-behaved and that appropriate convergence results for Yang–Mills connections extend to the perturbed setting. Others are algebraic in nature, and used to ensure that the perturbed curvature and covariant derivative behave in ways familiar from the unperturbed setting. For example, Corollary 2.5 says that the Bianchi identity holds in our perturbed setting (interestingly, it is not a trivial task to arrange for this identity to hold; see Remark 2.4). The Bianchi identity is used, via a perturbed version of (3), to show that a perturbed ASD connection has energy controlled by the value of its asymptotic limit; this is crucial for our long-time existence analysis.

Due to the fairly specific nature of the perturbations we require, we give an existence result for them in Appendix A. The construction uses holonomy perturbations, which are common in the literature when studying ASD connections. However, as pointed out by Kronheimer [17], care must be taken when considering bubbling phenomena in the presence of a holonomy perturbation. Kronheimer showed that even for ASD connections, the best one could generally hope for is convergence in $W^{1,p}$ away from the bubbling set. We will be considering bubbling phenomena for sequences of connections *along the flow*, which is a second order equation and so this $W^{1,p}$-convergence is initially concerning. The key observation that makes holonomy perturbations viable for our set-up is that any bubbling that cannot be a priori excluded results in connections that are ASD (satisfying a first order equation), as opposed to just Yang–Mills (satisfying a second order equation). As we show below, this is just enough to satisfactorily couple with the $W^{1,p}$-convergence. See Remark 5.9 for more details.

Though we work with a perturbation throughout the remainder of this paper, we make no regularity assumptions in our discussion of short-time existence in Sections 3 and 4. In particular, the results of those sections hold in the absence of any perturbation and without regularity hypotheses.

Remark 1.3. *(a) Guo [14] and Sà Earp [21] each consider similar flows on cylindrical-end Kähler manifolds. They employ Kähler techniques to obtain their convergence results.*

(b) See Janner [15] for a similar perturbed Yang–Mills flow over 3-manifolds.

The following remarks address authors working in the case where Z is a closed 4-manifold.

(c) Schlatter [25] proved long-time existence under the assumption that $F^+_{A_0}$ is L^2-small, and the bundle P has small Pontryagin number. Our approach is in many ways similar, with the restriction on the Pontryagin number being replaced by the index assumption on a.

(d) Waldron [32] has ruled out finite-time bubbling under the assumption that either F^+ or F^- does not concentrate in L^2; see also his paper [33] that rules out finite-time bubbling altogether.

(e) Feehan [11] has obtained similar infinite-time convergence results, where he uses the Łojasiewicz–Simon's inequality in place of our (rather strong) index and ASD-regularity assumptions.

Acknowledgments: The author is grateful to his thesis advisor Chris Woodward for his insight and valuable suggestions. He would also like to thank Paul Feehan, Tom Parker, Alex Waldron, and anonymous referees for their helpful comments and suggestions.

1.1 Applications

Here we discuss two applications of the main results in this paper. The first gives an alternative proof of a well-known gluing result in Donaldson/Floer theory. The second is an extension of these ideas to work in progress on the *quilted Atiyah–Floer conjecture*. The results of this section are not used elsewhere in the paper, though they do supply motivation for much of our set-up, including the specific perturbations used, our regularity hypotheses, and our focus on various constants that arise in the analysis.

1.1.1 Gluing ASD connections

For $i = 1, 2$, let Z_i be a connected, oriented, cylindrical end 4-manifold with asymptotically cylindrical metric g_i. Assume the ends of Z_1 are of the form $Y_1 \sqcup Y$ and those of Z_2 are of the form $\overline{Y} \sqcup Y_2$, where $\overline{Y}$ is the 3-manifold Y equipped with the opposite orientation. We require that Y is non-empty and connected, but we make no restrictions on the number of components of Y_1, Y_2.

In each of Z_1 and Z_2, cut off the end associated to Y, and then attach these together to obtain a new 4-manifold as follows:

$$Z(n) := \Big(Z_1 \setminus [n, \infty) \times Y \Big) \cup_{\{n\} \times Y} \Big(Z_2 \setminus [n, \infty) \times \overline{Y} \Big)$$

Then $Z(n)$ is a connected, oriented, cylindrical end 4-manifold with ends $Y_1 \cup Y_2$, and the g_i induce an asymptotically cylindrical metric on $Z(n)$. In this section, we show how Theorem 1.2 can be used to "glue" certain pairs of ASD connections on Z_1 and Z_2 to form an ASD connection on $Z(n)$, provided n is sufficiently large.

Fix principal G-bundles $P_i \to Z_i$ that are translationally-invariant on the end. Let $Q_i \to Y$ be the bundle on Y induced from P_i. We assume that Q_1 and Q_2 are isomorphic bundles, and we fix a bundle isomorphism $\rho : Q_2 \to Q_1$. Using this isomorphism, the bundles P_1, P_2 induce a bundle on $Z(n)$.

We assume all flat connections on Q_1 (and hence on Q_2) are irreducible. By working with a perturbation as in Theorem A.1, we may assume all (perturbed) flat connections on the Q_i are acyclic and all (perturbed) ASD connections on the $Z(n)$ and the Z_i are regular (an application of the Baire category theorem shows that all of these conditions can be met for all n, with the use of a single perturbation). To simplify the discussion, we will drop the bundles and perturbation from the terminology and notation.

We will write connections on $Y_1 \sqcup Y$ as $a_1 \cup b$, with a_1 (resp. b) representing a connection on Y_1 (resp. Y). Given a flat connection $a_1 \cup b$ on $Y_1 \sqcup Y$, we will write $\mathcal{M}_{\mathrm{ASD}}(Z_1; a_1 \cup b)$ for the moduli space of ASD connections on Z_1 that are asymptotic to $a_1 \cup b$; see Section 2.2 for a precise definition of the ASD moduli space. We use similar notation for Z_2 and $Z(n)$. Our regularity hypotheses imply these moduli spaces are all smooth, finite-dimensional manifolds, with dimension given by the index appearing in the statement of Theorem 1.2. When this index is zero, the moduli spaces consist of a finite set of points [6, Ch. 4]. The above-mentioned gluing result is as follows.

Corollary 1.4 (Parabolic Gluing)**.** *There is some $n_0 \geq 0$ so that the following holds for all $n \geq n_0$. For $i = 1, 2$, fix a flat connection a_i on Y_i. Suppose b is a flat connection on Y with*

$$\mathrm{Ind}_{Z_1}(a_1 \cup b) = \mathrm{Ind}_{Z_2}(\rho^* b \cup a_2) = 0. \tag{4}$$

Then there is an injection

$$\Psi_b : \mathcal{M}_{\mathrm{ASD}}(Z_1; a_1 \cup b) \times \mathcal{M}_{\mathrm{ASD}}(Z_2; \rho^* b \cup a_2) \to \mathcal{M}_{\mathrm{ASD}}(Z(n); a_1 \cup a_2).$$

Moreover, every element of $\mathcal{M}_{\mathrm{ASD}}(Z(n); a_1 \cup a_2)$ is in the image of Ψ_b for some b satisfying (4).

Our proof, sketched below, will use ASD connections on Z_1 and Z_2 to construct a nearly-ASD connection on $Z(n)$, and then use the Yang–Mills flow to obtain an actual ASD connection. Though this result itself is not new (e.g., various forms of it appear in [12],[2], and [6]), our use of the Yang–Mills flow differs from traditional approaches, which instead appeal to the implicit function theorem (IFT) to obtain the actual ASD connection. A similar proof-strategy was used by Waldron [32, Theorem 4.1], who gave a proof of Taubes' grafting procedure [29] using the Yang–Mills flow.

One upshot of the flow-theoretic proof is that, as we will see, essentially the only estimates required are L^2-estimates. This is in contrast to the standard IFT approaches, which require norms from higher Sobolev spaces to ensure that satisfactory embedding theorems hold. In the neck-stretching setting relevant to Corollary 1.4, the relative ease of working with L^2-estimates is similar to the relative ease in studying the metric-scaling properties of the semi-norm $\|\nabla f\|_{L^2}$ as opposed to studying those of the Sobolev norm $\|f\|_{W^{1,p}}$. Of course, with stronger hypotheses often come stronger results, and such is the case in the present setting. For example, in extending Corollary 1.4 to moduli spaces of higher dimension, one would like the induced map to be continuous. In the IFT setting, continuity follows readily from the stronger hypotheses. For flows, continuous dependence of the limit on initial conditions takes additional work to establish (we do not carry this out in the present paper).

A particularly interesting special case of Corollary 1.4 is when $Z_i = \mathbb{R} \times Y$ is a cylinder for $i = 1, 2$. In this case, Corollary 1.4 can be used to prove that the instanton Floer homology of Y [12] is an invariant of Y, and so independent of auxiliary choices like the metric and perturbation. More generally, gluing results such as Corollary 1.4 are the key analytic ingredients in establishing TQFT properties of the *relative Donaldson invariants*, which are topological invariants for cylindrical end 4-manifolds; see [2] and [6, Section 6.4]. In particular, these invariants do not depend on the choice of metrics (e.g., the value of n) or any choice of perturbation used to achieve the relevant regularity hypotheses. It is for reasons such as these that we have chosen in this paper to study the larger class of *perturbed* Yang–Mills flows mentioned in the introduction.

Proof of Corollary 1.4 (Sketch). Assume a_1, a_2 and b are, respectively, flat connections on Y_1, Y_2 and Y that satisfy (4). Let A_1 (resp. A_2) be an ASD connection on Z_1 (resp. Z_2) with asymptotic limits $a_1 \cup b$ (resp. $b \cup a_2$). The index assumption (4) implies $\mathrm{Ind}_{Z(n)}(a_1 \cup a_2) = 0$ for all n, and so all relevant moduli spaces consist of a finite set of points; see [6, Section 3.3.].

With the help of cut-off functions supported on the ends associated to Y, the connections A_1 and A_2 can be joined together to form a "preglued" connection A_{n0} on $Z(n)$ that (i) equals A_i on an open set in $Z(n)$ containing $Z_i \setminus [n-1, \infty) \times Y$ and (ii) is approximately ASD in the sense that

$$\lim_{n\to\infty} \|F^+_{A_{n0}}\|_{L^2(Z(n))} + \|d^*_{A_{n0}} F_{A_{n0}}\|_{L^2(Z(n))} = 0;$$

see [6, Section 4.4]. Let $\eta_n(a_1 \cup a_2) > 0$ be the constant from Theorem 1.2, where the subscript of n on this constant records its dependence on the Riemannian manifold $Z(n)$. By an argument similar to the one given in the proof of Theorem 5.7 below, one can show that the infimum $\inf_n \eta_n(a_1 \cup a_2) > 0$ is positive. In particular, it follows that A_{n0} satisfies the hypotheses of Theorem 1.2, provided n is sufficiently large. Let $A_n(A_1, A_2)$ be the ASD connection on $Z(n)$ obtained from the flow starting at A_{n0}. The gauge invariance of the Yang–Mills functional implies that the map $(A_1, A_2) \mapsto A_n(A_1, A_2)$ descends to a well-defined mapping Ψ_b as in the statement of the corollary.

To see that Ψ_b is injective, suppose $A_n(A_1, A_2) = A_n(A_1', A_2')$ for two pairs (A_1, A_2) and (A_1', A_2') of ASD connections. Let A_{n0} be the preglued connection associated to the pair (A_1, A_2), and let $A_{n0}(\tau)$ be the Yang–Mills flow starting at A_{n0}; similarly, write A_{n0}' and $A_{n0}'(\tau)$ for the preglued connection and flow associated to the A_i'. The regularity hypotheses imply that the supremum

$$C_n := \sup_{\tau \geq 0} \|F^+_{A_{n0}}(\tau)\|_{L^2(Z(n))} \|d^*_{A_{n0}(\tau)} F^+_{A_{n0}}(\tau)\|^{-1}_{L^2(Z(n))}$$

is finite, provided n is sufficiently large. Using an argument similar to the one given for Corollary 5.2, one can show that the C_n are uniformly bounded by some constant C. By Remark 6.8, we then have

$$\|A_{n0} - A_n(A_1, A_2)\|_{L^2(Z(n))} \leq C \|d^*_{A_{n0}} F_{A_{n0}}\|_{L^2(Z(n))}.$$

A similar estimate holds in the primed case. Combining this with the construction of the preglued connections and the assumption $A_n(A_1, A_2) = A_n(A_1', A_2')$, we have

$$\begin{aligned}
\sum_{i=1}^{2} \|A_i - A_i'\|_{L^2(Z_i \backslash [n-1,\infty) \times Y)} &\leq \|A_{n0} - A_{n0}'\|_{L^2(Z(n))} \\
&\leq \|A_{n0} - A_n(A_1, A_2)\|_{L^2(Z(n))} \\
&\qquad + \|A_{n0}' - A_n(A_1', A_2')\|_{L^2(Z(n))} \\
&\leq C(\|d^*_{A_{n0}} F_{A_{n0}}\|_{L^2(Z(n))} + \|d^*_{A_{n0}'} F_{A_{n0}'}\|_{L^2(Z(n))}),
\end{aligned}$$

which is going to zero in n. This implies $A_i = A_i'$, as desired.

To see that every element of $\mathcal{M}_{\mathrm{ASD}}(Z(n); a_1 \cup a_2)$ is in the image of Ψ_b for some b, assume otherwise. Then there are integers n_k diverging to ∞, and a sequence A_{n_k} of ASD connections whose gauge equivalence classes are not in the image Ψ_b for any b. By Uhlenbeck's compactness theorem, there is a flat connection b so that, after possibly passing to a subsequence and applying gauge transformations, the restriction of A_{n_k} to the slice $\{n_k\} \times Y \subseteq Z(n_k)$ converges pointwise to b; for a similar argument, see the proof of Claim 1 in the proof of Theorem 5.7. For $i = 1, 2$, use a cut-off function to construct a nearly-ASD connection on Z_i that agrees with A_{n_k} on $Z_i \backslash [n_k - 1, \infty) \times Y$ and is identically equal to b on $[n_k, \infty) \times Y$. Applying Theorem 1.2 to these nearly-ASD connections, we obtain ASD connections on Z_1 and Z_2 that map under Ψ_b to A_{n_k}. This is a contradiction. □

1.1.2 The quilted Atiyah–Floer conjecture

The specific inspiration for the present paper arose out of the author's work on the *quilted Atiyah–Floer conjecture* [10]. Here we briefly describe the conjecture and indicate where Theorem 1.2 fits into the picture.

Given a cylindrical end 4-manifold Z, the conjecture equates the above-mentioned relative Donaldson invariant of Z with the relative quilt invariants of Wehrheim–Woodward [35] [36] and Wehrheim [34]. The conjecture reduces to a claim that, under suitable regularity and metric assumptions, there is an identification between ASD connections on Z, and the minimizers of a different function E associated with Z. This function E is a version of the Dirichlet energy for Riemann surfaces.

These "suitable regularity assumptions" include an assumption that all ASD connections are regular. There is a similar regularity hypothesis on the minimizers of E, and the perturbations considered here are tailored to achieve all of these regularity hypotheses simultaneously. The aforementioned "metric assumptions" are that one works relative to a family g_ϵ of metrics on Z parametrized by $\epsilon > 0$ and degenerating, in a certain sense, as ϵ approaches 0.

An important special case is when $Z = S \times \Sigma$ is the product of a cylindrical end Riemannian 2-manifold (S, g_S) and a closed Riemannian 2-manifold $(\Sigma, \epsilon^2 g_\Sigma)$. In this product setting, the energy E is (a perturbation of) the usual Dirichlet energy for maps from S into the representation variety of Σ; the minimizers are the (perturbed) holomorphic curves. The conjecture in this product case was established by Dostoglou–Salamon [8, 9] for $S = \mathbb{R} \times S^1$ and by Salamon [23] when S has genus zero. Though several other special cases of the conjecture have been established, the conjecture in the general setting is open at the time of writing.

Given a minimizer e of E, one can construct a connection A_e on Z that is almost ASD in the sense that A_e has near-minimal Yang–Mills value. The idea now is to argue as in the proof of Corollary 1.4, and show that the constants relevant to the flow can be taken to be independent of ϵ. Though this is still work in progress, it would then follow from Theorem 1.2 that the Yang–Mills flow starting at A_e exists and converges to an ASD connection. This would establish an injection between the two sets of minimizers relevant to the quilted Atiyah–Floer conjecture. For example, in the product case $Z = S \times \Sigma$, this ϵ-independence of the constants does indeed hold, which can be seen via an argument similar to the n-independence discussed in the proof of Corollary 1.4. As one might expect, for more general Z, this ϵ-independence is more complicated and remains unknown. Nevertheless, as mentioned in Section 1.1.1, since the necessary estimates are only in L^2, they generally are much easier to work with than, say, estimates necessary for an implicit function theorem argument.

2 Perturbations

In Section 2.1 we define a certain class of perturbations that we will use to perturb the flow. After defining this class, we discuss perturbed versions of the Yang–Mills and Chern–Simons functionals; see Section 2.2. Section 2.3 discusses various elliptic operators that are relevant to the flow.

2.1 Definition of the perturbations

Here we define the relevant class of perturbations. We begin by discussing the asymptotic behavior down the cylindrical end Y, then we discuss the perturbation on the rest of Z.

2.1.1 Perturbations on Y

Let $Q \to Y$ be as in the introduction. Fix a map of the form

$$K : \mathcal{A}(Q) \longrightarrow \Omega^2(Y, \mathfrak{g}_Q), \qquad a \longmapsto K_a, \tag{5}$$

where $\mathcal{A}(Q)$ is the space of connections on Q. We will always assume this is gauge equivariant in the sense that

$$K_{u^*a} = \mathrm{Ad}(u^{-1})K_a$$

for all $a \in \mathcal{A}(Q)$ and all gauge transformations u on Q. We will refer to K as a *perturbation on* Y.

Denote the linearization of K at a by

$$dK_a : \Omega^1(Y, \mathfrak{g}_Q) \longrightarrow \Omega^2(Y, \mathfrak{g}_Q).$$

We will always assume K is chosen to satisfy the following axiom.

Axiom 1. *The operator dK_a is symmetric in the sense that*

$$\int_Y \langle dK_a(v) \wedge w \rangle = \int_Y \langle v \wedge dK_a(w) \rangle$$

for all $v, w \in \Omega^1(Y, \mathfrak{g}_Q)$ and $a \in \mathcal{A}(Q)$.

The notation $\langle \cdot \wedge \cdot \rangle$ combines the wedge on forms with the inner product (1) on $\mathfrak{g}$. The next example shows that the axiom is not difficult to arrange.

Example 2.1. *(a) Fix a gauge invariant function $H : \mathcal{A}(Q) \to \mathbb{R}$, and let $(dH)_a \in T_a^*\mathcal{A}(Q)$ be the derivative at a. Then define $K_a \in \Omega^2(Y, \mathfrak{g}_Q)$ by*

$$(dH)_a v = \int_Y \langle K_a \wedge v \rangle$$

for all $v \in \Omega^1(Y, \mathfrak{g}_Q)$. Then this satisfies Axiom 1.

(b) Here we give a variant of the previous construction that will be useful for our existence result in Appendix A. Suppose $\Sigma \subset Y$ is an embedded surface that is closed and oriented. Fix a gauge-invariant function $h : \mathcal{A}(Q|_\Sigma) \to \mathbb{R}$. For $\alpha \in \mathcal{A}(Q|_\Sigma)$, define a 1-form X_α by

$$dh_\alpha(\nu) = \int_\Sigma \langle X_\alpha \wedge \nu \rangle$$

for all $v \in \Omega^1(\Sigma, \mathfrak{g}_P)$. Next, thicken Σ up to a neighborhood $U \times \Sigma \subset Y$, for some interval U. Fix a function $f : U \to \mathbb{R}$ that is supported in the interior of U. Then declare

$$Y_a := df \wedge X_\alpha,$$

where we have written $a|_{U\times\Sigma} = \alpha + \psi\, dt$, with $t : U \times \Sigma \to U$ the projection. This also satisfies Axiom 1.

We set

$$F_{a,K} := F_a - K_a.$$

We will say a connection $a \in \mathcal{A}(Q)$ is *K-flat* if $F_{a,K} = 0$. A K-flat connection $a \in \mathcal{A}(Q)$ is *acyclic* if the matrix

$$\begin{pmatrix} *d_a - *dK_a & -d_a \\ -d_a^* & 0 \end{pmatrix} \tag{6}$$

is injective as an operator on $\Omega^1(Y, \mathfrak{g}_Q) \oplus \Omega^0(Y, \mathfrak{g}_Q)$; the Hodge star appearing here is the one on Y. The primary relevance of Axiom 1 is that it implies the matrix (6) is self-adjoint relative to the L^2-inner product on Y.

2.1.2 Perturbations on Z

Moving to the 4-manifold Z, we are interested in gauge equivariant maps of the form

$$\mathbf{K} : \mathcal{A}(P) \longrightarrow \Omega^2(Z, \mathfrak{g}_P), \qquad A \longmapsto \mathbf{K}_A.$$

We will assume any such $\mathbf{K}$ is translationally-invariant on the cylindrical end, in the following sense: Fix $A \in \mathcal{A}(P)$, and write

$$A|_{[0,\infty)\times Y} = a + p\, ds$$

so $a : [0, \infty) \to \mathcal{A}(Q)$ is a path of connections and $p : [0, \infty) \to \Omega^0(Y, \mathfrak{g}_Q)$ is a path of 0-forms. Then we assume there is some K as in (5) so that

$$\mathbf{K}_A|_{[0,\infty)\times Y} = K_a.$$

Any map $\mathbf{K}$ satisfying the above will be called a *perturbation*, and we will refer to K as the *induced perturbation on Y*. We will say that $\mathbf{K}$ satisfies Axiom 1 if the induced perturbation on Y satisfies Axiom 1.

Remark 2.2. *A particularly special case is when Z is the cylinder $\mathbb{R} \times Y$. Any perturbation on Y determines a canonical translationally-invariant perturbation on $\mathbb{R} \times Y$. We will use $\mathbf{K}^Y$ to denote perturbations on $\mathbb{R} \times Y$ obtained in this way.*

We will also assume our perturbations satisfy certain uniform bounds; these are given in Axiom 2, below. To state these, set

$$\Omega^k := \Omega^k(Z, \mathfrak{g}_P).$$

Then we will use $d^\ell \mathbf{K}_A : \otimes^\ell \Omega^1 \to \Omega^2$ to denote the ℓth derivative (relative to the background derivative on $\mathcal{A}(P)$) of the function $\mathbf{K}$ at $A \in \mathcal{A}(P)$.

Axiom 2 (Analytic Axiom)**.** *For any integers $\ell, k \geq 0$, and $p \in [1, \infty]$, there is a constant $C_{\mathbf{K}}(k, \ell, p)$ so that*

$$\begin{aligned}\|d^\ell \mathbf{K}_A(V_1, V_2, \dots, V_\ell)\|_{W^{k,p}(Z)} \\ \leq C_{\mathbf{K}}(k, \ell, p) \left(1 + \|F_{A,\mathbf{K}}\|^k_{W^{k-1,p}(Z)}\right) \|V_1\|_{W^{k,p}(Z)} \|V_2\|_{W^{k,p}(Z)} \cdots \|V_\ell\|_{W^{k,p}(Z)}\end{aligned}$$

for all connections A, and compactly supported 1-forms $V_1, \dots, V_\ell \in \Omega^1(Z, \mathfrak{g}_P)$.

For example, when $\ell = k = 0$, this gives a uniform bound of the form

$$\|\mathbf{K}_A\|_{L^p(Z)} \leq C_{\mathbf{K}}(0, 0, p)$$

for all connections A. Estimates of this type are satisfied by perturbations defined in terms of the holonomy; see Appendix A for a further discussion.

As in the 3-dimensional case, we set

$$F_{A,\mathbf{K}} := F_A - \mathbf{K}_A.$$

The linearization of the map $A \mapsto F_{A,\mathbf{K}}$ is the operator

$$d_{A,\mathbf{K}} := d_A - d\mathbf{K}_A : \Omega^1 \longrightarrow \Omega^2.$$

Just as the covariant derivative d_A is defined on forms of all degrees, we want to extend $d_{A,\mathbf{K}}$ to an operator on all forms on Z. This extension is necessary to perform integration by parts, and so we use this to motivate the definition. Specifically, we define maps

$$\Omega^0 \xrightarrow{d\mathbf{K}_A} \Omega^1 \xrightarrow{d\mathbf{K}_A} \Omega^2 \xrightarrow{d\mathbf{K}_A} \Omega^3 \xrightarrow{d\mathbf{K}_A} \Omega^4,$$

as follows:

- Declare $d\mathbf{K}_A : \Omega^0 \to \Omega^1$ to be the zero map.
- The map $d\mathbf{K}_A : \Omega^1 \to \Omega^2$ is the linearization of $A \mapsto \mathbf{K}_A$, as above.
- Declare $d\mathbf{K}_A : \Omega^2 \to \Omega^3$ to be the Banach space dual to $d\mathbf{K}_A : \Omega^1 \to \Omega^2$, using the identification $(\Omega^i)^* = \Omega^{4-i}$ coming from integration. That is, this extension satisfies

  $$\int_Z \langle d\mathbf{K}_A(W) \wedge V \rangle = -\int_Z \langle W \wedge d\mathbf{K}_A(V) \rangle$$

  for $W \in \Omega^2, V \in \Omega^1$ (the minus sign is to account for the grading).
- Define $d\mathbf{K}_A : \Omega^3 \to \Omega^4$ to be the zero map. Note that this is the Banach space dual to $d\mathbf{K}_A : \Omega^0 \to \Omega^1$.

Of course, the only interesting part of this is the extension to 2-forms. Due to Axiom 1 and the requirement that $\mathbf{K}$ agrees with K on the ends, the operator $d\mathbf{K}_A$ on Ω^2 is expressible, on the ends, in terms of dK_a.

The point is that $d_{A,\mathbf{K}} : \oplus_k \Omega^k \to \oplus_k \Omega^k$ is its own Banach space adjoint, up to a sign. Similarly, we can form the L^2-Hilbert space adjoint by setting

$$d^*_{A,\mathbf{K}} := -(-1)^{(4-k)(k-1)} * d_{A,\mathbf{K}} * : \Omega^k \longrightarrow \Omega^{k-1}.$$

This satisfies

$$(d_{A,\mathbf{K}} V, W) = \left(V, d^*_{A,\mathbf{K}} W\right)$$

for all compactly supported $V \in \Omega^{k-1}, W \in \Omega^k$, where

$$(\xi, \zeta) := \int_Z \langle \xi \wedge *\zeta \rangle$$

is the L^2-inner product on Z coming from the metric.

It will be convenient if $d_{A,\mathbf{K}}$ and $F_{A,\mathbf{K}}$ satisfy the Bianchi identity $d_{A,\mathbf{K}} F_{A,\mathbf{K}} = 0$, and similar algebraic identities. For this purpose, we impose our next axiom on $\mathbf{K}$.

Axiom 3 (Algebraic Axiom)**.** *The following holds for each $A \in \mathcal{A}(P)$:*

(i) $d\mathbf{K}_A \circ d\mathbf{K}_A = 0$

(ii) $d\mathbf{K}_A(\mathbf{K}_A) = 0$

(iii) $\langle \mathbf{K}_A \wedge \mathbf{K}_A \rangle = 0$

(iv) $d_A(\mathbf{K}_A) = -d\mathbf{K}_A(F_A)$.

Example 2.3. *To construct an example of a perturbation $\mathbf{K}$ satisfying this, repeat the construction of Example 2.1 (b), but interpret $U \times \Sigma$ as a neighborhood in Z (so U is a surface, as opposed to an interval). It is not hard to check that this satisfies Axiom 3. When the neighborhood $U \times \Sigma$ and the function f from Example 2.1 (b) are translationally-invariant on the end $[0, \infty) \times Y$, then the associated perturbation $\mathbf{K}$ is also translationally-invariant and satisfies Axiom 1.*

Remark 2.4. *The use of perturbations on $U \times \Sigma$ defined by coupling functions (or forms) on U with functions (or vector fields) on $\mathcal{A}(\Sigma)$ is an extension of the type of perturbations considered by Dostoglou–Salamon [8]. There are many alternative approaches to perturbations in the literature, but they do not all suffice for our purposes. For example, Kronheimer [17, Section 3] considers perturbations obtained by coupling 2-forms on $U \times \Sigma$ with functions on $\mathcal{A}(U \times \Sigma)$, but these generally do not satisfy Axiom 3, nor do they generally yield the desirable identities listed in Corollary 2.5.*

Note that the gauge equivariance of $\mathbf{K}$ automatically gives

$$d\mathbf{K}_A\,(d_A \phi) = [\mathbf{K}_A, \phi]$$

for all $\phi \in \Omega^0(Z, \mathfrak{g}_P)$. Combining this with Axiom 3, it is not hard to show that $d_{A,\mathbf{K}}$ behaves algebraically like the usual covariant derivative in the following sense.

Corollary 2.5. *Assume Axiom 3. Then the following hold for each $A \in \mathcal{A}(P)$:*

(Curvature Identity)	$d_{A,\mathbf{K}} \circ d_{A,\mathbf{K}}\phi = [F_{A,\mathbf{K}}, \phi], \quad \forall \phi \in \Omega^0(Z, \mathfrak{g}_P)$
(First Bianchi Identity)	$d_{A,\mathbf{K}} F_{A,\mathbf{K}} = 0$
(Second Bianchi Identity)	$d^*_{A,\mathbf{K}} d^*_{A,\mathbf{K}} F_{A,\mathbf{K}} = 0.$

Our final axiom is one that will be used to prove the relevant version of Uhlenbeck compactness. It too is satisfied by perturbations defined using the holonomy; see Appendix A.

Axiom 4 (Compactness Axiom)**.** *Suppose A_n is a sequence of connections with the property that there is some $1 \leq p < \infty$ and a finite set of points $\Omega \subset Z$ so that the A_n are bounded in $W^{1,p}(K)$ for each compact $K \subset Z \backslash \Omega$. Then $\mathbf{K}_{A_n}$ is Cauchy in $L^p(Z)$.*

2.2 Perturbed Chern–Simons and Yang–Mills theory

Let $\mathbf{K}$ be a perturbation, with K the induced perturbation on Y. We assume these satisfy Axioms 1, 2, 3, and 4.

Define the *perturbed Chern–Simons functional* by setting

$$\mathcal{CS}_{K,P} : \mathcal{A}(Q) \longrightarrow \mathbb{R}, \qquad \mathcal{CS}_{K,P}(a) := -\frac{1}{2}\int_Z \langle F_{A,\mathbf{K}} \wedge F_{A,\mathbf{K}} \rangle,$$

where A is any connection in $\mathcal{A}(P; a)$. It follows from the first Bianchi identity that this is independent of the choice of $A \in \mathcal{A}(P; a)$. Similarly, it depends on $\mathbf{K}$ only through its asymptotic value K.

The perturbed Chern–Simons functional is invariant under the set of gauge transformations on Q that can be extended to gauge transformations on P. The critical points of $\mathcal{CS}_{K,P}$ are precisely the K-flat connections, and the upper left-hand component of the matrix (6) represents the Hessian of $\mathcal{CS}_{K,P}$ relative to the L^2-inner product. Consequently, a K-flat connection is acyclic if and only if it is (i) irreducible and (ii) a non-degenerate critical point of $\mathcal{CS}_{K,P}$, modulo gauge.

Fix a K-flat connection a. The *perturbed Yang–Mills functional*, or *energy*, is defined by

$$\mathcal{YM}_{\mathbf{K}}(A) := \frac{1}{2}\|F_{A,\mathbf{K}}\|^2_{L^2(Z)} = \frac{1}{2}\int_Z \langle F_{A,\mathbf{K}} \wedge *F_{A,\mathbf{K}} \rangle$$

for $A \in \mathcal{A}(P; a)$. This is invariant under the action of the gauge group $\mathcal{G}(P; e)$ from the introduction.

Fix a smooth reference connection $A_{ref} \in \mathcal{A}(P; a)$, and define all Sobolev norms relative to A_{ref} and the Levi-Civita connection on Z; for example,

$$\|W\|^p_{W^{2,p}(Z)} = \|W\|^p_{L^p(Z)} + \|\nabla_{A_{ref}} AW\|^p_{L^p(Z)} + \|\nabla^2_{A_{ref}} W\|^p_{L^p(Z)}$$

where $\nabla_{A_{ref}}$ is the full covariant derivative. We will write

$$W^{k,p}(\Omega^i)$$

for the $W^{k,p}$-completion of the subspace of $\Omega^i(Z, \mathfrak{g}_P)$ consisting of forms with compact support; we use similar notation for the L^p-completion. Note that any two asymptotically cylindrical metrics produce equivalent Sobolev norms.

We will write

$$\mathcal{A}^{k,p}(P;a) = A_{ref} + W^{k,p}(\Omega^1)$$

for the completion of $\mathcal{A}(P;a)$ relative to the $W^{k,p}$-Sobolev norm. Then $\mathcal{YM}_{\mathbf{K}}$ extends smoothly to a real-valued function on $\mathcal{A}^{1,2}(P;a)$, and on $\mathcal{A}^{1,2}(P;a) \cap \mathcal{A}^{k,p}(P;a)$ whenever $k \geq 1$ and $(k+1)p > 4$.

We denote by

$$\mathcal{G}^{k+1,p}(P;e)$$

the $W^{k+1,p}$-completion of $\mathcal{G}(P;e)$. When $k \geq 1$ and $(k+1)p > 4$, the group structure on $\mathcal{G}(P;e)$ extends to give $\mathcal{G}^{k+1,p}(P;e)$ and $\mathcal{G}^{2,2}(P;e) \cap \mathcal{G}^{k+1,p}(P;e)$ each the structure of a Banach Lie group; see [6, Section 4.2]. These act smoothly on $\mathcal{A}^{k,p}(P;a)$ and $\mathcal{A}^{1,2}(P;a) \cap \mathcal{A}^{k,p}(P;a)$, respectively, and the Yang–Mills functional is invariant under these actions.

Remark 2.6. *The set $\mathcal{G}^{2,2}(P;e)$ does not inherit a group structure from $\mathcal{G}(P;e)$. This is due to the failure of the Sobolev multiplication theorem at the borderline level.*

The L^2-gradient of $\mathcal{YM}_{\mathbf{K}}$ is $d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}$. In particular, the critical points of $\mathcal{YM}_{\mathbf{K}}$ on $\mathcal{A}^{1,2}(P;a)$ are those connections A that satisfy

$$d^*_{A,\mathbf{K}} F_{A,\mathbf{K}} = 0.$$

We call these connections **K**-*YM*. For any $A \in \mathcal{A}^{1,2}(P;a)$, we have an identity

$$\mathcal{YM}_{\mathbf{K}}(A) = \|F^+_{A,\mathbf{K}}\|^2_{L^2(Z)} + \mathcal{CS}_{K,P}(a), \tag{7}$$

where

$$F^+_{A,\mathbf{K}} := \frac{1}{2}(1 + *)F_{A,\mathbf{K}}$$

is the anti-self dual part. We say a connection is **K**-*ASD* if $F^+_{A,\mathbf{K}} = 0$. The **K**-ASD connections are automatically **K**-YM by the first Bianchi identity. Moreover, it follows from (7) that if there are any **K**-ASD connections in $\mathcal{A}^{1,2}(P;a)$, then they are the *global* minimizers of $\mathcal{YM}_{\mathbf{K}}$ on $\mathcal{A}^{1,2}(P;a)$.

The linearization of the map $A \mapsto F^+_{A,\mathbf{K}}$ at a connection $A \in \mathcal{A}(P;a)$ is the operator

$$d^+_{A,\mathbf{K}} := \frac{1}{2}(1 + *)d_{A,\mathbf{K}} : \Omega^1(Z, \mathfrak{g}_P) \longrightarrow \Omega^+(Z, \mathfrak{g}_P),$$

where $\Omega^+(Z, \mathfrak{g}_P)$ is the $+1$ eigenspace for $*$ on $\Omega^2(Z, \mathfrak{g}_P)$. The coupled operator $d^+_{A,\mathbf{K}} \oplus d^*_{A,\mathbf{K}} : \Omega^1 \to \Omega^+ \oplus \Omega^0$ is elliptic and will play a special role in the analysis that follows.

Theorem 2.7. *Let a be an acyclic K-flat connection, and $A \in \mathcal{A}(P;a)$. Then the bounded operator*

$$d^+_{A,\mathbf{K}} \oplus d^*_{A,\mathbf{K}} : W^{k+1,p}(\Omega^1) \longrightarrow W^{k,p}(\Omega^+) \oplus W^{k,p}(\Omega^0)$$

is Fredholm for all $k \geq 0$ and $1 < p < \infty$.

A proof of Theorem 2.7 can be found in Proposition 3.6 and Section 3.4 of Donaldson's book [6]. Strictly speaking, Donaldson works with cylindrical metrics; however, it is not hard to see that his argument extends to handle asymptotically cylindrical metrics as well. Alternatively, this extension follows from [19, Theorem 6.1] by Lockhart and McOwen.

When a and A are as in Theorem 2.7, we denote the Fredholm index of the associated Fredholm operator by

$$\mathrm{Ind}_{K,P}(a) := \mathrm{Ind}(d^+_{A,\mathbf{K}} \oplus d^*_{A,\mathbf{K}}). \tag{8}$$

This index depends only on a, K (the asymptotic values of $A, \mathbf{K}$), and the topological type of the bundle P.

We will say that a $\mathbf{K}$-ASD connection A is *ASD-regular* if $d^+_{A,\mathbf{K}}$ is surjective. When a is irreducible, this is equivalent to the surjectivity of the operator $d^+_{A,\mathbf{K}} \oplus d^*_{A,\mathbf{K}}$. This regularity condition is a desirable property; indeed, our primary reason for introducing perturbations is to use the freedom in choice of $\mathbf{K}$ to achieve ASD-regularity for all relevant $\mathbf{K}$-ASD connections. This motivates the following definition.

Definition. *A perturbation* $\mathbf{K}$ *is* ASD-regular *if each of the following holds.*

- *All K-flat connections on Q are acyclic.*
- *All finite energy $\mathbf{K}$-ASD connections on P are ASD-regular.*
- *All finite energy $\mathbf{K}_Y$-ASD connections on $\mathbb{R} \times Q$ are ASD-regular.*

In the last bullet, $\mathbf{K}^Y$ is the perturbation on $\mathbb{R} \times Y$ induced from K as in Remark 2.2, and the $\mathbf{K}^Y$-ASD condition should be defined using the cylindrical metric $ds^2 + g^Y$ on $\mathbb{R} \times Y$.

One useful corollary of ASD-regularity is it asserts that, for each K-flat a, the moduli space

$$\mathcal{M}_{ASD,\mathbf{K}}(a) := \left\{ A \in \mathcal{A}^{1,p}(P; a) \,\middle|\, F^+_{A,\mathbf{K}} = 0 \right\} \Big/ \mathcal{G}^{2,p}(P; e)$$

of $\mathbf{K}$-ASD connections is a smooth manifold with dimension given by $\mathrm{Ind}_{K,P}(a)$. Here we need to assume $p > 2$ in order to have a good gauge group.

Remark 2.8. *Elsewhere in the literature, the term we are calling 'ASD-regular' is often simply called 'regular'. We have introduced the prefix 'ASD' to help distinguish the term from the function-theoretic notion of regularity.*

2.3 Elliptic operators

Fix a perturbation $\mathbf{K}$ as in Section 2.1, as well as a K-flat connection a. As usual, we assume that $\mathbf{K}$ satisfies Axioms 1, 2, 3, and 4; we make no assumptions on a (e.g., a may be reducible or degenerate). Consider the second order operator

$$\Delta_{A,\mathbf{K}} := d_{A,\mathbf{K}} d^*_{A,\mathbf{K}} + d^*_{A,\mathbf{K}} d_{A,\mathbf{K}}.$$

on $\Omega^k(Z, \mathfrak{g}_P)$. When there is no perturbation, we set $\Delta_A := \Delta_{A,0}$. The following provides a global elliptic regularity estimate for this operator; it is a variant of [28, Lemma 3.1].

Lemma 2.9. *For each $A \in \mathcal{A}(P; a)$, there is a constant C_A so that*

$$\|W\|^2_{W^{2,2}(Z)} \leq C_A \left(\|W\|^2_{L^2(Z)} + \|\Delta_{A,\mathbf{K}} W\|^2_{L^2(Z)} \right)$$

for all $W \in W^{2,2}(\Omega^i)$. The constant C_A depends on A only through the value of $\|A - A_{ref}\|_{\mathcal{C}^1(Z)}$.

Proof. It is not hard to see that there is an estimate

$$\|W\|_{W^{2,2}(Z)} \leq C(A) \left(\|W\|_{L^2(Z)} + \|\nabla_A^* \nabla_A W\|_{L^2(Z)} \right), \tag{9}$$

for all $W \in W^{2,2}(\Omega^i)$, where the constant depends on A through the norm $\|A - A_{ref}\|_{\mathcal{C}^1(Z)}$. By (9), it suffices to bound the L^2-norm of $\nabla_A^* \nabla_A W$ in terms of the norms of W and $\Delta_{A,\mathbf{K}} W$. For this, we note the following Weitzenböck formula

$$\begin{aligned} \nabla_A^* \nabla_A W \quad = \quad & \Delta_{A,\mathbf{K}} W + F_{A,\mathbf{K}} \# W + \mathrm{Rm} \# W \\ & + d\mathbf{K}_A(d^*_{A,\mathbf{K}} W) + d\mathbf{K}^*_A(d_{A,\mathbf{K}} W) + d_{A,\mathbf{K}}(d\mathbf{K}^*_A W) + d^*_{A,\mathbf{K}}(d\mathbf{K}_A W) \\ & + \mathbf{K}_A \# W + d\mathbf{K}_A(d\mathbf{K}^*_A W) + d\mathbf{K}^*_A(d\mathbf{K}_A W), \end{aligned}$$

which the reader can check reduces to the familiar *unperturbed* Weitzenböck formula upon expanding and canceling the perturbation terms on the right. Here we use # to denote an algebraic bilinear operator. Obtaining the necessary bounds at this stage is straight-forward; the relevant estimates for the perturbation terms are supplied by Axiom 2. □

Lemma 2.9 has extensions to other Sobolev norms as well. As a corollary, we have the following, which will be used to solve the linearized flow equation in Section 3.

Theorem 2.10. *Let a be any K-flat connection, and $A \in \mathcal{A}^{1,2}(P; a)$. Then the operator $\Delta_{A,\mathbf{K}}$ is self-adjoint as an unbounded operator on $L^2(\Omega^i)$ for $0 \leq i \leq 4$.*

Proof. Since $\Delta_{A,\mathbf{K}}$ is symmetric and positive, the self-adjoint property for $\Delta_{A,\mathbf{K}}$ follows from standard arguments using the regularity estimates as in Lemma 2.9. (Alternatively, when $\mathbf{K}$ is small, that $\Delta_{A,\mathbf{K}}$ is self-adjoint then follows immediately from the Kato–Rellich Theorem and the fact that Δ_A is self-adjoint.) □

We will now give more refined estimates that use a curvature assumption to obtain constants that are independent of the connection A. This will be useful in our bubbling analysis of Section 4. The following is a variant of [28, Lemma 3.3].

Lemma 2.11. *There are constants $C, \delta > 0$ with the following significance. Suppose $A \in \mathcal{A}^{1,2}(P; a)$ and $R > 0$ satisfy*

$$\sup_{x \in Z} \int_{B_R(x)} |F_A|^2 \leq \delta.$$

Then

$$\|W\|^2_{L^4(Z)} + \|\nabla_A W\|^2_{L^2(Z)} \leq C \left(R^{-2} \|W\|^2_{L^2(Z)} + \|d_{A,\mathbf{K}} W\|^2_{L^2(Z)} + \|d^*_{A,\mathbf{K}} W\|^2_{L^2(Z)} \right)$$

for all smooth forms $W \in W^{1,2}(\Omega^i)$.

Proof. Kato's inequality $|d|W|| \leq |\nabla_A W|$ combines with the Sobolev embedding $W^{1,2}(Z) \subset L^4(Z)$ for real-valued functions to give an estimate of the form

$$\|W\|_{L^4(Z)} \leq C \left(\|W\|_{L^2(Z)} + \|\nabla_A W\|_{L^2(Z)} \right)$$

with a constant C that is independent of A and W. It therefore suffices to establish the claimed estimate for $\|\nabla_A W\|_{L^2(Z)}$.

For this, let $(\cdot, \cdot)$ denote the L^2-inner product on Z. The Weitzenböck formula from the proof of Lemma 2.9 gives

$$\begin{aligned}
\|\nabla_A W\|^2_{L^2(Z)} &= (\nabla^*_A \nabla_A W, W) \\
&= \|d_{A,\mathbf{K}} W\|^2 + \|d^*_{A,\mathbf{K}} W\|^2 + (F_{A,\mathbf{K}} \# W, W) + (\mathrm{Rm} \# W, W) \\
&\quad + 2(d^*_{A,\mathbf{K}} W, d\mathbf{K}^*_A W) + 2(d_{A,\mathbf{K}} W, d\mathbf{K}_A W) \\
&\quad + (\mathbf{K}_A \# W, W) + \|d\mathbf{K}_A W\|^2 + \|d\mathbf{K}^*_A W\|^2 \\
&\leq 2 \left(\|d_{A,\mathbf{K}} W\|^2 + \|d^*_{A,\mathbf{K}} W\|^2 + \|d\mathbf{K}_A W\|^2 + \|d\mathbf{K}^*_A W\|^2 \right) \\
&\quad + (F_A \# W, W) + (\mathrm{Rm} \# W, W) + (\mathbf{K}_A \# W, W).
\end{aligned}$$

The perturbation and Riemannian curvature terms can be estimated using Axiom 2 and the fact that Z has bounded geometry. As for the curvature term $(F_A \# W, W)$, it follows as in [28, Lemma 3.3] that there is a constant C, depending only on the Riemannian manifold Z, so that

$$|(F_A \# W, W)| \leq \delta C \Big(R^{-2} \|W\|_{L^2(Z)} + \|\nabla_A W\|_{L^2(Z)} \Big),$$

where δ, R are as in the statement of the lemma. The lemma then follows by taking δ small enough so $\delta C < 1$. □

3 Short-time existence

Fix a perturbation $\mathbf{K}$, and let K be the induced perturbation on Y. We assume throughout this section that $\mathbf{K}$ satisfies Axioms 1, 2, and 3 from Section 2.1

(Axiom 4 and ASD-regularity are *not* assumed here). Suppose $a \in \mathcal{A}(Q)$ is a K-flat connection on Q, not necessarily acyclic. Fix a smooth reference connection $A_{ref} \in \mathcal{A}^{1,2}(P;a)$, and define all Sobolev norms relative to this connection.

Consider the perturbed Yang–Mills functional $\mathcal{YM}_{\mathbf{K}} : \mathcal{A}^{1,2}(P;a) \to \mathbb{R}$. The *(perturbed) Yang–Mills flow* is the negative gradient flow of $\mathcal{YM}_{\mathbf{K}}$:

$$\partial_\tau A = -d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}, \qquad A(0) = A_0, \tag{10}$$

where $A_0 \in \mathcal{A}^{1,2}(P;a)$ is some fixed initial condition, and the unknown A is a path in $\mathcal{A}^{1,2}(P;a)$. Our main short-time existence theorem is as follows.

Theorem 3.1 (Short-time existence)**.** *Let $\mathbf{K}$ and $a \in \mathcal{A}(Q)$ be as above. Fix $4 < p < \infty$, as well as an initial condition $A_0 \in \mathcal{A}^{1,2}(P;a) \cap \mathcal{A}^{2,p}(P;a)$. Then there is some $\tau_0 > 0$, and a solution $A : [0,\tau_0) \to \mathcal{A}^{1,2}(P;a)$ to the perturbed Yang–Mills flow (10), with regularity*

$$A \in \mathcal{C}^1\Big((0,\tau_0), \mathcal{A}^{1,q}(P;a) \Big) \cap \mathcal{C}^0\Big((0,\tau_0), \mathcal{A}^{2,q}(P;a) \Big) \tag{11}$$

for $2 \leq q \leq p$. The curvature has regularity

$$F_A \in \mathcal{C}^1\Big((0,\tau_0), W^{1,q}(Z) \Big) \cap \mathcal{C}^0\Big((0,\tau_0), W^{2,q}(Z) \Big) \tag{12}$$

for $2 \leq q \leq p$. At time $\tau = 0$, the path $A(\tau)$ is $W^{2,2}(Z)$-continuous

$$\lim_{\tau \searrow 0} \|A(\tau) - A_0\|_{W^{2,2}(Z)} = 0. \tag{13}$$

If the K-flat connection a is irreducible, then the solution A is unique. If $A_0 \in \mathcal{C}^\infty(Z)$ is smooth, then the solution $A \in \mathcal{C}^\infty([0,\tau_0) \times Z)$ is smooth in all variables.

Our proof is given below. Theorem 1.1 from the introduction follows by considering the case $\mathbf{K} = 0$ (this trivially satisfies Axioms 1, 2, and 3).

Remark 3.2. *(a) We use Hilbert space techniques to show the $W^{2,2}(Z)$-continuity at time 0 in (13). Since the initial condition A_0 is in $\mathcal{A}^{2,p}(P;a)$, it seems likely that (13) can be improved to*

$$\lim_{\tau \searrow 0} \|A(\tau) - A_0\|_{W^{2,p}(Z)} = 0.$$

However, we do not pursue this further.

(b) The flow (10) is not parabolic due to its invariance under the action of the gauge group. One consequence of this is that the flow is typically not smoothing. For example, suppose the initial condition A_0 is a Yang–Mills connection. Then the constant path $A(\tau) = A_0$ clearly solves (10), but it only has as much regularity as A_0. By applying a gauge transformation with regularity no higher than $\mathcal{G}^{3,p}(P;e)$, one can always construct Yang–Mills connections that are in $\mathcal{A}^{2,p}(P;a)$ but do not have any higher regularity. This shows that we cannot generally expect $A(\tau)$ to have any higher spatial regularity than A_0.

Note that (11) implies $A(\tau) \in \mathcal{A}^{2,p}(P;a)$ for all $\tau < \tau_0$. In this sense, the spatial regularity claimed in (11) is sharp.

(c) In the absence of a perturbation, Struwe's argument [28] carries over directly to our setting to prove short-time existence and the weak regularity result in [28, Theorem 2.3(i)]. This follows from two observations: First, the operator Δ_A from Section 2.3 is self-adjoint, and this is sufficient to import Struwe's discussion of the linearized problem. To import his discussion of the nonlinear problem, note that Struwe estimates the nonlinear terms through the borderline Sobolev embedding $W^{1,2}(Z) \subset L^4(Z)$, which continues to hold in our non-compact setting.

On the other hand, it takes a little more work to extend Struwe's argument to handle the case of non-zero perturbations, and to establish the specific regularity claims of Theorem 3.1. Consequently, in the proof we give below, we appeal to the more general framework laid out by Feehan [11].

Proof of Theorem 3.1. In light of Remark 3.2 (c), and since short-time existence in the closed case is well-treated in the literature (see [7, 28, 16, 11]), we will only sketch the basic proof of Theorem 3.1, emphasizing the aspects that are new to our situation. We refer primarily to the monograph [11] by Feehan, since it is quite exhaustive and provides a nice overview of the various approaches. Once we establish existence, then we work to establish the claimed regularity.

The (now standard) first step in establishing short-time existence for the Yang–Mills flow is to follow Donaldson's variant of the 'de Turck trick'. Here, one first solves the equation

$$\partial_\tau B + d^*_{B,\mathbf{K}} F_{B,\mathbf{K}} + d_{B,\mathbf{K}} d^*_{B,\mathbf{K}} (B - A_{ref}) = 0, \qquad B(0) = A_0, \tag{14}$$

for a path B in $\mathcal{A}^{1,2}(P;a)$, where A_{ref} is the fixed smooth reference connection. To solve (14), note that the linearization of its left-hand side produces the operator $\partial_\tau + \Delta_{B,\mathbf{K}}$. Since $\Delta_{B,\mathbf{K}}$ is self-adjoint and positive, it follows from general theory that a solution B to (14) exists on $[0,\tau_0)$ for some $\tau_0 > 0$. Standard bootstrapping for the heat equation implies that this solution has regularity

$$B \in \mathcal{C}^0\left([0,\tau_0)\,, \mathcal{A}^{1,2}(P;a) \cap \mathcal{A}^{2,p}(P;a)\right) \cap \mathcal{C}^\infty\left((0,\tau_0) \times Z\right).$$

If A_0 is in $\mathcal{C}^\infty(Z)$, then the flow B is in $\mathcal{C}^\infty([0,\tau_0) \times Z)$. For a reference on these regularity assertions, see [11, Theorems 16.4, 16.5].

The next step is to transform our solution B from (14) into a solution of the perturbed Yang–Mills flow (10). To do this, solve the equation

$$u^{-1}\partial_\tau u = -d^*_{B,\mathbf{K}}(B - A_{ref}), \qquad u(0) = e \tag{15}$$

for a path u of gauge transformations on Z. Given the regularity of B, this has a unique solution u, with regularity

$$u \in \mathcal{C}^0\left([0,\tau_0)\,, \mathcal{G}^{1,p}(P;e)\right) \cap \mathcal{C}^1\left((0,\tau_0)\,, \mathcal{G}^{1,p}(P;e)\right).$$

Here we are using the assumption that $p > 4$ in order to obtain good Sobolev multiplication results (e.g., a well-defined gauge group). Moreover, if A_0 is $\mathcal{C}^\infty(Z)$, then $d^*_{B,\mathbf{K}}(B - A_{ref})$ is smooth, and so u is $\mathcal{C}^\infty$ on $[0, \tau_0) \times Z$. See [11, Lemma 20.1]. We set

$$A(\tau) := \left(u(\tau)^{-1}\right)^* B(\tau),$$

and so the regularity on u and B give

$$A \in \mathcal{C}^0\left([0, \tau_0), \mathcal{A}^{0,p}(P; a)\right) \cap \mathcal{C}^1\left((0, \tau_0), \mathcal{A}^{0,p}(P; a)\right).$$

Suppose for the moment that A is smooth. Then one can check directly that A satisfies the flow (10). When a is irreducible, all elements of $\mathcal{A}^{1,2}(P; a)$ are irreducible. In particular, $A(\tau)$ is irreducible for all τ, so it follows from the argument of [28, Section 6] that the solution A to (14) is unique. See also [11, Section 19.2]. More generally, these conclusions hold provided A has high enough regularity to express $d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}$ classically (e.g., if this is in L^p). That is, to finish the proof, it suffices to assume $A_0 \in \mathcal{A}^{1,2} \cap \mathcal{A}^{2,p}$, and show that A has the claimed regularity (11), (12), and (13).

Remark 3.3. *(a) Note that the inhomogeneous term in (15) is smooth for positive time, but not necessarily at time zero since $B(0) = A_0$. Any irregularity of B at time $\tau = 0$ will generally lead to irregularity of $u(\tau)$ and hence $A(\tau)$, even for positive τ. This is expected, given the observation that the flow (10) is not smoothing; see Remark 3.2 (b).*

(b) If A_0 has less regularity than $W^{2,p}$, then the above proof breaks down. For example, if A_0 is only in $\mathcal{A}^{1,2}(P; a)$, then u is only in $\mathcal{G}^{0,2}(P; e)$, which is neither a Lie group, nor does it act on the space of connections.

In what follows, we will use notation such as

$$\mathcal{C}^\ell(W^{k,p}) := \mathcal{C}^\ell((0, \tau_0), W^{k,p}(Z))$$

for the space of $\mathcal{C}^\ell$ maps from $(0, \tau_0)$ into any space of $W^{k,p}$-sections on Z.

We begin our regularity discussion by showing that A and F_A are each $\mathcal{C}^0(L^p \cap \mathcal{C}^0)$. Towards this end, note that gauge equivariance of the curvature and covariant derivative give

$$F_{A,\mathbf{K}} = \mathrm{Ad}(u) F_{B,\mathbf{K}}, \qquad d^*_{A,\mathbf{K}} F_{A,\mathbf{K}} = \mathrm{Ad}(u) d^*_{B,\mathbf{K}} F_{B,\mathbf{K}}.$$

The connection B is smooth for positive time, and u is continuous, so this shows that $F_{A,\mathbf{K}}$ and $d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}$ are in $\mathcal{C}^0(L^p \cap \mathcal{C}^0)$. The same regularity holds in the absence of the perturbation, and so our regularity is sufficient to interpret the flow equation $\partial_\tau A = -d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}$ classically as an equation in $\mathcal{C}^0(L^p \cap \mathcal{C}^0)$. We know that the initial condition A_0 is in $L^p(Z) \cap \mathcal{C}^0(Z)$, so it follows from the identity

$$A(\tau) = A_0 + \int_0^\tau \partial_\tau A$$

that A is in $\mathcal{C}^0(L^p \cap \mathcal{C}^0)$.

Now we bootstrap. We claim that if $A_0 \in W^{k,p}(Z)$, then

$$A, F_A \in \mathcal{C}^0(W^{k-1,p}) \Longrightarrow A, F_A \in \mathcal{C}^0(W^{k,p}).$$

We illustrate the argument for $k = 1$; the more general case is similar.

To show that F_A is in $\mathcal{C}^0(W^{1,p})$, it suffices to show that

$$d_{A_{ref}} F_A, \qquad d^*_{A_{ref}} F_A$$

are each in $\mathcal{C}^0(W^{0,p})$. For this, use the identity

$$d_{A_{ref}} = d_A + \left[A_{ref} - A \wedge \cdot\right],$$

and the Bianchi identity to write

$$\begin{aligned} d_{A_{ref}} F_A &= d_A F_A + \left[A_{ref} - A \wedge F_A\right] \\ &= \left[A_{ref} - A \wedge F_A\right] \\ \\ d^*_{A_{ref}} F_A &= d^*_A F_A - * \left[A_{ref} - A \wedge *F_A\right] \\ &= \mathrm{Ad}(u) d^*_B F_B + * \left[A - A_{ref} \wedge *F_A\right]. \end{aligned}$$

The right-hand side of each of these is in $\mathcal{C}^0(W^{0,p})$, so this gives $F_A \in \mathcal{C}^0(W^{1,p})$. Similar computations/claims hold in the presence of perturbations.

To show that $A \in \mathcal{C}^0(W^{1,p})$, it suffices to show that

$$d_{A_{ref}}(A - A_{ref}), \qquad d^*_{A_{ref}}(A - A_{ref})$$

are both in $\mathcal{C}^0(W^{0,p})$. For the first of these, write

$$d_{A_{ref}}(A - A_{ref}) \quad = \quad F_A - F_{A_{ref}} - \tfrac{1}{2}\left[A - A_{ref} \wedge A - A_{ref}\right].$$

which is clearly in $\mathcal{C}^0(W^{0,p})$. For the second of these, we first claim that the time derivative $\partial_\tau d^*_{A_{ref}}(A - A_{ref})$ is in $\mathcal{C}^0(W^{0,p})$. To see this, write

$$\begin{aligned} \partial_\tau d^*_{A_{ref}}(A - A_{ref}) &= -d^*_{A_{ref}} d^*_{A,\mathbf{K}} F_{A,\mathbf{K}} \\ \\ &= * \left[A_{ref} - A \wedge *d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\right] - *d\mathbf{K}_A(*d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}) \end{aligned}$$

where we used the second Bianchi identity in the second line. The right-hand side is in $\mathcal{C}^0(W^{0,p})$. Recall that we have also assumed that the initial condition A_0 is in $W^{1,p}(Z)$. Then the fact that $d^*_{A_{ref}}(A - A_{ref})$ is in $\mathcal{C}^0(W^{0,p})$ now follows from the identity

$$d^*_{A_{ref}}(A(\tau) - A_{ref}) = d^*_{A_{ref}}(A_0 - A_{ref}) + \int_0^\tau \partial_\tau d^*_{A_{ref}}(A - A_{ref}),$$

and the observations we have just made about the right-hand side. This finishes the inductive step for the case $k = 1$; note that we only used the regularity on A_0 at the end.

In summary, we have shown that if $A_0 \in A^{2,p}(P; a)$, then

$$A \in \mathcal{C}^0((0,\tau_0), \mathcal{A}^{2,p}(P;a)), \qquad F_A \in \mathcal{C}^0((0,\tau_0), W^{2,p}(Z)),$$

and if A_0 is smooth, then A and F_A are smooth in the spatial variables. We now focus on establishing the claimed regularity in the time-variable. We have $\partial_\tau A = -d^*_{A,\mathbf{K}} F_{A,\mathbf{K}} \in \mathcal{C}^0(W^{1,p})$, and so it follows that

$$A \in \mathcal{C}^1((0,\tau_0), \mathcal{A}^{1,p}(P;a)).$$

As for the curvature, the flow gives

$$\partial_\tau F_{A,\mathbf{K}} = -d_{A,\mathbf{K}} d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}.$$

We have already seen that the right-hand side is in $\mathcal{C}^0(L^p)$, and similar arguments show that it is in $\mathcal{C}^0(W^{1,p})$. Hence, $F_{A,\mathbf{K}} \in \mathcal{C}^1(W^{1,p})$.

To finish the proof, we need to verify that A is continuous at $\tau = 0$ in the $W^{2,2}(Z)$-topology. We will show convergence in the $W^{1,2}(Z)$-topology; the higher order case is similar. Using the flow, we have

$$A(\tau_b) - A(\tau_a) = \int_{\tau_a}^{\tau_b} \partial_\tau A \, d\tau = -\int_{\tau_a}^{\tau_b} d^*_{A,\mathbf{K}} F_{A,\mathbf{K}} \, d\tau. \tag{16}$$

Take the $W^{1,2}$-norm of both sides to get

$$\begin{aligned} \|A(\tau_b) - A(\tau_a)\|_{W^{1,2}(Z)} &\leq \int_{\tau_a}^{\tau_b} \|d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|_{W^{1,2}(Z)} \, d\tau \\ &\leq C \int_{\tau_a}^{\tau_b} \Big(\|d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|_{L^2(Z)} \\ &\qquad + \|d^*_{A_{ref}} d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|_{L^2(Z)} \\ &\qquad + \|d_{A_{ref}} d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|_{L^2(Z)} \Big) d\tau. \end{aligned} \tag{17}$$

We want to show that the right-hand side goes to zero as τ_a, τ_b go to zero. For the first term, we have

$$\begin{aligned} \int_{\tau_a}^{\tau_b} \|d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|_{L^2(Z)} \, d\tau &\leq |\tau_a - \tau_b| \sup_{[\tau_a,\tau_b]} \|d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|_{L^2(Z)} \\ &= |\tau_a - \tau_b| \sup_{[\tau_a,\tau_b]} \|d^*_{B,\mathbf{K}} F_{B,\mathbf{K}}\|_{L^2(Z)}. \end{aligned}$$

where we used the fact that A and B are gauge equivalent. The continuity properties of B at $\tau = 0$ imply that the supremum here is bounded independent of $\tau_a, \tau_b > 0$ (assuming they are far from the maximal time τ_0). In particular, the right-hand side of the above goes to zero as τ_a, τ_b go to zero.

As for the second term in (17), to show

$$\lim_{\tau_a,\tau_b\searrow 0}\int_{\tau_a}^{\tau_b}\|d^*_{A_{ref}}d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\|_{L^2(Z)}\,d\tau=0,$$

use the second Bianchi identity, and argue as we did for the first term.

It remains to show that the third term in (17) goes to zero:

$$\lim_{\tau_a,\tau_b\searrow 0}\int_{\tau_a}^{\tau_b}\|d_{A_{ref}}d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\|_{L^2(Z)}\,d\tau=0.$$

For this, differentiate and use the flow equation to get

$$\begin{aligned}\frac{d}{d\tau}\frac{1}{2}\|d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\|^2_{L^2(Z)} \quad = \quad & -\|d_{A,\mathbf{K}}d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\|^2_{L^2(Z)}\\ & +\left(*\left[d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\wedge *F_{A,\mathbf{K}}\right],d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\right)\\ & +\left(*d^2\mathbf{K}_A(d^*_{A,\mathbf{K}}F_{A,\mathbf{K}},*F_{A,\mathbf{K}}),d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\right)\end{aligned} \tag{18}$$

Here $d^2\mathbf{K}_A$ is the second derivative of $\mathbf{K}$ at A. Note that the last two terms on the right are bounded by some constant C that is independent τ, provided τ is sufficiently small. Integrating (18) over $[\tau_a,\tau_b]$ then gives

$$\begin{aligned}\int_{\tau_a}^{\tau_b}\|d_{A_{ref}}d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\|_{L^2(Z)}\,d\tau \quad \leq \quad & \tfrac{1}{2}\|d^*_{A(\tau_a),\mathbf{K}}F_{A(\tau_a),\mathbf{K}}\|^2_{L^2(Z)}\\ & -\tfrac{1}{2}\|d^*_{A(\tau_b),\mathbf{K}}F_{A(\tau_b),\mathbf{K}}\|^2_{L^2(Z)}+|\tau_b-\tau_a|C\\ = \quad & \tfrac{1}{2}\|d^*_{B(\tau_a),\mathbf{K}}F_{B(\tau_a),\mathbf{K}}\|^2_{L^2(Z)}\\ & -\tfrac{1}{2}\|d^*_{B(\tau_b),\mathbf{K}}F_{B(\tau_b),\mathbf{K}}\|^2_{L^2(Z)}+|\tau_b-\tau_a|C\end{aligned}$$

for some constant C. The continuity of B at $\tau=0$ shows that this is going to zero when τ_a,τ_b approach 0. □

4 Energy concentration

Let $\mathbf{K}$ and $a\in\mathcal{A}(Q)$ be as in the introduction to Section 3. As in the closed case [28, 26], the maximal existence time for the flow is determined by concentration of energy. When this maximal existence time is finite, Yang–Mills bubbles form at isolated points. These claims are made precise in Propositions 4.1 and 4.3, below. We will repeatedly use fact that the $L^2(Z)$-norms of $F_{A,\mathbf{K}}$ and $F^+_{A,\mathbf{K}}$ are non-increasing along the flow. Indeed, the relation (7), the flow (10), and the first Bianchi identity give

$$\begin{aligned}\tfrac{d}{d\tau}\|F^+_{A(\tau),\mathbf{K}}\|^2_{L^2(Z)} &= \tfrac{d}{d\tau}\tfrac{1}{2}\|F_{A(\tau),\mathbf{K}}\|^2_{L^2(Z)}\\ &= -\|d^*_{A,\mathbf{K}}F_{A,\mathbf{K}}\|^2_{L^2(Z)} = -4\|d^*_{A,\mathbf{K}}F^+_{A,\mathbf{K}}\|^2_{L^2(Z)}.\end{aligned} \tag{19}$$

Throughout, we denote by $B_R(z)\subset Z$ the R-ball centered at $z\in Z$.

Proposition 4.1 (Energy concentration)**.** *There is some constant $\eta > 0$ so that the following holds. Let A be a solution of the flow (10), satisfying the conclusions of Theorem 3.1. Then the maximal existence time for A is characterized by*

$$\overline{\tau} := \sup \left\{ \tau_0 > 0 \;\middle|\; \exists R > 0, \sup_{z \in Z,\, 0 \leq \tau \leq \tau_0} \int_{B_R(z)} |F_{A(\tau)}|^2 < \eta \right\}. \tag{20}$$

If $\overline{\tau} < \infty$ is finite, then at $\tau = \overline{\tau}$, the curvature concentrates at a finite number of points

$$\{z_1, \ldots, z_K\} \subset Z$$

in the sense that

$$\forall 1 \leq k \leq K, \forall R > 0, \qquad \limsup_{\tau \nearrow \overline{\tau}} \int_{B_R(z_k)} |F_{A(\tau)}|^2 \geq \eta. \tag{21}$$

We will refer to the z_k as the *bubbling points*. The proof of this proposition is given below.

Remark 4.2. *(a) Let η_{S^4} be the infimum of $\|F_A\|^2_{L^2(S^4)}$ over all non-flat Yang–Mills connections A on principal G-bundles over S^4 (it is well-known that this infimum is positive, however see also Theorem 5.7). The proof of Proposition 4.1 will show that we can take $\eta = \eta_{S^4}$ in the statement of the proposition. Note that this constant η_{S^4} depends only on the Lie group G and the choice in (1) of* Ad*-invariant inner product on $\mathfrak{g}$.*

(b) The curvature $F_{A(\tau)}$ appearing in (20) and (21) can be replaced by the perturbed curvature $F_{A(\tau),\mathbf{K}}$ to yield the same result with the same constants. This is because Axiom 2 implies the norm $\|\mathbf{K}_A\|_{L^\infty(Z)}$ is bounded independent of A, and so $\int_{B_R(z)} |\mathbf{K}_A|^2 \leq CR^4$ for some uniform constant C.

(c) It is conceivable, a priori, that energy concentration can occur along points that escape off the end of Z. However, part of the claim of Proposition 4.1 is that, given an initial condition A_0, all of the (finite-time) bubbling points are confined to a compact subset of Z. We thank an anonymous referee for pointing this out.

By rescaling around each bubbling point, one can show that a Yang–Mills bubble on S^4 forms as τ approaches the maximal flow time $\overline{\tau}$.

Proposition 4.3 (Bubble formation)**.** *At each bubbling point, a non-flat Yang–Mills connection on a bundle over S^4 separates, in the following sense:*

In the notation of Proposition 4.1, fix $1 \leq k \leq K$, as well as sequences $\tau_n \nearrow \overline{\tau}$, and $R_n \searrow 0$ indexed by $n \geq 0$. Let d be the trivial connection on $B_{R_0}(z_k)$ relative to some fixed trivialization of the bundle over $B_{R_0}(z_k)$. Write

$$A(\tau)|_{B_{R_0}(z_k)} = d + M(\tau)$$

for some τ-dependent $\mathfrak{g}$-valued 1-form M. Define a connection on $B_{1/R_n}(0) \subset \mathbb{R}^4$ by

$$A_n(x) := d + R_n M(\tau_n; z_k + R_n x). \tag{22}$$

Then the A_n converge, modulo gauge and in $W^{1,p}_{loc}(\mathbb{R}^4)$, to a non-flat Yang–Mills connection on $\mathbb{R}^4$ with finite energy. This Yang–Mills connection extends to a non-flat Yang–Mills connection on some bundle over S^4.

In the statement of the above proposition, we are implicitly assuming that $R_0 > 0$ is small enough so that each ball $B_{R_0}(z_k) \subset Z$ is contractible.

Proof of Propositions 4.1 and 4.3. Suppose $A(\tau)$ is a solution of (10) on $[0, \tau_0)$ with the regularity of Theorem 3.1. Let $\delta > 0$ be as in the statement of Lemma 2.11. As argued by Struwe in [28, Lemma 3.6], it follows from Lemma 2.11 that if there is some $R > 0$ with

$$\sup_{z \in Z,\, 0 \le \tau < \tau_0} \int_{B_R(z)} |F_{A(\tau)}|^2 < \delta,$$

then $A(\tau)$ can be continuously extended to $\tau = \tau_0$ (hence extended for a positive time past τ_0, as well). In particular, the quantity $\overline{\tau}$ from the statement of Proposition 4.1, does indeed characterize the maximal existence time (with δ temporarily in place of η).

We will now justify the claim of Remark 4.2 (a), by showing that if

$$\exists \eta > 0, \forall R > 0, \qquad \sup_{z \in Z,\, 0 \le \tau < \overline{\tau}} \int_{B_R(z)} |F_{A(\tau)}|^2 \ge \eta, \tag{23}$$

then

$$\forall R > 0, \quad \sup_{z \in Z,\, 0 \le \tau < \overline{\tau}} \int_{B_R(z)} |F_{A(\tau)}|^2 \ge \eta_{S^4}, \tag{24}$$

with η_{S^4} as defined in Remark 4.2 (a). Our argument here holds for any $\overline{\tau} \in (0, \infty]$. Assuming (23), fix sequences $\tau_n \nearrow \overline{\tau}$ and $R_n \searrow 0$, and find points $x_n \in Z$ with

$$\int_{B_{R_n}(x_n)} |F_{A(\tau_n)}|^2 \ge \sup_{z \in Z} \frac{1}{2} \int_{B_{R_n}(z)} |F_{A(\tau_n)}|^2.$$

At this stage, it may be the case that the x_n are unbounded; however, see the claim below. Then

$$\limsup_n \int_{B_{R_n}(x_n)} |F_{A(\tau_n)}|^2 \ge \frac{1}{2}\eta > 0 \tag{25}$$

is positive. Define connections A'_n on $B_{1/R_n}(0) \subset \mathbb{R}^4$ as in the right-hand side of (22), but with x_n in place of z_k. These rescaled connections are Yang–Mills on $B_{1/R_n}(0)$ relative to a metric that $\mathcal{C}^\infty$-converges to the standard metric, and relative to a perturbation that $\mathcal{C}^\infty$-converges to 0. It therefore follows from the argument given by Schlatter [26] that a subsequence of the A'_n converge, in the sense described in Proposition 4.3, to produce a Yang–Mills connection A'_∞ on some bundle over S^4. Then (25) implies that A'_∞ is not flat. It follows from the definition of η_{S^4} that

$$\limsup_n \int_{B_{R_n}(x_n)} |F_{A(\tau_n)}|^2 \ge \eta_{S^4}, \tag{26}$$

from which (24) follows.

To finish the proof, we need to identify the bubbling points, and argue that there are finitely many. We begin with the following.

Claim. *If $\overline{\tau} < \infty$, then the x_n are confined to a compact subset of Z.*

We will prove the claim shortly. Assuming it for now, we can pass to a subsequence and assume the x_n converge to some $z_1 \in Z$. Clearly the connections defined in (22) also produce a non-flat Yang–Mills connection in the limit, so this proves Proposition 4.3. Now repeating the above argument for all sequences x_n where energy concentrates, we obtain points $z_1, z_2, \ldots \in Z$ where bubbles form; we may assume the z_k are distinct. Since each bubble requires an energy of at least η_{S^4}, and the total energy of $A(\tau)$ is bounded by the energy of A_0,

$$\mathcal{YM}_{\mathbf{K}}(A(\tau)) \leq \mathcal{YM}_{\mathbf{K}}(A_0),$$

there can be only a finite number $\{z_1, \ldots, z_K\}$ of such bubbling points. This finishes the proof of Proposition 4.1, provided we can establish the claim.

To prove the claim, we will show that there is some $s_0 \geq 0$ with

$$\sup_{0 \leq \tau < \overline{\tau}} \int_{[s_0, \infty) \times Y} |F_{A(\tau), \mathbf{K}}|^2 \leq \eta_{S^4}/2. \tag{27}$$

The claim will then follow immediately from (27), (26), and Remark 4.2 (b). We therefore aim to prove (27). Fix a smooth bump function $\varphi : \mathbb{R} \to \mathbb{R}$ that vanishes for $s \leq 0$ and is identically 1 for $s \geq 1$. This can be chosen so the derivative is bounded:

$$\sup_s |\varphi'(s)|^2 \leq 3/2.$$

Since the initial condition A_0 has finite energy, we can find $\sigma \geq 0$ with

$$\int_{[\sigma, \infty) \times Y} |F_{A_0, \mathbf{K}}|^2 < \eta_{S^4}/4.$$

Let $\phi_\sigma : Z \to \mathbb{R}$ be given by $\phi_\sigma(s, y) = \varphi(s/\sigma - 1)$ on $[0, \infty) \times Y$, and extended by zero to the rest of Z. Then relative to any cylindrical metric, this satisfies $\sup_z |\nabla \phi_\sigma|^2 \leq 3/(2\sigma)$. If we are working with an asymptotically cylindrical metric, then by increasing σ, we may assume this satisfies

$$\sup_z |\nabla \phi_\sigma(z)|^2 \leq 2/\sigma.$$

Our argument now is similar to [28, p.133]. For $A = A(\tau)$, use the flow to write

$$\begin{aligned}
\int_Z \phi_\sigma^2 |d_{A,\mathbf{K}}^* F_{A,\mathbf{K}}|^2 + \frac{d}{d\tau} \frac{1}{2} \int_Z \phi_\sigma^2 |F_{A,\mathbf{K}}|^2 &= 2((\nabla \phi_\sigma) \wedge d_{A,\mathbf{K}}^* F_{A,\mathbf{K}}, \phi_\sigma F_{A,\mathbf{K}}) \\
&\leq \int_Z \phi_\sigma^2 |d_{A,\mathbf{K}}^* F_{A,\mathbf{K}}|^2 + \int_Z |\nabla \phi_\sigma|^2 |F_{A,\mathbf{K}}|^2 \\
&\leq \int_Z \phi_\sigma^2 |d_{A,\mathbf{K}}^* F_{A,\mathbf{K}}|^2 + \frac{1}{\sigma} \mathcal{YM}_{\mathbf{K}}(A(\tau)).
\end{aligned}$$

The Yang–Mills functional is non-increasing along the flow, so this gives

$$\frac{d}{d\tau}\frac{1}{2}\int_Z \phi_\sigma^2 |F_{A,\mathbf{K}}|^2 \leq \frac{1}{\sigma}\mathcal{YM}_{\mathbf{K}}(A_0).$$

Integrate over the interval $[0, \tau]$ to obtain

$$\begin{aligned}\int_{[2\sigma,\infty)\times Y} |F_{A(\tau),\mathbf{K}}|^2 &\leq \int_{[\sigma,\infty)\times Y} |F_{A_0,\mathbf{K}}|^2 + \frac{2\tau}{\sigma}\mathcal{YM}_{\mathbf{K}}(A_0) \\ &\leq \frac{\eta_{S^4}}{4} + \frac{2\overline{\tau}}{\sigma}\mathcal{YM}_{\mathbf{K}}(A_0).\end{aligned}$$

Then (27) follows by taking $\sigma \geq 8\overline{\tau}\mathcal{YM}_{\mathbf{K}}(A_0)/\eta_{S^4}$. □

Even in the presence of bubbling, the connections $A(\tau)$ converge in a rather weak sense on the complement of the bubbling set on Z. When bubbles form in finite time, we obtain a statement familiar from the closed setting [26, Theorem 1.3] in the sense that we have L^2-convergence on the full 4-manifold.

Proposition 4.4 (Convergence with finite-time bubbling). *Let A be as in the statement of Proposition 4.1, and assume the maximal existence time $\overline{\tau} < \infty$ is finite. Then there is a finite-energy connection*

$$A_1 \in \mathcal{A}^{0,2}\Big(P; a\Big) \cap \mathcal{A}^{1,2}_{loc}\Big(P|_{Z\setminus\{z_1,\ldots,z_K\}}\Big)$$

with the property that, as τ increases to $\overline{\tau}$, the connections $A(\tau)$ converge to A_1 in $L^2(Z) \cap W^{1,2}_{loc}(Z\setminus\{z_1,\ldots,z_K\})$. Moreover,

$$\mathcal{YM}_{\mathbf{K}}(A_1) + \eta_{S^4}\sum_{k=1}^{K} n_k \leq \liminf_{\tau\nearrow\overline{\tau}} \mathcal{YM}_{\mathbf{K}}(A(\tau)), \tag{28}$$

for some positive integers n_k, where η_{S^4} is as in Remark 4.2 (a).

Remark 4.5. *It follows from Uhlenbeck's theorem on removal of singularities [30, Theorem 2.1] that the limiting connection A_1 from Proposition 4.4 extends over the bubbling points z_k by possibly modifying the underlying bundle. The finite-energy and $L^2(Z)$-convergence then imply that A_1 is gauge equivalent to a connection in $\mathcal{A}^{1,2}(P_1; a)$ for some principal G-bundle $P_1 \to Z$. More precisely, the bundle P_1 is cylindrical on the end in the sense that*

$$P_1|_{[s_1,\infty)\times Y} \cong [s_1,\infty) \times Q$$

and there is a bundle isomorphism

$$U_1 : P_1|_{Z\setminus\{z_1,\ldots,z_K\}} \xrightarrow{\cong} P|_{Z\setminus\{z_1,\ldots,z_K\}},$$

so that $U_1^* A_1$ extends to a $W^{1,2}$ connection defined on all of Z. The bundle Q appearing here is the same one associated with the ends of P, and $s_1 \geq 0$ is large enough so *that $[s_1, \infty) \times Y$ does not contain any of the bubbling points z_k. It will essentially follow from the proof of Proposition 4.4 that U_1 converges down the cylindrical ends to the identity map. In particular, $U^* A_1$ is asymptotic to a, and so is an element of $\mathcal{A}^{1,2}(P_1; a)$.*

Intuitively, the bundle isomorphism U_1 reflects the bubbles associated to the z_k. That U_1 is asymptotic to the identity captures the fact that there are no bubbling points that escape down the ends.

Proof of Proposition 4.4. First we will prove $L^2(Z)$-convergence of $A(\tau)$ as τ increases to $\overline{\tau} < \infty$. Integrating (19) over some interval $[\tau_a, \tau_b]$ gives

$$\int_{\tau_a}^{\tau_b} \|d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|^2_{L^2(Z)} = \frac{1}{2}\|F^+_{A(\tau_a),\mathbf{K}}\|^2_{L^2(Z)} - \frac{1}{2}\|F^+_{A(\tau_a),\mathbf{K}}\|^2_{L^2(Z)}. \qquad (29)$$

Next, recall the identity (16). Take the $L^2(Z)$-norm of both sides of (16), and then using Hölder's inequality in the time-variable to get

$$\begin{aligned} \|A(\tau_b) - A(\tau_a)\|_{L^2(Z)} &\leq \int_{\tau_a}^{\tau_b} \|d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|_{L^2(Z)}\, d\tau \\ &\leq |\tau_b - \tau_a|^{1/2} \Big(\int_{\tau_a}^{\tau_b} \|d^*_{A,\mathbf{K}} F_{A,\mathbf{K}}\|^2_{L^2(Z)} \Big)^{1/2}. \end{aligned}$$

Combining this with (29) gives

$$\|A(\tau_b) - A(\tau_a)\|^2_{L^2(Z)} \leq |\tau_b - \tau_a| \sup_{[\tau_a, \tau_b]} \|F^+_{A,\mathbf{K}}\|^2_{L^2(Z)}.$$

By (19), the quantity $\|F^+_{A,\mathbf{K}}\|^2_{L^2(Z)}$ is non-increasing along the flow, so we have

$$\|A(\tau_b) - A(\tau_a)\|^2_{L^2(Z)} \leq |\tau_b - \tau_a| \|F^+_{A_0,\mathbf{K}}\|^2_{L^2(Z)}.$$

This implies that $A(\tau)$ is $L^2(Z)$-Cauchy as $\tau \nearrow \overline{\tau}$. In particular, the $A(\tau)$ converge in $L^2(Z)$ to some

$$A_1 \in \mathcal{A}^{0,2}(P; a).$$

The $W^{1,2}_{loc}$-convergence to A_1 on $Z \backslash \{z_1, \ldots, z_K\}$ can now be shown using Schlatter's argument for the proof of [26, Theorem 1.2 (i)], which is local in nature and hence not sensitive to the cylindrical ends.

Finally, we establish the energy inequality (28). For this, let τ_n, R_n be as in the statement of Proposition 4.3; we may assume the τ_n are increasing, and the R_n are small. Consider the complement

$$Z_n := Z \backslash \bigcup_k B_{R_n}(z_k),$$

of the R_n-balls around the bubbling points. Then

$$\mathcal{YM}_{\mathbf{K}}(A(\tau_n)) = \frac{1}{2}\int_{Z_n} |F_{A(\tau_n),\mathbf{K}}|^2 + \frac{1}{2}\sum_k \int_{B_{R_n}(z_k)} |F_{A(\tau_n),\mathbf{K}}|^2.$$

The energy is non-increasing along the flow, so the sequence $\mathcal{YM}_{\mathbf{K}}(A(\tau_n))$ converges to the liminf of $\mathcal{YM}_{\mathbf{K}}(A(\tau))$. Hence

$$\limsup_{n\to\infty} \frac{1}{2}\int_{Z_n} |F_{A(\tau_n),\mathbf{K}}|^2 + \eta_{S^4}\sum_k n_k \leq \liminf_{\tau \nearrow \overline{\tau}} \mathcal{YM}_{\mathbf{K}}(A(\tau)).$$

On the other hand, for each compact set $S \subset Z$, the $W^{1,2}_{loc}$-convergence of the $A(\tau)$ gives

$$\frac{1}{2}\int_S |F_{A_1,\mathbf{K}}|^2 \leq \limsup_{n\to\infty} \frac{1}{2}\int_{Z_n} |F_{A(\tau_n),\mathbf{K}}|^2.$$

Since this bound is plainly independent of the compact set S, we obtain (28). □

5 Long-time existence

In this section we prove the long-time existence assertions of Theorem 1.2.

Theorem 5.1 (Long-time existence)**.** *Assume* $\mathbf{K}$ *is an ASD-regular perturbation satisfying Axioms 1, 2, 3, and 4. Then there is an integer* $\mathcal{I}_G > 0$ *so that if* a *is any* K*-flat connection on* Q *with* $\mathrm{Ind}_{K,P}(a) < \mathcal{I}_G$*, then there is a positive constant* $\eta(a)$ *so the following holds.*

Fix $p > 4$*, as well as* $A_0 \in \mathcal{A}^{1,2}(P;a) \cap \mathcal{A}^{2,p}(P;a)$*, and assume this satisfies*

$$\|F^+_{A_0}\|^2_{L^2(Z)} < \eta(a).$$

Then there is a unique solution $A : [0,\infty) \to \mathcal{A}^{1,2}(P;a)$ *to the flow (10), and this has the regularity asserted in Theorem 3.1. Moreover, there is no energy quantization at finite or infinite time, in the sense that*

$$\lim_{\delta\to 0^+} \sup_{\tau \geq 0, z\in Z} \int_{B_\delta(z)} |F_{A(\tau)}|^2 = 0.$$

Our proof is given in Section 5.4, and is motivated by the work of Schlatter [26]. The index $\mathrm{Ind}_{K,P}$ is the index of the $\mathbf{K}$-ASD operator, and is defined in (8). We define the constant $\mathcal{I}_G$ in Section 5.1; it depends only on G. Section 5.2 defines a notion of convergence that naturally arises when one considers gauge theory on cylindrical ends; this notion will be relevant for our proofs in the sections that follow. In Section 5.3, we give a fairly concrete definition of the constant $\eta(a)$ from the theorem. In addition to depending on a, the constant $\eta(a)$ also depends on the metric g as well as the perturbation $\mathbf{K}$.

For future reference, we also establish the following regularity estimate, which is proved in Section 5.4.

Corollary 5.2. *In the setting of Theorem 5.1, there are* $C, \tau_0 > 0$ *so the following holds. Suppose* $2 \leq q \leq 4$ *and* $\tau \geq \tau_0$*. Then*

$$\|W\|_{L^q(Z)} \leq C\|d^*_{A(\tau),\mathbf{K}} W\|_{L^2(Z)}$$

for all self-dual 2-forms $W \in W^{1,2}(\Omega^+)$, and

$$\|V\|_{L^q(Z)} \leq C\|d^+_{A(\tau),\mathbf{K}} V\|_{L^2(Z)}$$

for all 1-forms V in the image of $d^*_{A(\tau),\mathbf{K}} : W^{2,2}(\Omega^+) \to W^{1,2}(\Omega^1)$.

5.1 Characteristic classes and the constant $\mathcal{I}_G$

In this section, we define the constant $\mathcal{I}_G$ that appears in Theorem 5.1. Its definition relies on a characteristic class, and a certain rational number associated to the index. We begin by describing the former of these.

Fix a principal G-bundle R over a closed, connected, oriented 4-manifold X. Then the relevant characteristic class is

$$\kappa(R) := \frac{1}{2}\int_X \langle F_A \wedge F_A \rangle,$$

which depends only on the topological type of R. The following examples relate this to standard characteristic classes (recall from (1) that the inner product $\langle \cdot, \cdot \rangle$ is induced from an immersion $G \to \mathrm{U}(N)$).

Example 5.3. *(a) Suppose $G = \mathrm{SU}(N)$ for $N \geq 2$, and the immersion $G \to \mathrm{U}(N)$ from above is the inclusion. Then the Chern-Weil formula gives*

$$\kappa(R) = 2c_2(R)\,[X] \in 2\mathbb{Z}.$$

(b) Suppose $G = \mathrm{U}(N)$ for $N \geq 2$, and the immersion $G \to \mathrm{U}(N)$ is just the identity. Then

$$\kappa(R) = 2\left(c_2(R) - \frac{1}{2}c_1^2(R)\right)[X] \in \mathbb{Z}.$$

(c) Suppose $G = \mathrm{SO}(r)$ for $r \geq 2$, and the immersion $G \to \mathrm{SU}(N) \subset \mathrm{U}(N)$ is given by the complexified adjoint action of G on $\mathfrak{g}^{\mathbb{C}}$. Then the induced inner product on $\mathfrak{g}$ is $-(2\pi^2)^{-1}$ times the Killing form, and

$$\kappa(R) = -2(r-2)p_1(R)\,[X] \in 2(r-2)\mathbb{Z}$$

where p_1 is the first Pontryagin class. Note that this vanishes for $r = 2$, reflecting the fact that $\mathrm{SO}(2)$ is abelian.

(d) Suppose $G = \mathrm{PU}(r)$ for $r \geq 2$, and the embedding $G \hookrightarrow \mathrm{SU}(N) \subset \mathrm{U}(N)$ is given by the complexified adjoint action, as in (c). Then

$$\kappa(R) = 2q_4(R)\,[X] \in 2\mathbb{Z},$$

where $q_4(R) \in H^4(X, \mathbb{Z})$ is a $\mathrm{PU}(r)$-generalization of the first Pontryagin class; see [38].

More generally, we have the following.

Lemma 5.4. *Fix G and $\langle \cdot, \cdot \rangle$ as above, and let X be a closed, connected, oriented 4-manifold X. Then $\kappa(R)$ is an integer for every principal G-bundle $R \to X$. If G is not abelian, then there are principal G-bundles R for which $\kappa(R)$ is non-zero.*

Proof. Given R, consider the classifying map $\psi_R : X \to BG \subset B\mathrm{U}(N)$, where we have used the embedding $G \hookrightarrow \mathrm{U}(N)$ to identify BG as a sub-complex of $B\mathrm{U}(N)$. Then by Example 5.3 (b), the class

$$\kappa(R) = \psi_R^*(2c_2 - c_1^2)\,[X]$$

is the pullback of $2c_2 - c_1^2$. This shows that $\kappa(R)$ is an integer.

Now assume G is not abelian. Then there is a Lie group homomorphism $\phi : \mathrm{SU}(2) \to G$ with finite kernel $|\ker \phi| < \infty$. Fix any $\mathrm{SU}(2)$-bundle $R' \to X$ with $c_2(R') \neq 0$, and define

$$R := R' \times_{\mathrm{SU}(2)} G,$$

where $\mathrm{SU}(2)$ acts on G by the homomorphism ϕ. Then $\kappa(R) = 2|\ker \phi| c_2(R')$, which is non-zero. See [13, Appendix E] for a similar discussion in the case where $G = \mathrm{SO}(3)$. $\square$

Let $Q \to Y$ be a principal G-bundle on a 3-manifold, and suppose u is a gauge transformation on Q. The mapping torus of u is a bundle Q_u over the 4-manifold $S^1 \times Y$. It is not hard to show that, if $a \in \mathcal{A}(Q)$ is any connection, then

$$\mathcal{CS}_{K,P}(u^*a) - \mathcal{CS}_{K,P}(a) = \kappa(Q_u) \in \mathbb{Z}. \tag{30}$$

Now suppose that a is a K-flat connection. By the argument of [6, Prop. 3.16], we have the following action-index identity

$$n_G\left(\mathcal{CS}_{K,P}(u^*a) - \mathcal{CS}_{K,P}(a)\right) = \mathrm{Ind}_{K,P}(u^*a) - \mathrm{Ind}_{K,P}(a). \tag{31}$$

Here $n_G \geq 0$ is a number depending only on the Lie group G, and the choice of inner product on $\mathfrak{g}$. When G is not abelian, then n_G is uniquely determined by (31) since there are a, u for which both sides are non-zero. In particular, the number n_G is positive; it is also rational by (30) and the fact that the index takes integer values. When G is abelian, both sides of (31) are zero for all a, u; this reflects the triviality of the group $\pi_3(G) = 0$. In the abelian case, we are therefore free to declare $n_G := 1$.

Example 5.5. *(a) Suppose $G = \mathrm{SU}(N)$ and the embedding $G \hookrightarrow \mathrm{U}(N)$ is the identity. Then $n_G = 2r$. See [6, Prop. 3.16].*

(b) Suppose $G = \mathrm{PU}(r)$ and the embedding $G \hookrightarrow \mathrm{SU}(N) \subset \mathrm{U}(N)$ is given by the complexified adjoint action. Then $n_G = 1$.

With these preliminaries, we define the extended real number

$$\mathcal{I}_G := \inf_{R \to S^4} n_G\,|\kappa(R)|,$$

where the infimum is over all principal G-bundles $R \to S^4$ for which $\kappa(R) \neq 0$. When G is abelian, we have $\mathcal{I}_G = \infty$ since $\kappa(R) = 0$ for all R. When G is not abelian, it follows from Lemma 5.4 that $\mathcal{I}_G \in \mathbb{Z}$ is a (finite) positive integer. The significance of $\mathcal{I}_G$ is that bubbling cannot occur in any $\mathbf{K}$-ASD moduli space of dimension smaller than $\mathcal{I}_G$; see Section 2.2 for more details on this moduli space. See also [1, Section 8] for a similar discussion.

5.2 Broken trajectories

Fix a K-flat connection a on Q. Here we define 'broken trajectories', which are objects that often appear as limits of sequences in $\mathcal{A}(P; a)$.

Definition. *A* broken trajectory on Z asymptotic to $a \in \mathcal{A}(Q)$ *consists of the following:*

- *a finite number of K-flat connections $a^0, a^1, \ldots, a^J \in \mathcal{A}(Q)$ with $a^J = a$;*
- *for each $0 \leq j \leq J$, a gauge transformation u_j on Q;*
- *a connection $A^0 \in \mathcal{A}^{1,2}(P; u_0^* a^0)$ asymptotic to $u_0^* a^0$;*
- *a finite number of connections $A^1, \ldots, A^J \in \mathcal{A}^{1,2}(\mathbb{R} \times Q)$ satisfying*

$$\lim_{s \to -\infty} A^j|_{\{s\} \times Y} = a^{j-1}, \qquad \lim_{s \to +\infty} A^j|_{\{s\} \times Y} = u_j^* a^j.$$

A broken trajectory is a broken $\mathbf{K}$-YM trajectory *(resp.* broken $\mathbf{K}$-ASD trajectory*) if A^0 is $\mathbf{K}$-YM (resp. $\mathbf{K}$-ASD) and for $1 \leq j \leq J$, the A^j is $\mathbf{K}^Y$-YM (resp. $\mathbf{K}^Y$-ASD).*

Here the perturbation $\mathbf{K}^Y$ is as in Remark 2.2. We will typically denote a broken trajectory by $(A^0; A^1, \ldots, A^J)$, with the asymptotic K-flat connections a^j and gauge transformations u_j suppressed.

Remark 5.6. *A more refined notion would be to require that the gauge transformations u_j have degree zero (i.e., to require that $\kappa(Q_{u_j}) = 0$), since otherwise the connections $u_j^* a^j$ and a^j would be very far apart from an energy perspective. However, we will have no need to specify such a criterion explicitly, as it is effectively implicit in the estimates we establish (e.g., see Case 1.1 in the proof of Theorem 5.7).*

Now we define the relevant notion of convergence. Let $\mathcal{H}$ be a function space (e.g., $\mathcal{C}^\infty$ or $W^{1,p}$).

Definition. *A sequence $(A_n)_{n \in \mathbb{N}}$ of connections* converges in $\mathcal{H}$, modulo gauge, *to a broken trajectory $(A^0; A^1, \ldots, A^J)$ if the following holds.*

- *There is a sequence $(U_{0,n})_{n \in \mathbb{N}}$ of gauge transformations on Z and, for each $1 \leq j \leq J$, a sequence $(U_{j,n})_{n \in \mathbb{N}}$ of gauge transformations on $\mathbb{R} \times Y$.*
- *For each $1 \leq j \leq J$, there is a sequence $(s_{j,n})_{n \in \mathbb{N}}$ of positive real numbers that increases to ∞.*

These are required to satisfy the following.

- *For each* $1 \leq j \leq J-1$, $\lim_{n\to\infty} s_{j+1,n} - s_{j,n} = \infty$.
- *For each compact subset* $B \subset Z$*, the sequence* $U_{0,n}^* A_n$ *converges in* $\mathcal{H}(B)$ *to* A^0.
- *Fix* $1 \leq j \leq J$*, and let*

$$\tau_{s_{j,n}}^* U_{j,n}^* A_n$$

  denote the connection on $[-s_{j,n}, \infty) \times Y$ *obtained by translating* $U_{j,n}^* A_n|_{[0,\infty)\times Y}$. *Then for each compact set* $B \subset \mathbb{R} \times Y$*, the sequence* $\tau_{s_{j,n}}^* U_{j,n}^* A_n$ *converges in* $\mathcal{H}(B)$ *to* A^j.

The sequence converges modulo bubbling *if the convergence to* A^0 *holds on the complement of a finite set of points on* Z*, and, for* $1 \leq j \leq J$*, the convergence to each* A^j *is on the complement of a finite set of points in* $\mathbb{R} \times Y$ *(this set is allowed to depend on* j*).*

For more details, see [20, Chapter 6] or [12]; see also [22] for a nice treatment of the closely related case of holomorphic curves.

5.3 A positive energy gap

In this section, we give a fairly concrete description of the constant $\eta(a)$ appearing in the statement of Theorem 5.1. Specifically, let n_G be as in (31). Then $\eta(a)$ can be taken to be the minimum of the numbers $1, 1/n_G$ and the following three numbers:

A. The infimum

$$\inf_A \|F_A^+\|^2_{L^2(S^4)}.$$

Here the infimum ranges over all connections A on (any principal G-bundle over) S^4 that are Yang–Mills, not ASD, and satisfy $\mathcal{YM}(A) \leq \mathcal{CS}_{K,P}(a) + 1$.

B. The infimum

$$\inf_A \|F_{A,\mathbf{K}}^+\|^2_{L^2(Z)}.$$

Here the infimum ranges over all connections A on P that are $\mathbf{K}$-YM, not $\mathbf{K}$-ASD, and satisfy $\mathcal{YM}_{\mathbf{K}}(A) \leq \mathcal{CS}_{K,P}(a) + 1$.

C. The infimum

$$\inf_A \|F_{A,\mathbf{K}^Y}^+\|^2_{L^2(\mathbb{R}\times Y)}.$$

Here the infimum ranges over all connections A on $\mathbb{R} \times Q$ that are $\mathbf{K}^Y$-YM, not $\mathbf{K}^Y$-ASD, and satisfy $\mathcal{YM}_{\mathbf{K}^Y}(A) \leq \mathcal{CS}_{K,P}(a) + 1$.

For the quantity in C, the perturbation $\mathbf{K}^Y$ is as in Remark 2.2, and the metric on $\mathbb{R} \times Y$ is $ds^2 + g^Y$.

The next theorem justifies this choice of $\eta(a)$ by stating that each of the infima in A-C is positive. In light of the identity (7), this positivity can be viewed as saying that the perturbed Yang–Mills functional has a positive energy gap above the minimum energy level given by the **K**-ASD connections. (Of course, this minimum energy level is only a theoretical lower bound, since **K**-ASD connections may not exist.)

Theorem 5.7. *The quantities in A, B, and C are positive.*

Before proving this theorem, we establish some preliminary estimates. It follows from our regularity assumptions that if A (resp. A') is a **K**-ASD connection on Z (resp. $\mathbf{K}^Y$-ASD connection on $\mathbb{R} \times Y$), then there are constants $C_0(A), C_0(A')$ so that

$$\begin{array}{rcl} \|W\|_{L^q(Z)} & \leq & C_0(A)\|d_{A,\mathbf{K}}W\|_{L^2(Z)}, \\ \|W'\|_{L^q(\mathbb{R}\times Y)} & \leq & C_0(A')\|d_{A',\mathbf{K}^Y}W'\|_{L^2(\mathbb{R}\times Y)}, \end{array} \tag{32}$$

for all $2 \leq q \leq 4$ and all smooth, compactly supported self-dual 2-forms W (resp. W') on Z (resp. $\mathbb{R} \times Y$). There are similar estimates associated to K-flat connections. The next lemma extends these to connections with small slice-wise curvature on the end.

Lemma 5.8. *Assume all K-flat connections on Y are non-degenerate. There are constants $T_0, C_0, \delta_0 > 0$ so that the following holds. Let $I \subset [T_0, \infty)$ be an interval (possibly unbounded), and suppose A is a connection on $I \times Y \subset Z$. If*

$$\sup_{s\in I} \int_{\{s\}\times Y} |F_{A,\mathbf{K}}|^2 < \delta_0,$$

then

$$\|W\|_{L^q(I\times Y)} \leq C_0 \|d_{A,\mathbf{K}}W\|_{L^2(I\times Y)}$$

for all $2 \leq q \leq 4$ and all smooth self-dual 2-forms W compactly supported in $I \times Y$.

Proof. We will show that the lemma holds with $T_0 = 0$ in the case where $g = g_{cyl}$ is cylindrical; the general asymptotically cylindrical case follows by choosing $T_0 \geq 0$ large enough so that $g - g_{cyl}$ is $\mathcal{C}^1$-small on $[T_0, \infty) \times Y$.

Since all K-flat connections are non-degenerate, it follows that there is a constant C_1 so that

$$\|w\|_{L^2(Y)} + \|w\|_{L^4(Y)} \leq C_1(\|d_{a,K}w\|_{L^2(Y)} + \|d^*_{a,K}w\|_{L^2(Y)})$$

for all 2-forms $w \in \Omega^2(Y, \mathfrak{g}_Q)$ and all K-flat connections a. This constant C_1 can be chosen to be independent of a because this estimate is gauge invariant and the moduli space of K-flat connections is finite. Using this estimate and

Uhlenbeck compactness, it is not hard to show that there are $C_2, \delta > 0$ so that if a is a connection with $\|F_{a,K}\|_{L^2(Y)} < \delta$, then

$$\|w\|^2_{L^2(Y)} + \|w\|^2_{L^4(Y)} \leq C_2(\|d_{a,K}w\|^2_{L^2(Y)} + \|d^*_{a,K}w\|^2_{L^2(Y)}) \tag{33}$$

for all 2-forms w. Take δ_0 from the lemma to be no greater than this δ and no greater than $1/4C_2$.

Let A, W be as in the statement of the lemma, and write $A = a + pds$ and $W = w + ds \wedge v$ in terms of their components on the end $[0, \infty) \times Y$. The self-duality condition on W implies $v = *_Y w$, and the curvature of A can be expressed as

$$F_{A,\mathbf{K}} = F_{a,K} + ds \wedge (\partial_s a - d_{a,K} p).$$

The small curvature hypothesis on A implies that, for each $s \in I$, the estimate (33) holds with $a = a(s)$ and $w = w(s)$. Integrating this estimate over I gives

$$\|W\|^2 = 2\|w\|^2 \leq 2C_2(\|d_{a,K}w\|^2 + \|d^*_{a,K}w\|^2) \tag{34}$$

where, here and below, all unspecified norms and inner products are L^2 on $I \times Y$. To estimate the right-hand side, note that we have

$$d_{A,\mathbf{K}}W = d_{a,K}w + ds \wedge (\nabla_s w - d_{a,K} *_Y w),$$

where $\nabla_s = \partial_s + p$. Taking the L^2-norm of both sides gives

$$\begin{aligned}
\|d_{A,\mathbf{K}}W\|^2 &= \|d_{a,K}w\|^2 + \|d^*_{a,K}w\|^2 + \|\nabla_s w\|^2 \\
&\quad -2(\nabla_s w, d_{a,K} *_Y w) \\
&= \|d_{a,K}w\|^2 + \|d^*_{a,K}w\|^2 + \|\nabla_s w\|^2 \\
&\quad -(\nabla_s w, d_{a,K} *_Y w) + (w, \nabla_s d_{a,K} *_Y w) \\
&= \|d_{a,K}w\|^2 + \|d^*_{a,K}w\|^2 + \|\nabla_s w\|^2 \\
&\quad -(\nabla_s w, d_{a,K} *_Y w) + (w, d_{a,K} *_Y \nabla_s w) \\
&\quad + \int_{I \times Y} \langle w \wedge *_Y [\partial_s a - d_{a,K} p \wedge *_Y w] \rangle \\
&= \|d_{a,K}w\|^2 + \|d^*_{a,K}w\|^2 + \|\nabla_s w\|^2 \\
&\quad + \int_{I \times Y} \langle w \wedge *_Y [\partial_s a - d_{a,K} p \wedge *_Y w] \rangle.
\end{aligned}$$

where, we integrated by parts in ∇_s, used the identity $\nabla_s d_{a,K} = d_{a,K}\nabla_s + \partial_s a - d_{a,K}p$, and then integrated by parts in the self-adjoint operator $d_{a,K}*$. Since $\partial_s a - d_{a,K}p$ is a component of the curvature, we can use the small curvature assumption to bound this cross term as follows

$$\begin{aligned}
\Big| \int_{I \times Y} \langle w \wedge *_Y [\partial_s a - d_{a,K} p \wedge *_Y w] \rangle \Big| &\leq 2\delta_0 \int_I \|w\|^2_{L^4(Y)}\, ds \\
&\leq 2\delta_0 C_2(\|d_{a,K}w\|^2 + \|d^*_{a,K}w\|^2) \\
&\leq \tfrac{1}{2}(\|d_{a,K}w\|^2 + \|d^*_{a,K}w\|^2),
\end{aligned}$$

where we used (33) and $\delta_0 \leq 1/4C_0$. Having bounded the cross term, we then have

$$\|d_{a,K}w\|^2_{L^2(I\times Y)} + \|d^*_{a,K}w\|^2_{L^2(I\times Y)} + \|\nabla_s w\|^2_{L^2(I\times Y)} \quad \leq \quad 2\|d_{A,\mathbf{K}}W\|^2_{L^2(I\times Y)}.$$

Combining this with (34) proves the lemma for $q = 2$. The result for $q = 4$ follows from the $q = 2$ case and Lemma 2.11. The result for general $2 \leq q \leq 4$ follows from the interpolation estimate $\|\cdot\|_{L^q} \leq \|\cdot\|^{1-\theta}_{L^2}\|\cdot\|^{\theta}_{L^4}$, where $\theta = 4(1/2 - 1/q)$. □

Proof of Theorem 5.7. We will prove that the quantity in B is positive; the positivity of the quantities in A and C can be proven similarly. We argue by contradiction. If the infimum in B is zero, then we can find a sequence A_n of smooth **K**-YM connections A_n with $\mathcal{YM}_{\mathbf{K}}(A_n) \leq \mathcal{CS}_{K,P}(a) + 1$ and $F^+_{A_n,\mathbf{K}} \neq 0$, but with the property that the self-dual curvatures are converging to zero

$$\|F^+_{A_n,\mathbf{K}}\|_{L^2(Z)} \longrightarrow 0. \tag{35}$$

Each A_n is **K**-YM with finite energy. Since all K-flat connections are non-degenerate, it follows from standard arguments that A_n is asymptotic down the end to a K-flat connection a_n and $A_n \in \mathcal{A}^{1,2}(P; a_n)$; e.g., see [6, Section 4.1]. To simplify the discussion, we assume $a_n = a$ for all n. The general case reduces to this one by the gauge invariance of the problem, the *uniform* energy bound on the A_n, and the fact that there are only finitely many gauge equivalence classes of K-flat connections.

Next, we want to take a limit of the A_n. For this, we prove the following variant of Uhlenbeck's compactness theorem.

Claim 1. *After passing to a subsequence, the A_n converge weakly in $W^{1,4}$, modulo gauge and bubbling, to a broken* **K***-ASD trajectory $(A^0; A^1, \ldots, A^J)$ asymptotic to a.*

We begin our proof of the claim by focusing on bubbling phenomena on Z. We note that, due to the perturbation, we need to be a little careful (see Remark 5.9 (b)); our argument for handling this is similar to that of [17, Prop. 11]. Suppose $z \in Z$ is a point where energy can concentrate, in the sense that there is some $\delta > 0$ so $\inf_{r>0} \int_{B_r(z)} |F_{A_n}|^2 \geq \delta$ for all n. Then conformally rescaling produces a sequence of connections A'_n defined on increasing, exhausting subsets of $\mathbb{R}^4$; see [29]. These connections are Yang–Mills relative to a metric that is converging to the standard euclidean metric, and with a perturbation that is converging pointwise to zero. It follows from Uhlenbeck's compactness theorem that the limit of the A'_n exists and is a non-flat Yang–Mills connection on $\mathbb{R}^4$ that has finite energy; in particular, it extends to a non-flat Yang–Mills connection A'_∞ on S^4. Moreover, the condition $\|F^+_{A_n,\mathbf{K}}\|_{L^2} \to 0$ is preserved under conformal changes, and so A'_∞ is actually ASD. Each non-flat ASD connection on S^4 has energy at least $\mathcal{I}_G/n_G > 0$. Since energy is conformally-invariant, the uniform energy bound on the A_n implies that this energy concentration can occur at, at most, a finite number of points in Z; let $\Omega \subset Z$ be this finite bubbling set. (We note that, a priori, there may be additional energy concentration along points that escape down the cylindrical end; these points are not contained in Ω.)

Fix a compact set $B \subset Z \backslash \Omega$. By definition of Ω, there is no energy concentration on B, and so a local version of Lemma 2.11 implies that there is a uniform $L^4(B)$-bound on the curvatures $F_{A_n,\mathbf{K}}$. We can therefore combine this with Uhlenbeck's weak compactness theorem [31, 37], as well as a diagonal argument [7, Lemma 4.4.6], to conclude the following: After passing to a subsequence and applying suitable gauge transformations, the A_n converge weakly in $W^{1,4}$ on compact subsets of $Z \backslash \Omega$. Let A^0 be the limiting connection. By Axiom 4, the perturbations $\mathbf{K}_{A_n}$ are converging to $\mathbf{K}_{A^0}$ in L^4, and so A^0 is $\mathbf{K}$-ASD by (35).

Remark 5.9. *(a) Though we do not make use of this below, we note that a Coulomb gauge argument [30, Cor. 4.3] can then be used to bootstrap this slightly to strong $W^{1,4}$-convergence on compact subsets of $Z \backslash \Omega$.*

(b) As noted by Kronheimer in the discussion following [17, Prop. 11], when considering holonomy perturbations (as we intend to do here; see Appendix A), in the presence of bubbling, it is unlikely to expect convergence in any Sobolev norm $W^{k,p}$ for $k \geq 2$. This is due to the non-local behavior of holonomy perturbations, which tend to spread out singularities in the curvature. Nevertheless, holonomy perturbations do satisfy Axiom 4, which, as we have just seen, is sufficient for our purposes.

(c) We want to emphasize that the assumption that F^+ is going to zero is used in a subtle, but crucial way in the argument above. In the absence of such an assumption, we would want to show that the limit A^0 is $\mathbf{K}$-YM. Since the $\mathbf{K}$-YM equation is second order, this would require some sort of higher order version of Axiom 4, similar to "if the A_n are weakly $W^{2,p}$-bounded on compact subsets in the complement of a finite bubbling set, then the $\mathbf{K}_{A_n}$ are Cauchy in $W^{1,p}(Z)$". As mentioned in Remark 5.9 (b), it is unlikely that any such higher order version will hold for the perturbations we have in mind.

Returning to the proof of Claim 1, we now address the additional terms that may appear in the broken trajectory. Here our strategy is similar to that of [22]; see also [12] and [20, Chapter 6]. The $\mathbf{K}$-ASD connection A^0 has finite-energy, and so is asymptotic to some K-flat connection a^0. Now we address cases.

Case 1.1: For each $\epsilon > 0$, there is some T so that

$$\sup_{s \geq T} \int_{\{s\} \times Y} |F_{A_n,\mathbf{K}}|^2 < \epsilon, \qquad \forall n. \tag{36}$$

We will show that this implies there is some gauge transformation u so that $u^* a^0 = a$; this will therefore prove Claim 1 for Case 1.1. To find this gauge transformation, let $\mathcal{A}^*(Q)$ denote the set of irreducible L^4-connections on Q, and $\mathcal{G}^{1,4}(Q)$ the group of $W^{1,4}$-gauge transformations. The quotient $\mathcal{B}^* := \mathcal{A}^*(Q)/\mathcal{G}^{1,4}(Q)$ is a smooth Banach manifold containing the (gauge equivalence classes of) K-flat connections $\mathcal{B}_\flat \subset \mathcal{B}^*$. Non-degeneracy and Uhlenbeck compactness imply that $\mathcal{B}_\flat$ is a finite set in $\mathcal{B}^*$. Fix $s \geq 0$, and let g_s^Y denote the restriction of the metric g to the slice $\{s\} \times Y \cong Y$. It follows that there is some $\epsilon_0 > 0$ so that the set

$$\left\{ [a'] \in \mathcal{B}^* \;\middle|\; \exists\, [a_\flat] \in \mathcal{B}_\flat, \|a' - a_\flat\|_{L^4(Y), g_s^Y} < \epsilon_0 \right\}$$

deformation retracts to $\mathcal{B}_\flat$. The embedding $W^{1,2} \subset L^4$ and gauge-invariance implies that there is some $\epsilon_0' > 0$ so that

$$\begin{aligned} &\left\{ [a'] \in \mathcal{B}^* \;\middle|\; \|F_{a',K}\|_{L^2(Y), g_s^Y} < \epsilon_0' \right\} \\ &\qquad \subseteq \left\{ [a'] \in \mathcal{B}^* \;\middle|\; \exists\, [a_\flat] \in \mathcal{B}_\flat, \|a' - a_\flat\|_{L^4(Y), g_s^Y} < \epsilon_0 \right\} \end{aligned}$$

(e.g., argue as in [30, Cor. 4.3]). Since the metrics g_s^Y converge in s to g^Y, these constants ϵ_0, ϵ_0' can be taken to be independent of $s \geq 0$ (e.g., these constants depend continuously on g_s^Y in the $\mathcal{C}^1(Y)$-topology, and the path $s \mapsto g_s^Y$ is precompact in the space of $\mathcal{C}^1$-metrics on Y). Let

$$\mathcal{P} \subseteq \left\{ [a] \in \mathcal{B}^* \;\middle|\; \|F_{a,K}\|_{L^2(Y), g_s^Y} < \epsilon_0' \right\}$$

be the component containing $[a]$, where a is the K-flat asymptotic limit of the A_n. It follows from these considerations that the intersection $\mathcal{P} \cap \mathcal{B}_\flat = \{[a]\}$ contains exactly one K-flat connection class.

Returning to (36), take $\epsilon = \epsilon_0'$. Then it follows that the path $[T, \infty) \to \mathcal{B}^*$ defined by $s \mapsto [a_n(s)]$ is entirely contained in $\mathcal{P}$. On the other hand, since A^0 is asymptotic to a^0, and the A_n converge to A^0 on compact sets, it follows that there is a sequence $s_n \to \infty$ so that $a_n(s_n)$ converges in $L^4(Y)$ to a^0. Hence $[a^0] \in \mathcal{P}$, and so $u_0^* a^0 = a$ for some gauge transformation u_0. This finishes the proof of Claim 1, assuming the hypothesis of Case 1.1.

Case 1.2: The hypothesis of Case 1.1 fails.

In this case, it follows that there is some $\epsilon_1 > 0$ and a sequence $(s_{1,n})_n$ diverging to ∞ so that

$$\int_{\{s_{1,n}\} \times Y} |F_{A_n, \mathbf{K}}|^2 \geq \epsilon_1 \tag{37}$$

for all n. On the other hand, since the A_n have uniformly bounded energy, for each $\epsilon > 0$, the set of points s where the estimate of (36) fails must have finite measure. Hence, by possibly redefining the $s_{1,n}$, for each $\epsilon > 0$, there is some T so that

$$\sup_{T \leq s \leq s_{1,n} - T} \int_{\{s\} \times Y} |F_{A_n, \mathbf{K}}|^2 < \epsilon, \qquad \forall n. \tag{38}$$

Now consider the translated connections $\tau_{s_{1,n}}^* A_n$. These are Yang–Mills on increasing and exhausting subsets of $\mathbb{R} \times Y$, relative to metrics that are converging to $ds^2 + g^Y$ and perturbations that are converging to $\mathbf{K}^Y$. We can therefore repeat the argument from above to conclude that a subsequence converges weakly in $W^{1,4}$ on compact subsets, modulo gauge and bubbling, to a limiting finite-energy $\mathbf{K}^Y$-ASD connection A^1 on $\mathbb{R} \times Y$. Arguing as in Case 1.1, it follows from (38) that A^1 is asymptotic to a^0 at $-\infty$, modulo gauge. Turning to the end at $+\infty$, if the translated connections $\tau^*_{s_{1,n}} A_n$ satisfy the condition of Case 1.1, then they are asymptotic at $+\infty$ to $u_1^* a$ for some gauge transformation u_1; this would finish the proof of Claim 1 in this case. Otherwise, if the condition of Case 1.1 is not satisfied, repeat the analysis of Case 1.2.

Continuing inductively, we can then find times $s_{1,n}, s_{2,n}, \ldots$ and asymptotic $\mathbf{K}$-ASD connections $A_1, A_2, \ldots$ as appearing in Claim 1. We need to show that this process terminates in a finite number of steps $J \geq 0$. Since the A_n have uniformly bounded energy, it suffices to show there is some $\hbar > 0$ so that, each time the condition of Case 1.1 fails, an energy of at least $\hbar$ is required (hence J is at most this uniform energy bound divided by $\hbar$). To see this, let $\hbar > 0$ be the minimum of $\mathcal{I}_G / n_G$ and

$$\inf_{a_\flat, a'_\flat} |\mathcal{CS}_{K,P}(a_\flat) - \mathcal{CS}_{K,P}(a'_\flat)| \tag{39}$$

where the infimum ranges over all K-flat connections $a_\flat$ and $a'_\flat$ on Q, with $\mathcal{CS}_{K,P}(a_\flat) \neq \mathcal{CS}_{K,P}(a'_\flat)$; that this infimum is positive follows because the moduli space of K-flat connections is finite. Recall that the failure of the condition of Case 1.1 implies an estimate of the form (37). This implies that either bubbling occurred or A^j is not K-flat and so has energy $\mathcal{CS}_{K,P}(a^j) - \mathcal{CS}_{K,P}(a^{j-1}) > 0$. The former case requires energy of at least $\mathcal{I}_G / n_G$ in the sequence A_n, and the latter requires energy at least that of the infimum (39). This completes the proof of Claim 1.

We will now use the convergence of Claim 1 to show that, after possibly passing to a subsequence, we have

$$F^+_{A_n, \mathbf{K}} = 0 \tag{40}$$

for all n sufficiently large. Once we have shown this, we will be done with the proof of the theorem, since (40) clearly contradicts the assumption that the A_n are not $\mathbf{K}$-ASD.

We will consider the following three cases that increase gradually in complexity.

Case 2.1: There is no energy concentration.

Case 2.2: Energy concentration occurs, but is confined to a compact set in Z.

Case 2.3: Energy concentration occurs, but it is not confined to a compact set in Z.

Proof in Case 2.1: In this case, we are assuming the A_n converge weakly in $W^{1,4}$ to $(A^0; A^1, \ldots, A^J)$, with no energy concentration at any step. We will show that there is a constant C so that

$$\|W\|_{L^q(Z)} \leq C \|d_{A_n, \mathbf{K}} W\|_{L^2(Z)} \tag{41}$$

for all $2 \leq q \leq 4$, all sufficiently large n, and all self-dual 2-forms $W \in W^{1,2}(\Omega^+)$. Then (40) will follow by applying this to $W = F^+_{A_n,\mathbf{K}}$ and using the $\mathbf{K}$-YM condition $d_{A_n,\mathbf{K}} F^+_{A_n,\mathbf{K}} = 0$. (Strictly speaking, we only need (41) for one value of q; we prove this more general estimate for later use.)

To simplify the notation, we will carry out the proof under the assumption that $g = g_{cyl}$ is cylindrical and only one trajectory breaks off in the limit (so $J = 1$); the more general case is similar, albeit more notationally cumbersome. The discussion of Cases 1.1 and 1.2, and the assumption $J = 1$, imply that there is some $T > 0$ and a sequence $(s_n)_n$ diverging to ∞, so that

$$\sup_{s \in [T, s_n - T] \cup [s_n, \infty)} \int_{\{s\} \times Y} |F_{A_n,\mathbf{K}}|^2 < \delta_0 \tag{42}$$

for all n; here δ_0 is as in Lemma 5.8. By increasing T, if necessary, we may assume $T \geq T_0$, where T_0 is the obvious constant from Lemma 5.8.

We will use the intervals in (42) to decompose Z into regions, relative to which we can carry out a patching argument similar to the one given by Floer in [12, Lemma 2d.2]. For this, we will need a suitably chosen bump function, which we define now. Let $\beta : \mathbb{R} \to [0, 1]$ be a smooth bump function that is identically 0 on $(-\infty, 0]$ and identically 1 on $[1, \infty)$. For $L \geq 1$, set $\beta_L(x) := \beta(x/L)$. Then the support of $d\beta_L$ is contained in $[0, L]$, and $|d\beta_L| \leq \|\beta\|_{\mathcal{C}^1}/L$; in particular, $\|d\beta_L\|_{L^4(\mathbb{R})} \leq \|\beta\|_{\mathcal{C}^1}/L^{3/4}$. Let C_0 be the maximum of the constant C_0 from Lemma 5.8, and the constants $C_0(A), C_0(A')$ from (32) applied to the connections $A = A^0$ and $A' = A^1$. Then fix L large enough so that

$$\|d\beta_L\|_{L^4(\mathbb{R})} \leq \frac{1}{64 C_0}.$$

Consider the cover of Z given by

$$U_0(n) := Z_0 \cup [0, T + L] \times Y, \qquad U_1(n) := [T, s_n - T] \times Y$$

$$U_2(n) := [s_n - T - L, s_n + L] \times Y, \qquad U_3(n) := [s_n, \infty) \times Y.$$

We will show that

$$\|W_i\|_{L^q(U_i(n))} \leq 4 C_0 \|d_{A_n,\mathbf{K}} W_i\|_{L^2(U_i(n))}, \qquad 0 \leq i \leq 3 \tag{43}$$

for all $2 \leq q \leq 4$, all sufficiently large n, and all smooth self-dual 2-forms W_i compactly supported in $U_i(n)$. Before proving this, we show how it is used to prove the global version (41) and therefore finish the proof in Case 2.1.

Assume that (43) holds, and fix n. Note that any common overlap of the $U_i(n)$ has length either 0 or L. Using various translations and reflections of β_L along the end of Z, it is easy to create a partition of unity

$$\beta_0(n), \beta_1(n), \beta_2(n), \beta_3(n)$$

on Z with $\mathrm{supp}(\beta_i(n)) \subset U_i(n)$ and so that $d\beta_i(n)$ is supported in the intersection of interval $U_i(n)$ with the adjacent interval(s) (e.g., $\mathrm{supp}(\beta_1(n)) \subseteq U_1(n) \cap (U_0(n) \cup U_2(n))$). It follows that

$$\|d\beta_i(n)\|_{L^4(Z)} = 2\|d\beta_L\|_{L^4(\mathbb{R})} \leq \frac{1}{32C_0}, \qquad 0 \leq i \leq 3.$$

for all n. Let W be a smooth, compactly supported self-dual 2-form on Z. Then (43) gives

$$\begin{aligned} \|W\|_{L^2(Z)} + \|W\|_{L^4(Z)} &\leq \sum_{i=0}^{3} \|\beta_i W\|_{L^2(U_i)} + \|\beta_i W\|_{L^4(U_i)} \\ &\leq 8C_0 \sum_{i=0}^{3} \|d_{A_n,\mathbf{K}}(\beta_i W)\|_{L^2(U_i)} \\ &\leq 8C_0 \sum_{i=0}^{3} \frac{1}{32C_0}\|W\|_{L^4(U_i)} + \|d_{A_n,\mathbf{K}} W\|_{L^2(U_i)} \\ &= \tfrac{1}{2}\|W\|_{L^4(Z)} + 16C_0\|d_{A_n,\mathbf{K}}W\|_{L^2(Z)}. \end{aligned}$$

Hence

$$\|W\|_{L^2(Z)} + \|W\|_{L^4(Z)} \leq 32C_0\|d_{A_n,\mathbf{K}}W\|_{L^2(Z)}.$$

Now (41) follows by interpolation and the fact that the smooth, compactly supported, self-dual 2-forms are dense in $W^{1,2}(\Omega^+)$.

To finish the proof in Case 2.1, it therefore suffices to verify (43). The estimate for $i = 1$ and $i = 3$ follows immediately from Lemma 5.8, which is valid by the small curvature estimate (42). As for $U_2(n)$, since there is no energy concentration, it follows from Uhlenbeck's compactness theorem that the translated connections $\tau_{s_n}^* A_n$ converge to A^1 weakly in $W^{1,4}([-T-L, L] \times Y)$. By the compactness of the embedding $W^{1,4} \hookrightarrow L^4$ on compact sets, we may assume n is large enough so that

$$\|\tau_{s_n}^* A_n - A^1\|_{L^4([-T-L,L]\times Y)} \leq 1/8C_0,$$

for all n sufficiently large. It follows from the second estimate in (32) that if W is a self-dual 2-form with support in $U_2(n)$, then

$$\begin{aligned} &\|W\|_{L^2(U_2(n))} + \|W\|_{L^4(U_2(n))} \\ &\qquad = \|\tau_{s_n}^* W\|_{L^2([-T-L,L]\times Y)} + \|\tau_{s_n}^* W\|_{L^4([-T-L,L]\times Y)} \\ &\qquad \leq 2C_0\|d_{A^1,\mathbf{K}}\tau_{s_n}^* W\|_{L^2([T-L,L]\times Y)} \\ &\qquad \leq 2C_0\|d_{A_n,\mathbf{K}}W\|_{L^2(U_2(n))} \\ &\qquad\qquad + 4C_0\|\tau_{s_n}^* A_n - A^1\|_{L^4([-T-L,L]\times Y)}\|W\|_{L^4(U_2(n))} \\ &\qquad \leq 2C_0\|d_{A_n,\mathbf{K}}W\|_{L^2(U_2(n))} + \tfrac{1}{2}\|W\|_{L^4(U_2(n))}. \end{aligned}$$

This implies (43) for $i = 2$. The proof for $i = 0$ is similar and left to the reader (it is slightly easier since there is no need for translations). This finishes the proof in Case 2.1.

Proof in Case 2.2: To simplify the discussion, we assume that energy concentrates at exactly one point $z_0 \in Z$. We also assume that there is no trajectory breaking occurs in the limit (so $J = 0$). The more general case is left to the reader and is just a matter of combining the techniques of Case 2.1 with those given below.

The assumption that energy does not concentrate at any $z \in Z \setminus \{z_0\}$ means that, for each $R > 0$, the limit

$$\lim_{\delta \to 0^+} \sup_{n \in \mathbb{N}, z \in Z \setminus B_R(z_0)} \int_{B_\delta(z)} |F_{A_n}|^2 = 0 \tag{44}$$

is zero. On the other hand, the assumption of energy concentration at z_0 means that there is some $\eta > 0$ so that

$$\lim_{\delta \to 0^+} \sup_{n \in \mathbb{N}} \int_{B_\delta(z_0)} |F_{A_n}|^2 \geq \eta. \tag{45}$$

To prove (40), we invoke the following trick that reduces the analysis to the cylindrical end setting without bubbling. Consider the punctured 4-manifold $Z' := Z - \{z_0\}$ with metric g. The metric is conformally equivalent to a metric g' on Z' that is asymptotically cylindrical. Our strategy is to reinterpret everything in terms of g'.

By conformal invariance of L^2-norms on 2-forms, the estimate (45) continues to hold with all metric quantities interpreted as being relative to g'. That is, when viewed as being defined on the asymptotically cylindrical manifold (Z', g'), the sequence A_n does not exhibit energy concentration. Then the same Uhlenbeck compactness from before implies that a subsequence converges weakly in $W^{1,4}$ to a broken trajectory $(A^0; B^1, \ldots, B^L)$, where A^0 is as above (viewed as a connection on Z') and the B^ℓ are finite-energy connections on $\mathbb{R} \times S^3$. (Here we are using the assumption that the original sequence, relative to g, did not have any trajectory breaking; otherwise, these additional trajectories would need to be incorporated in this g'-limit as well.) It follows from Axiom 2 that, relative to the metric g', the perturbations $\mathbf{K}_{A_n}$ converge to zero in L^∞ on the end $[0, \infty) \times S^3$ associated to z_0. This implies that the B^ℓ are ASD (no perturbation). The 2-point conformal compactification of $\mathbb{R} \times S^3$ into S^4 identifies the B^j with the ASD bubble(s) forming at z_0.

We claim that this is a setting to which the argument of Case 2.1 applies. Indeed, all flat connections on S^3 are non-degenerate (up to gauge, the only one is the trivial connection and non-degeneracy follows from the topological condition $H^1(S^3) = 0$). Hence the result of Lemma 5.8 applies with $Y = S^3$ and $K = 0$. Similarly, it is well-known that all ASD connections on S^4 are regular, and so all ASD connections on $\mathbb{R} \times S^3$ are regular since this is a conformally-invariant notion and all asymptotic limits are non-degenerate. This provides regularity estimates for the B^j that can be used in place of (32). Now the argument of Case 2.1 carries over verbatim to give a constant C so that

$$\|W\|_{L^2(Z'),g'} \leq C\|d_{A_n,\mathbf{K}}W\|_{L^2(Z'),g'}$$

for all sufficiently large n and all g'-self-dual 2-forms W of Sobolev class $W^{1,2}(Z')$. Since the Hodge star on 2-forms is conformally-invariant, the g-self dual part of any 2-forms equals the g'-self dual part. That is, we can apply the above estimate to $W = F^+_{A_n,\mathbf{K}}$ to get that $F^+_{A_n,\mathbf{K}}$ vanishes on Z' for all sufficiently large n. This clearly implies (40).

Proof in Case 2.3: In this case, bubbles can form on any of the trajectories A^j. The result in this setting follows by incorporating translations and then arguing as in Case 2.2. □

5.4 Proofs of Theorem 5.1 and Corollary 5.2

Proof of Theorem 5.1. By Theorem 3.1, there is some maximal time $\overline{\tau} \in (0, \infty]$ for which the flow $A(\tau)$ starting at $A_0 \in \mathcal{A}^{1,2}(P; a)$ exists for all $\tau \in [0, \overline{\tau})$. First assume $\overline{\tau}_1 := \overline{\tau}$ is finite. Then it follows from Proposition 4.4 and Remark 4.5, that there is some bundle $P_1 \to Z$ and a connection

$$A_1 \in \mathcal{A}^{1,2}(P_1, a)$$

so that the $A(\tau)$ converge to a pullback of A_1, and

$$\{\text{energy of bubbles}\} + \mathcal{YM}_{\mathbf{K}}(A_1) \leq \mathcal{YM}_{\mathbf{K}}(A_0).$$

In fact, we can say a little more: We have assumed that $\|F^+_{A_0,\mathbf{K}}\|^2_{L^2(Z)}$ is no greater than the constant $\eta(a)$. By (19), this L^2-norm is non-increasing along the flow, and is conformally invariant. In particular, it follows from A in the definition of $\eta(a)$ in Section 5.3 that each Yang–Mills bubble is in fact ASD. Energy quantization for ASD connections on S^4 implies that each has energy at least $\mathcal{I}_G / n_G$, where n_G is as in (31) and $\mathcal{I}_G$ is the one appearing in the statement of Theorem 5.1. In particular, the assumption that $\overline{\tau} < \infty$ implies that at least one bubble forms, and so

$$\mathcal{I}_G / n_G + \mathcal{YM}_{\mathbf{K}}(A_1) \leq \mathcal{YM}_{\mathbf{K}}(A_0). \tag{46}$$

Now we want to start the flow over again, but with initial condition A_1 in place of A_0. However, from what we have at this point, it is not clear whether A_1 has enough regularity to apply the Short-Time Existence Theorem 3.1. To obtain a suitably regular connection, consider Donaldson's flow (14) with initial condition A_1. Let A_1' be the value of this flow at any fixed positive time. This flow is smoothing, so A_1' smooth. The second Bianchi identity shows that this flow is energy non-increasing, and so $\mathcal{YM}_{\mathbf{K}}(A_1') \leq \mathcal{YM}_{\mathbf{K}}(A_1)$; in particular, (46) holds with A_1' in place of A_1. By relabeling A_1' as A_1, we may therefore assume that A_1 is a *smooth* connection in $\mathcal{A}^{1,2}(P_1, a)$ satisfying (46).

Now repeat the above procedure with A_1 in place of A_0. Continuing inductively, there are a number of times $\overline{\tau}_1, \ldots, \overline{\tau}_L$ at which bubbles can form. Associated to each $\overline{\tau}_\ell$ is a bundle P_ℓ and a smooth connection $A_\ell \in \mathcal{A}^{1,2}(P_\ell, a)$ satisfying

$$\ell \mathcal{I}_G / n_G + \mathcal{YM}_{\mathbf{K}}(A_\ell) \leq \mathcal{YM}_{\mathbf{K}}(A_0). \tag{47}$$

This shows that there can be only finitely many such times $L \geq 1$ at which bubbles form. Then the flow starting at A_L exists for all time. We will denote this flow by $A(\tau)$, with the understanding that this notation is valid for $\tau \geq \overline{\tau}_L$.

Now we wish to take the infinite-time limit. Towards this end, we make the following claim.

Claim 1. *There is a gauge transformation u on Q, and a broken* **K***-ASD trajectory $(A^0; A^1, \ldots, A^J)$ that is asymptotic to $u^* a$, and satisfies*

$$\mathcal{YM}_{\mathbf{K}}(A^0) + \sum_{j=1}^{J} \mathcal{YM}_{\mathbf{K}_Y}(A^j) \leq \mathcal{YM}_{\mathbf{K}}(A_L).$$

Note that the analysis from Proposition 4.4 is no longer valid in the infinite-time regime, so there is indeed something to be shown here. Our argument is similar to that of Claim 1 from the proof of Theorem 5.7, and so we will be relatively brief, placing emphasis on the new features.

Fix a sequence $\tau_n \to \infty$. The curvature of the $A(\tau_n)$ can only concentrate on a finite bubbling set $B \subset Z$. On the complement of B, we can appeal to Uhlenbeck's weak compactness theorem [37, Theorem B]. This implies that, after possibly passing to a subsequence, there are gauge transformations U_n so that the $U_n^* A(\tau_n)$ converge weakly in $W^{1,4}_{loc}(Z \backslash B)$ to a connection

$$A^0 \in \mathcal{A}^{1,4}_{loc}\left(P|_{Z \backslash B}\right).$$

By standard infinite-time analysis for flows, it follows that A^0 is **K**-YM on the complement of the bubbling set B. It also satisfies the energy bound

$$\mathcal{YM}_{\mathbf{K}}(A^0) \leq \mathcal{YM}_{\mathbf{K}}(A_0) = \mathcal{CS}_{K,P}(a) + \|F^+_{A_0}\|^2_{L^2(Z)} < \mathcal{CS}_{K,P}(a) + \eta(a).$$

This implies two things. First, it implies A^0 has finite energy, and so extends over B by removal of singularities to a finite-energy **K**-YM connection defined on all of Z. Second, it implies that A^0 is actually **K**-ASD. Indeed, recall we have defined $\eta(a)$ so that $\eta(a) \leq 1$, which gives $\mathcal{YM}_{\mathbf{K}}(A^0) < \mathcal{CS}_{K,P}(a) + 1$. We also have

$$\|F^+_{A^0,\mathbf{K}}\|^2_{L^2(Z)} \leq \liminf_{\tau \to \infty} \|F^+_{A(\tau),\mathbf{K}}\|^2_{L^2(Z)} \leq \|F^+_{A_0,\mathbf{K}}\|^2_{L^2(Z)} < \eta(a).$$

That A^0 is **K**-ASD now follows from B in the definition of $\eta(a)$ in Section 5.3.

Now argue as in Cases 1.1 and 1.2 in the proof of Theorem 5.7 to produce a broken **K**-ASD trajectory $(A^0; A^1, \ldots, A^J)$, with A^J asymptotic to $u^* a$ for some gauge transformation u. That each A^j is actually $\mathbf{K}_Y$-ASD (as opposed to $\mathbf{K}^Y$-YM) follows from the same argument we used to show A^0 is $\mathbf{K}$-ASD; this time one should use C from the definition of $\eta(a)$ in place of B. The energy estimate in the statement of Claim 1 follows as in the proof of Proposition 4.4.

With Claim 1 in hand, we combine it with (47) and $L \geq 1$ to get

$$\begin{aligned} \mathcal{I}_G/n_G + \mathcal{YM}_{\mathbf{K}}(A^0) + \sum_{j=1}^{J} \mathcal{YM}_{\mathbf{K}^Y}(A^j) &\leq \mathcal{YM}_{\mathbf{K}}(A_0) \\ &< \mathcal{CS}_{K,P}(a) + \eta(a) \\ &\leq \mathcal{CS}_{K,P}(a) + 1/n_G; \end{aligned}$$

in the last line, we used one of the defining conditions on $\eta(a)$ from Section 5.3. Hence

$$\mathcal{YM}_{\mathbf{K}}(A^0) + \sum_{j=1}^{J} \mathcal{YM}_{\mathbf{K}^Y}(A^j) < \mathcal{CS}_{K,P}(a) + 1/n_G - \mathcal{I}_G/n_G. \tag{48}$$

We will now give a lower bound on the energies appearing on the left. Since the connections in $(A^0; A^1, \ldots, A^J)$ are $\mathbf{K}$-ASD/$\mathbf{K}^Y$-ASD, these energies depend only on the asymptotic limits. To compute them, let a_0 be the asymptotic limit of A^0 at $+\infty$. By applying an overall gauge transformation to A^1, we can assume that its asymptotic limit at $-\infty$ is equal to a_0. Repeating inductively, we may assume that

$$a_j := \lim_{s\to+\infty} A^j|_{\{s\}\times Y} = \lim_{s\to-\infty} A^{j+1}|_{\{s\}\times Y}$$

It follows that $a_J = v^*a$ for some gauge transformation v (v may be different from u appearing in Claim 1; this difference does not play a crucial role in the argument). Then since A^0 is $\mathbf{K}$-ASD, we have

$$\mathcal{YM}_{\mathbf{K}}(A^0) = \mathcal{CS}_{K,P}(a_0).$$

The version of this for A^j with $j > 0$ is

$$\mathcal{YM}_{\mathbf{K}^Y}(A^j) = \mathcal{CS}_{K,P}(a_j) - \mathcal{CS}_{K,P}(a_{j-1});$$

indeed, it is a routine exercise to check that if A is any connection on $\mathbb{R} \times Q$ that decays sufficiently fast to $b^\pm$ at $\pm\infty$, then $\mathcal{CS}_{K,P}(b^+) - \mathcal{CS}_{K,P}(b^-) = \frac{1}{2}\int_{\mathbb{R}\times Y} \langle F_A \wedge F_A \rangle$. Adding the energies then gives

$$\begin{aligned} \mathcal{YM}_{\mathbf{K}}(A^0) + \textstyle\sum_{j=1}^{J} \mathcal{YM}_{\mathbf{K}^Y}(A^j) &= \mathcal{CS}_{K,P}(a_0) \\ &\quad + \textstyle\sum_{j=1}^{J} \mathcal{CS}_{K,P}(a_j) - \mathcal{CS}_{K,P}(a_{j-1}) \\ &= \mathcal{CS}_{K,P}(v^*a) \\ &= \mathcal{CS}_{K,P}(a) + \mathrm{Ind}_{K,P}(v^*a)/n_G \\ &\qquad - \mathrm{Ind}_{K,P}(a)/n_G, \end{aligned}$$

where we used (31) in the last line. We have assumed $\mathrm{Ind}_{K,P}(a) < \mathcal{I}_G$. Since $\mathrm{Ind}_{K,P}(a)$ and $\mathcal{I}_G$ are integers, their difference is at least 1, and so

$$\begin{aligned}\mathcal{YM}_{\mathbf{K}}(A^0) + \sum_{j=1}^J \mathcal{YM}_{\mathbf{K}^Y}(A^j) \quad \geq \quad & \mathcal{CS}_{K,P}(a) + \mathrm{Ind}_{K,P}(v^*a)/n_G \\ & + 1/n_G - \mathcal{I}_G/n_G.\end{aligned}$$

Comparing this with (48) we find

$$\mathrm{Ind}_{K,P}(v^*a) < 0.$$

Our desired contradiction will now follow from the next claim.

Claim 2. *The integer* $\mathrm{Ind}_{K,P}(v^*a)$ *is non-negative.*

To see this, recall that the quantity $\mathrm{Ind}_{K,P}(v^*a)$ is the expected dimension of the moduli space $\mathcal{M}_{ASD,\mathbf{K}}(v^*a)$ of $\mathbf{K}$-ASD connections that are asymptotic to v^*a. We have assumed that $\mathbf{K}$ is ASD-regular, which in particular means that all non-empty moduli space are smooth and of the expected dimension (which is necessarily non-negative since there are no smooth manifolds with negative dimension). It follows from Floer's gluing theorem [12] applied to the broken trajectory

$$(A^0; A^1, \ldots, A^J)$$

that there is some $\mathbf{K}$-ASD connection in $\mathcal{A}^{1,p}(P; v^*a)$. Hence $\mathcal{M}_{ASD,\mathbf{K}}(v^*a)$ is non-empty, and must therefore have non-negative dimension. This proves the claim.

This concludes our argument for long-time existence. Note that this same type of argument also excludes energy quantization at infinite time. □

Proof of Corollary 5.2. If the first estimate in the statement of the corollary does not hold, then there are $\tau_n \to \infty$ and compactly supported self-dual 2-forms W_n so that

$$\|W_n\|_{L^q(Z)} > n\|d_{A(\tau_n),\mathbf{K}} W_n\|_{L^2(Z)}$$

for all n. It follows from Theorem 5.1 that there is no energy concentration in the sequence $A_n := A(\tau_n)$. Moreover, the argument of Claim 1 in the proof of Theorem 5.1 shows that, after possibly passing to a subsequence, the A_n converge weakly in $W^{1,4}$ to a broken $\mathbf{K}$-ASD trajectory. This puts us effectively in the situation of Case 2.1 in the proof of Theorem 5.7. In particular, we can repeat the proof of (41) to get a constant C so that

$$\|W\|_{L^q(Z)} \leq C\|d_{A_n,\mathbf{K}} W\|_{L^2(Z)}$$

for all $2 \leq q \leq 4$, all n, and all self-dual 2-forms W in $W^{1,2}(Z)$. This, of course, is a contradiction.

We turn now to the second estimate of the corollary. By Lemma 2.11 and interpolation, it suffices to prove the following:

If A is any connection for which

$$\|W\|_{L^2(Z)} \leq C_0 \|d^*_{A,\mathbf{K}} W\|_{L^2(Z)}$$

for all smooth self-dual 2-forms W with compact support, then

$$\|V\|_{L^2(Z)} \leq C_0 \|d^+_{A,\mathbf{K}} V\|_{L^2(Z)},$$

where V is any smooth compactly supported 1-form in the image of $d^*_{A,\mathbf{K}} : \Omega^+ \to \Omega^1$.

To see this, let $V = d^*_{A,\mathbf{K}} W$ for some smooth compactly supported $W \in \Omega^+$. Then

$$\begin{aligned} \|V\|^2_{L^2(Z)} &\leq (d^+_{A,\mathbf{K}} V, W) \\ &\leq \|d^+_{A,\mathbf{K}} V\|_{L^2(Z)} \|W\|_{L^2(Z)} \\ &\leq C_0 \|d^+_{A,\mathbf{K}} V\|_{L^2(Z)} \|d^*_{A,\mathbf{K}} W\|_{L^2(Z)} \end{aligned}$$

The result now follows by dividing both sides by $\|V\|_{L^2(Z)} = \|d^*_{A,\mathbf{K}} W\|_{L^2(Z)}$. □

6 Infinite-time convergence

In this section we finish the proof of Theorem 1.2 by establishing convergence of the flow at infinite time. More generally, we prove the following.

Theorem 6.1 (Infinite-time convergence)**.** *In the setting of Theorem 5.1, for all $2 \leq q \leq p$, the $A(\tau)$ converge exponentially in $W^{2,q}(Z)$, as τ approaches ∞, to a unique $\mathbf{K}$-ASD connection $A_\infty \in \mathcal{A}^{2,2}(P;a) \cap \mathcal{A}^{2,p}(P;a)$. If A_0 is smooth, then $A(\tau)$ converges exponentially in $\mathcal{C}^\infty(Z)$ to A_∞ in the sense that, for every $k \geq 0$ there are constants $C, \kappa > 0$ so that*

$$\|A(\tau) - A_\infty\|_{W^{k,2}(Z)} \leq Ce^{-\kappa\tau}.$$

The proof of Theorem 6.1 is carried out in Section 6.2. Our basic analytic arguments follow those of [26, 25, 32]. See also Feehan's book [11] for a thorough treatment of the asymptotics of the flow in the absence of ASD-regularity hypotheses. Section 6.1 is preliminary in nature, establishing various estimates on the self-dual curvature and showing that this self-dual curvature converges to zero exponentially in all derivatives.

Remark 6.2. *(a) The argument of our Long-Time Existence Theorem 5.1 shows that the flow converges, modulo gauge, at infinite time to a* broken *$\mathbf{K}$-ASD trajectory. Theorem 6.1 refines this by showing the flow $A(\tau)$ converges at infinite time* on all of *Z to an actual $\mathbf{K}$-ASD connection A_∞. In particular, A_∞ is an element of the same space as the initial connection A_0.*

(b) Similar to the situation of Remark 3.2 (b), given the regularity $A_0 \in \mathcal{A}^{2,p}(P;a)$ on the initial condition, the claimed regularity $A_\infty \in \mathcal{A}^{2,p}(P;a)$ is the best we can achieve, in general.

6.1 Curvature estimates

Throughout this section, we assume that A satisfies the hypotheses of Theorem 6.1. We assume further that $A(\tau)$ is smooth for all τ sufficiently large. All unspecified Sobolev- and L^p-norms are on all of Z.

In this section, we prove certain exponential decay results for the curvature $F^+_{A,\mathbf{K}}$ and certain derivatives $\mathcal{D}^k_A F_{A,\mathbf{K}}$. Here

$$\mathcal{D}^k_A : \Omega^+(Z, \mathfrak{g}_P) \oplus \Omega^1(Z, \mathfrak{g}_P) \longrightarrow \Omega^+(Z, \mathfrak{g}_P) \oplus \Omega^1(Z, \mathfrak{g}_P)$$

is an order-k differential operator that we define now. When $k = 0$, we define $\mathcal{D}^0_A$ to be the identity map. Assume $k = 2\ell > 0$. Then on Ω^+, we define

$$\mathcal{D}^{2\ell}_A|_{\Omega^+} := (d^+_{A,\mathbf{K}} d^*_{A,\mathbf{K}})(d^+_{A,\mathbf{K}} d^*_{A,\mathbf{K}}) \ldots (d^+_{A,\mathbf{K}} d^*_{A,\mathbf{K}}) : \Omega^+ \longrightarrow \Omega^+$$

where the term $d^+_{A,\mathbf{K}} d^*_{A,\mathbf{K}}$ appears ℓ times. Similarly, on Ω^1 we define this as

$$\mathcal{D}^{2\ell}_A|_{\Omega^1} := (d^*_{A,\mathbf{K}} d^+_{A,\mathbf{K}})(d^*_{A,\mathbf{K}} d^+_{A,\mathbf{K}}) \ldots (d^*_{A,\mathbf{K}} d^+_{A,\mathbf{K}}) : \Omega^1 \longrightarrow \Omega^1.$$

For $k = 2\ell + 1 > 0$ odd, we set

$$\begin{aligned} \mathcal{D}^{2\ell+1}_A|_{\Omega^+} &:= d^*_{A,\mathbf{K}} \mathcal{D}^{2\ell}_A|_{\Omega^+} : \Omega^+(Z, \mathfrak{g}_P) \longrightarrow \Omega^1(Z, \mathfrak{g}_P) \\ \mathcal{D}^{2\ell+1}_A|_{\Omega^1} &:= d^+_{A,\mathbf{K}} \mathcal{D}^{2\ell}_A|_{\Omega^1} : \Omega^1(Z, \mathfrak{g}_P) \longrightarrow \Omega^+(Z, \mathfrak{g}_P). \end{aligned}$$

When $k < 0$, we define $\mathcal{D}^k_A$ to be the zero map.

The specific relevance of the operator $\mathcal{D}^k_A$ for us is that it arises in the flow. For example, we can write

$$\partial_\tau A = -2\mathcal{D}^1_A F^+_{A,\mathbf{K}}, \qquad \partial_\tau F^+_{A,\mathbf{K}} = -2\mathcal{D}^2_A F^+_{A,\mathbf{K}}. \tag{49}$$

Note that each $\mathcal{D}^k_A$ is symmetric relative to the L^2-inner product, and

$$\mathcal{D}^{k+\ell}_A = \mathcal{D}^k_A \circ \mathcal{D}^\ell_A, \qquad \text{for } k, \ell \geq 0. \tag{50}$$

Now we can state the main results of this section.

Theorem 6.3. *There are positive constants C, κ so the following holds for all $\tau \geq 0$ and $k = 0, 1, 2$:*

$$\|\mathcal{D}^k_{A(\tau)} F^+_{A(\tau),\mathbf{K}}\|_{L^2(Z)} + \int_\tau^{\tau+1} \|\mathcal{D}^3_A F^+_{A,\mathbf{K}}\|_{L^2(Z)}\, d\tau \leq Ce^{-\kappa\tau}.$$

Theorem 6.4. *For every $c_0 > 0$ and integer $k \geq 3$, there are $C, \kappa > 0$ such that if*

$$\|A(\tau) - A_{ref}\|_{W^{k-1,2}(Z)} \leq c_0, \qquad \forall \tau \geq 0,$$

then the following holds for all $\tau \geq 0$:

$$\|\mathcal{D}^k_{A(\tau)} F^+_{A(\tau),\mathbf{K}}\|_{L^2(Z)} + \int_\tau^{\tau+1} \|\mathcal{D}^{k+1}_A F^+_{A,\mathbf{K}}\|_{L^2(Z)}\, d\tau \leq Ce^{-\kappa\tau}.$$

Together, these imply bounds for $\mathcal{D}^k_A F^+_{A,\mathbf{K}}$ for all $k \geq 0$. As suggested by these theorems, the case $k = 2$ is a certain borderline case, after which different techniques become necessary. In our proof of Theorem 6.3, we will see that, when $k \leq 2$, there are very few lower order terms that need to be estimated, and these do not require any further hypotheses on the connections A. This reflects the fact that the Yang–Mills functional is quadratic. However, when $k \geq 3$ there are higher order terms that we have been unable to estimate without the hypotheses on the connections $A(\tau)$ appearing in Theorem 6.4. Nevertheless, we will see in the next section that, almost by happy accident, the conclusions of Theorem 6.3 are sufficient to imply that the hypotheses of Theorem 6.4 hold for $k = 3$. This will then allow us to prove infinite-time convergence inductively.

Before giving the proofs of Theorems 6.3 and 6.4, we establish the following preliminary estimates.

Lemma 6.5. *For each $k \geq 0$, there is some C_k so the following holds for all $\tau \geq 0$ and all $2 \leq q \leq 4$:*

$$\|\mathcal{D}^k_{A(\tau)} F^+_{A(\tau),\mathbf{K}}\|_{L^q(Z)} \leq C_k \|\mathcal{D}^{k+1}_{A(\tau)} F^+_{A(\tau),\mathbf{K}}\|_{L^2(Z)}. \tag{51}$$

Proof. We will work under the assumption that the initial condition A_0 is *not* $\mathbf{K}$-YM; the case where A_0 is $\mathbf{K}$-YM is relatively easy and left to the reader.

We first claim that $A(\tau)$ is not $\mathbf{K}$-YM for any $\tau \geq 0$. Indeed, suppose $A(\tau_0)$ were $\mathbf{K}$-YM for some $\tau_0 > 0$. Then the uniqueness of the flow would imply $A(\tau) = A(\tau_0)$ for all $\tau \geq \tau_0$. It then follows from Corollary 5.2 that there is a constant C so that

$$\|W\|_{L^q(Z)} \leq C \|d_{A(\tau),\mathbf{K}} W\|_{L^2(Z)}$$

for $\tau \in [\tau_0, \infty)$ and all self-dual 2-forms W. Since the $A(\tau)$ converge in $L^4(Z)$ to $A(\tau_0)$ as $\tau \nearrow \tau_0$, there is some $\delta > 0$ so that the same estimate holds for all τ in the slightly larger interval $[\tau_0 - \delta, \infty)$. For such τ, apply this estimate with $W = F^+_{A(\tau),\mathbf{K}}$ to get that $A(\tau)$ is $\mathbf{K}$-ASD for all $\tau \geq \tau_0 - \delta$; hence $A(\tau)$ is constant for all $\tau \geq \tau_0 - \delta$. Arguing in this way, it follows that $A(\tau)$ is $\mathbf{K}$-ASD (and constant in τ) for all $\tau \geq 0$. This contradicts the assumption that A_0 is not $\mathbf{K}$-YM.

To prove the lemma, fix $\tau \geq 0$. When $k = 0$, the existence of a constant $C_0 = C_0(\tau)$ satisfying (51) is now obvious form the fact that $A(\tau)$ is not $\mathbf{K}$-YM. For $k > 0$, the existence of such a constant $C_k = C_k(\tau)$ follows from the identities

$$(d^+_{A,\mathbf{K}} d^*_{A,\mathbf{K}} U, U) = \|d^*_{A,\mathbf{K}} U\|^2_{L^2(Z)}, \qquad (d^*_{A,\mathbf{K}} d^+_{A,\mathbf{K}} V, V) = \|d^+_{A,\mathbf{K}} V\|^2_{L^2(Z)},$$

for $U \in \Omega^+, V \in \Omega^1$, which show that $d^+_{A,\mathbf{K}}$ (resp. $d^*_{A,\mathbf{K}}$) is injective on the image of $d^*_{A,\mathbf{K}}$ (resp. $d^+_{A,\mathbf{K}}$). For each k, the constants $C_k(\tau)$ can be chosen to depend continuously on τ. In particular, for any $\tau_0 \geq 0$, the supremum $C_k := \sup_{0 \leq \tau \leq \tau_0} C_k(\tau) > 0$ is positive. Apply this with τ_0 equal to the constant from Corollary 5.2. This proves the result for all $0 \leq \tau \leq \tau_0$. The result for $\tau \geq \tau_0$ follows from Corollary 5.2. $\square$

Proof of Theorem 6.3. To simplify the notation, we prove this theorem under the assumption that the perturbation $\mathbf{K} = 0$ is zero. With the use of Axiom 2, it is not hard to adapt our proofs to handle non-zero $\mathbf{K}$ as well.

It is convenient to introduce the algebraic operator Λ defined by

$$\Lambda V := \begin{cases} -*V & \text{if } V \in \Omega^3(Z, \mathfrak{g}_P) \\ \frac{1}{2}(1+*)V & \text{if } V \in \Omega^+(Z, \mathfrak{g}_P) \end{cases}$$

(We will not need an extension of this to forms of other degrees.)

To also set up for the proof of Theorem 6.4 below, we begin by working with all $k \geq 0$. Differentiating $\|\mathcal{D}_A^k F_A^+\|_{L^2}^2$ in τ and using (49) gives

$$\begin{aligned} \tfrac{d}{d\tau}\|\mathcal{D}_A^k F_A^+\|_{L^2}^2 &= 2\Big(\mathcal{D}_A^k \partial_\tau F_A^+, \mathcal{D}_A^k F_A^+\Big) \\ &\quad +2\sum_{\ell=0}^{k-1}\Big(\mathcal{D}_A^{k-\ell-1}\Lambda\Big[\partial_\tau A \wedge \mathcal{D}_A^\ell F_A^+\Big], \mathcal{D}_A^k F_A^+\Big) \\ &= -4\Big(\mathcal{D}_A^k \mathcal{D}_A^2 F_A^+, \mathcal{D}_A^k F_A^+\Big) \\ &\quad -4\sum_{\ell=0}^{k-1}\Big(\mathcal{D}_A^{k-\ell-1}\Lambda\Big[\mathcal{D}_A^1 F_A^+ \wedge \mathcal{D}_A^\ell F_A^+\Big], \mathcal{D}_A^k F_A^+\Big), \end{aligned}$$

where we used the flow in the second line. The identity (50) and the fact that $\mathcal{D}_A^j$ is symmetric then gives

$$\begin{aligned} \tfrac{d}{d\tau}\|\mathcal{D}_A^k F_A^+\|_{L^2}^2 &= -4\|\mathcal{D}_A^{k+1} F_A^+\|_{L^2}^2 \\ &\quad -4\sum_{\ell=0}^{k-1}\Big(\mathcal{D}_A^{k-\ell-1}\Lambda\Big[\mathcal{D}_A^1 F_A^+ \wedge \mathcal{D}_A^\ell F_A^+\Big], \mathcal{D}_A^k F_A^+\Big) \end{aligned} \tag{52}$$

Claim 1. *Suppose that for all $\delta > 0$ there is some $\tau_\delta > 0$ so that*

$$\Big|\Big(\mathcal{D}_A^{k-\ell-1}\Lambda\Big[\mathcal{D}_A^1 F_A^+ \wedge \mathcal{D}_A^\ell F_A^+\Big], \mathcal{D}_A^k F_A^+\Big)\Big| \leq \delta\|\mathcal{D}_A^{k+1}F_A^+\|_{L^2}^2 \tag{53}$$

for all $\tau > \tau_\delta$ and all $0 \leq \ell \leq k-1$. Then there are constants C, κ so that $\|\mathcal{D}_A^k F_A^+\|_{L^2}^2 \leq Ce^{\kappa\tau}$.

To prove the claim, use (52) and Lemma 6.5 to write

$$\begin{aligned} \tfrac{d}{d\tau}\|\mathcal{D}_A^k F_A^+\|_{L^2}^2 &\leq -2C_k^{-1}\|\mathcal{D}_A^k F_A^+\|_{L^2}^2 - 2\|\mathcal{D}_A^{k+1}F_A^+\|_{L^2}^2 \\ &\quad +4\sum_{\ell=0}^{k-1}\Big|\Big(\mathcal{D}_A^{k-\ell-1}\Lambda\Big[\mathcal{D}_A^1 F_A^+ \wedge \mathcal{D}_A^\ell F_A^+\Big], \mathcal{D}_A^k F_A^+\Big)\Big|. \end{aligned}$$

If the condition (53) holds, then take $\delta = 1/2k$ to obtain

$$\frac{d}{d\tau}\|\mathcal{D}_A^k F_A^+\|_{L^2}^2 \leq -2C_k^{-1}\|\mathcal{D}_A^k F_A^+\|_{L^2}^2.$$

Then exponential decay for $\|\mathcal{D}_A^k F_A^+\|_{L^2}^2$ follows from this identity and integration.

Now we specialize to the situation of Theorem 6.3, and show that the condition of Claim 1 holds for $k = 0, 1, 2$. When $k = 0$, this is trivial since the hypothesis of (53) is empty (i.e., there are no terms in the summation in (52)). This gives exponential decay for $\|F_{A(\tau)}^+\|_{L^2}^2$.

When $k = 1$, there is only one term to bound, and this is

$$
\begin{aligned}
\left|\left(\Lambda\left[\mathcal{D}_A^1 F_A^+ \wedge F_A^+\right], \mathcal{D}_A^1 F_A^+\right)\right| &\leq 8\|F_A^+\|_{L^2}\|\mathcal{D}_A^1 F_A^+\|_{L^4}^2 \\
&\leq 8C_1\|F_A^+\|_{L^2}\|\mathcal{D}_A^2 F_A^+\|_{L^2}^2,
\end{aligned}
$$

where we used Lemma 6.5. By the $k = 0$ case, we can assume τ is sufficiently large so that $\|F_A^+\|_{L^2}$ is as small as we wish. This establishes (53), which gives exponential decay for $\|\mathcal{D}_{A(\tau)}^1 F_{A(\tau)}^+\|_{L^2}^2$.

When $k = 2$, we need to estimate the two terms

$$
\left|\left(\Lambda\left[\mathcal{D}_A^1 F_A^+ \wedge \mathcal{D}_A^1 F_A^+\right], \mathcal{D}_A^2 F_A^+\right)\right|, \qquad \left|\left(\Lambda\left[\mathcal{D}_A^1 F_A^+ \wedge F_A^+\right], \mathcal{D}_A^3 F_A^+\right)\right|.
$$

The first of these terms can be estimated as in the $k = 1$ case. As for the second term, use Lemma 6.5 to write

$$
\begin{aligned}
\left|\left(\Lambda\left[\mathcal{D}_A^1 F_A^+ \wedge F_A^+\right], \mathcal{D}_A^3 F_A^+\right)\right| &\leq \|F_A^+\|_{L^4}\left(\|\mathcal{D}_A^1 F_A^+\|_{L^4}^2 + \|\mathcal{D}_A^3 F_A^+\|_{L^2}^2\right) \\
&\leq C_0(C_1 C_2 + 1)\|\mathcal{D}_A^1 F_A^+\|_{L^2}\|\mathcal{D}_A^3 F_A^+\|_{L^2}^2.
\end{aligned}
$$

Then exponential decay for $\|\mathcal{D}_{A(\tau)}^2 F_{A(\tau)}^+\|_{L^2}^2$ follows form the claim and the $k = 1$ case.

All that remains is to prove exponential decay for

$$
\int_\tau^{\tau+1} \|\mathcal{D}_A^3 F_A^+\|_{L^2}\, d\tau.
$$

For this, use Hölder's inequality in the time variable to write

$$
\begin{aligned}
\left(\int_\tau^{\tau+1} \|\mathcal{D}_A^3 F_A^+\|_{L^2} d\tau\right)^2 &\leq \int_\tau^{\tau+1} \|\mathcal{D}_{A(\tau)}^3 F_{A(\tau)}^+\|_{L^2}^2 d\tau \\
&\leq \tfrac{1}{4}\|\mathcal{D}_{A(\tau)}^2 F_{A(\tau)}^+\|_{L^2}^2 \\
&\quad - \int_\tau^{\tau+1}\left(\Lambda\left[\mathcal{D}_A^1 F_A^+ \wedge \mathcal{D}_A^1 F_A^+\right], \mathcal{D}_A^2 F_A^+\right) d\tau \\
&\quad - \int_\tau^{\tau+1}\left(\Lambda\left[\mathcal{D}_A^1 F_A^+ \wedge F_A^+\right], \mathcal{D}_A^3 F_A^+\right) d\tau,
\end{aligned}
$$

where we used (52). When we verified (53) in the previous paragraph, we showed that the second and third lines of this are bounded by

$$
\frac{1}{2}\int_\tau^{\tau+1} \|\mathcal{D}_{A(\tau)}^3 F_{A(\tau)}^+\|_{L^2}^2 d\tau,
$$

provided τ is sufficiently large. This gives

$$\Big(\int_{\tau}^{\tau+1}\|\mathcal{D}_A^3 F_A^+\|_{L^2}d\tau\Big)^2 \leq \frac{1}{2}\|\mathcal{D}^2_{A(\tau)}F^+_{A(\tau)}\|^2_{L^2},$$

and the desired estimate follows from the exponential decay of $\|\mathcal{D}^2_{A(\tau)}F^+_{A(\tau)}\|^2_{L^2}$. □

Our proof of Theorem 6.4 will be similar to the proof just given, except it will take more work to estimate the lower order terms in (52). For this, will be be interested in how Sobolev norms depend on the connection A used to define them. To keep track of this dependence, we will write

$$\|V\|_{W^{k,p}(Z),A} := \Big(\sum_{0\leq\ell\leq k}\|\nabla_A^\ell V\|^p_{L^p(Z)}\Big)^{1/p}$$

for the $W^{k,p}$-Sobolev norm on $\Omega^\bullet(Z,\mathfrak{g}_P)$ defined using A and the Levi-Civita connection on Z. The relevant estimates are as follows.

Lemma 6.6. *There is a constant C so that if $k\geq 0$ is any integer and $A\in\mathcal{A}(P;a)$ is smooth, then*

$$\|\cdot\|_{W^{k,4}(Z),A}\leq C\|\cdot\|_{W^{k+1,2}(Z),A}.$$

If, in addition $k\geq 3$, then

$$\|\cdot\|_{L^\infty(Z)}\leq C\|\cdot\|_{W^{k,2}(Z),A}.$$

Proof. Let W be a section of $\otimes^\ell T^*Z\otimes\mathfrak{g}_P$ for some $\ell\geq 0$. Consider the real-valued function $f=|W|$. Then the Sobolev embedding $W^{1,2}(Z)\subset L^4(Z)$ for real-valued functions on Z gives

$$\|W\|_{L^4(Z)}=\|f\|_{L^4(Z)}\leq C(\|f\|_{L^2(Z)}+\|df\|_{L^2(Z)}).$$

The covariant derivative ∇_A on $\otimes^\ell T^*Z\otimes\mathfrak{g}_P$ is a metric connection, so Kato's inequality gives the pointwise estimate $|df|\leq|\nabla_A W|$. Then the first estimate of the lemma follows by taking $W=\nabla_A^\ell V$ for $0\leq\ell\leq k+1$. The second estimate is similar and relies on the embedding $W^{k,2}(Z)\subset L^\infty(Z)$ for real-valued functions; this embedding holds provided $k\geq 3$. □

Lemma 6.7. *Fix a smooth reference connection A_{ref}. For each $c_0>0$ and integer $k\geq 3$, there are constants c, C so that the following holds. If $A\in\mathcal{A}(P;a)$ is a smooth connection satisfying*

$$\|A-A_{ref}\|_{W^{k-1,2}(Z),A_{ref}}\leq c_0,$$

then

$$c\|\cdot\|_{W^{k,2}(Z),A}\leq\|\cdot\|_{W^{k,2}(Z),A_{ref}}\leq C\|\cdot\|_{W^{k,2}(Z),A}. \tag{54}$$

Proof. We prove the estimate on the right of (54); the other estimate is similar. Fix $1 \leq \ell \leq k$ and $V \in \Omega^\bullet(Z, \mathfrak{g}_P)$. It suffices to bound $\|\nabla^\ell_{A_{ref}} V\|_{L^2(Z)}$ in terms of $\|V\|_{W^{k,2}(Z),A}$. We have an identity of the form

$$\nabla^\ell_{A_{ref}} V = \nabla^\ell_A V + \sum_{0 \leq j \leq \ell-1} \nabla^{\ell-j-1}_{A_{ref}}((A - A_{ref}) \# \nabla^j_A V),$$

where here and below we are using # to denote a bilinear algebraic operation. Taking the L^2-norm of each side, using the product rule, and then Hölder's inequality gives

$$\begin{aligned}
\|\nabla^\ell_{A_{ref}} V\|_{L^2(Z)} &\leq \|\nabla^\ell_A V\|_{L^2(Z)} + \|(\nabla^{\ell-1}_{A_{ref}}(A - A_{ref})) \# V\|_{L^2(Z)} \\
&\quad + C \sum_{1 \leq j \leq \ell-1} \|\nabla^{\ell-j-1}_{A_{ref}}(A - A_{ref})\|_{L^4(Z)} \|\nabla^j_A V\|_{L^4(Z)} \\
&\leq \|\nabla^\ell_A V\|_{L^2(Z)} + C' \|A - A_{ref}\|_{W^{\ell-1,2}(Z),A_{ref}} \|V\|_{L^\infty(Z)} \\
&\quad + C' \sum_{1 \leq j \leq \ell-1} \|\nabla^{\ell-j}_{A_{ref}}(A - A_{ref})\|_{L^2(Z)} \|\nabla^{j+1}_A V\|_{L^2(Z)} \\
&\leq C'' \|A - A_{ref}\|_{W^{k-1,2}(Z),A_{ref}} \|V\|_{W^{k,2}(Z),A} \\
&\leq c_0 C''' \|V\|_{W^{k,2}(Z),A}
\end{aligned}$$

where we used Lemma 6.6. □

Proof of Theorem 6.4. We prove the theorem for $\mathbf{K} = 0$, leaving the case of nonzero $\mathbf{K}$ to the reader. We will prove exponential decay of $\|\mathcal{D}^k_{A(\tau)} F^+_{A(\tau)}\|_{L^2(Z)}$ by induction on k, using Theorem 6.3 as the base case.

We therefore fix $k \geq 3$ and assume the claim of Theorem 6.4 holds for all integers less than k. By Claim 1 appearing in the proof of Theorem 6.3, it suffices to verify the condition (53). Towards this end, we begin by working with the $\ell = k - 1$ term

$$\left| \left(\Lambda \left[\mathcal{D}^1_A F^+_A \wedge \mathcal{D}^{k-1}_A F^+_A \right], \mathcal{D}^k_A F^+_A \right) \right|$$

Here we can use Lemma 6.5 and Theorem 6.3 to write

$$\begin{aligned}
\left| \left(\Lambda \left[\mathcal{D}^1_A F^+_A \wedge \mathcal{D}^{k-1}_A F^+_A \right], \mathcal{D}^k_A F^+_A \right) \right| &\leq \|\mathcal{D}^1_A F^+_A\|_{L^2} \|\mathcal{D}^{k-1}_A F^+_A\|_{L^4} \|\mathcal{D}^k_A F^+_A\|_{L^4} \\
&\leq C e^{-\kappa\tau} \|\mathcal{D}^{k+1}_A F^+_A\|^2_{L^2},
\end{aligned}$$

which is the desired estimate for $\ell = k - 1$, provided τ is sufficiently large.

We may therefore assume that $0 \leq \ell \leq k - 2$. Then we have

$$\begin{aligned}
&\left| \left(\mathcal{D}^{k-\ell-1}_A \Lambda \left[\mathcal{D}^1_A F^+_A \wedge \mathcal{D}^\ell_A F^+_A \right], \mathcal{D}^k_A F^+_A \right) \right| \\
&\qquad = \left| \left(\mathcal{D}^{k-\ell-2}_A \Lambda \left[\mathcal{D}^1_A F^+_A \wedge \mathcal{D}^\ell_A F^+_A \right], \mathcal{D}^{k+1}_A F^+_A \right) \right| \\
&\qquad \leq C' \|\mathcal{D}^{k+1}_A F_A\|_{L^2} \sum_{j=1}^{k-\ell-1} \|F^+_A\|_{W^{j,4},A} \|F^+_A\|_{W^{k-1-j,4},A}.
\end{aligned}$$

By Lemma 6.6, this is bounded by a uniform constant times

$$\|\mathcal{D}_A^{k+1} F_A^+\|_{L^2} \sum_{j=1}^{k-\ell-1} \|F_A^+\|_{W^{j+1,2},A} \|F_A^+\|_{W^{k-j,2},A}$$

We will be done if we can show that the summand $\|F_A^+\|_{W^{j+1,2},A}\|F_A^+\|_{W^{k-j,2},A}$ is small for all $1 \leq j \leq k-\ell-1$. Note that $j+1$ and $k-j$ are both no greater than $k-1$. Then we can use our hypotheses on $A - A_{ref}$ and Lemma 6.7 to write

$$\|F_A^+\|_{W^{j+1,2},A} \|F_A^+\|_{W^{k-j,2},A} \quad \leq \quad C'' \|F_A^+\|^2_{W^{k-1,2},A_{ref}}$$

Elliptic regularity for the operator $\mathcal{D}^{k-1}_{A_{ref}}$ on self-dual 2-forms gives a bound of the form

$$\|F_A^+\|^2_{W^{k-1,2},A_{ref}} \leq C''' \Big(\|F_A^+\|_{L^2} + \|\mathcal{D}^{k-1}_{A_{ref}} F_A^+\|^2_{L^2} \Big).$$

As in the proof of Lemma 6.7, we can use the bound on $\|A - A_{ref}\|_{W^{k-1,2},A_{ref}}$ to convert from $\mathcal{D}^{k-1}_{A_{ref}}$ to $\mathcal{D}^{k-1}_A$, bounding this further by a constant times

$$\|F_A^+\|_{L^2} + \|\mathcal{D}_A^{k-1} F_A^+\|^2_{L^2}.$$

By the inductive hypothesis, this can be made as small as we want. By Claim 1 in the proof of Theorem 6.3, the decay estimate for $\|\mathcal{D}^k_{A(\tau)} F^+_{A(\tau),\mathbf{K}}\|_{L^2(Z)}$ follows.

The decay estimate for

$$\int_\tau^{\tau+1} \|\mathcal{D}_A^{k+1} F^+_{A,\mathbf{K}}\|_{L^2(Z)}$$

is obtained by the argument we gave for the analogous term appearing in Theorem 6.3. □

6.2 Proof of Theorem 6.1

Let A be a solution to the flow, as in Theorem 6.1, and set

$$\|\cdot\|_{L^p} := \|\cdot\|_{L^p(Z)}, \qquad \|\cdot\|_{W^{k,p}} := \|\cdot\|_{W^{k,p}_{A_{ref}}(Z)}$$

where A_{ref} is a fixed smooth reference connection on Z. We begin by reducing the general case to the case where A is smooth. Recall from the proof of Theorem 3.1 that the solution A has the form

$$A(\tau) = (u(\tau)^{-1})^* B(\tau),$$

where u is a path of gauge transformations, B is a path of connections, and where B is *smooth* for positive time. Fix any $\tau^\circ > 0$, and set

$$A^\circ(\tau) := (u^\circ)^* A(\tau + \tau^\circ), \qquad u^\circ := u(\tau^\circ).$$

Then this satisfies the flow (10) with the *smooth* initial condition $A^\circ(0) = B(\tau^\circ)$. In particular, A° is smooth in all variables and exists for all $\tau \geq 0$. Assume we can show that $A^\circ(\tau)$ converges exponentially in $\mathcal{C}^\infty$ to a some **K**-ASD connection A°_∞. Then we claim that, for each $2 \leq q \leq p$, the $A(\tau)$ converge exponentially in $W^{2,q}$ to

$$A_\infty := ((u^\circ)^{-1})^* A^\circ_\infty.$$

To see this, it suffices to show that u° has regularity $W^{3,q}$ on Z. We already know from the proof of Theorem 3.1 that u° has regularity $W^{1,p}$, and it is not hard to see that this can be refined to having regularity in $W^{1,2} \cap W^{1,p} \subset W^{1,q}$. The higher regularity on u° now follows from a two-step bootstrapping argument using the identity

$$du^\circ = u^\circ A(\tau^\circ) - B(\tau^\circ)u^\circ$$

and the fact that $B(\tau^\circ)$ is smooth, and $A(\tau^\circ)$ has regularity $W^{2,q}$.

It therefore suffices to prove Theorem 6.1 under the assumption that A is smooth in all variables. This puts us in a setting where the discussion of Section 6.1 is valid. We remind the reader of the operators $\mathcal{D}^k_A$, defined in that section. Our proof at this stage is carried out in several steps.

Step 1. *The $A(\tau)$ converge exponentially in L^2 to some $A_\infty \in \mathcal{A}^{0,2}(P; a)$.*

Integrate $\partial_\tau A = -2\mathcal{D}^1_A F^+_{A,\mathbf{K}}$ over an interval $[\tau_a, \tau_b]$ to get

$$A(\tau_a) - A(\tau_b) = 2\int_{\tau_a}^{\tau_b} \mathcal{D}^1_A F^+_{A,\mathbf{K}}. \tag{55}$$

Take the L^2-norm of both sides and then use Theorem 6.3 to get

$$\|A(\tau_a) - A(\tau_b)\|_{L^2} \leq 2\int_{\tau_a}^{\tau_b} \|\mathcal{D}^1_A F^+_{A,\mathbf{K}}\|_{L^2} \leq 2C_1\kappa^{-1}(e^{-\kappa\tau_a} - e^{-\kappa\tau_b}).$$

This shows that $\{A(\tau)\}_\tau$ is L^2-Cauchy and so converges in $L^2(Z)$ to some limiting connection $A_\infty \in \mathcal{A}^{0,2}(P; a)$. This argument also shows exponential convergence in L^2:

$$\|A(\tau) - A_\infty\|_{L^2} \leq 2C_1\kappa^{-1}e^{-\kappa\tau}, \qquad \tau \geq 0. \tag{56}$$

Remark 6.8. *Let C_0 be any constant for which*

$$\|F^+_{A(\tau),\mathbf{K}}\|_{L^2} \leq C_0\|d_{A(\tau),\mathbf{K}} F^+_{A(\tau),\mathbf{K}}\|_{L^2}$$

for all $\tau \geq 0$. An inspection of the proofs of Theorem 6.3 and Lemma 6.5 show that if $\|F^+_{A_0,\mathbf{K}}\|_{L^2} \leq 1/(16C_0)$, then the constants C_1 and κ appearing in (56) can be taken to be

$$C_1 = \frac{1}{2}\|d^*_{A_0,\mathbf{K}} F_{A_0,\mathbf{K}}\|_{L^2}, \qquad \kappa = C_0^{-1}.$$

Step 2. *The connections $A(\tau)$ converge exponentially in L^4 to A_∞.*

Take the L^4-norm of (55), then use Lemma 6.5 and Theorem 6.3 to get

$$\begin{aligned} \|A(\tau_a) - A(\tau_b)\|_{L^4} &\leq 2\int_{\tau_a}^{\tau_b} \|\mathcal{D}_A^1 F_{A,\mathbf{K}}^+\|_{L^4}\, d\tau \\ &\leq 2C_2 \int_{\tau_a}^{\tau_b} \|\mathcal{D}_A^2 F_{A,\mathbf{K}}^+\|_{L^2}\, d\tau \\ &\leq 2C_2 C\kappa^{-1}(e^{-\kappa\tau_a} - e^{-\kappa\tau_b}), \end{aligned}$$

As in Step 1, this implies that $A(\tau)$ converges exponentially in L^4; the limit is necessarily A_∞.

The next step shows that the $L^4 \cap L^2$-convergence we just established implies convergence in $W^{1,2}$. To set up for an induction argument later, we prove a more general result.

Step 3. *Fix an integer $k \geq 1$ and assume that the $A(\tau)$ converge exponentially in $W^{k-1,4} \cap W^{k-1,2}$. Then the $A(\tau)$ converge exponentially in $W^{k,2}$.*

Recall the Sobolev norms are defined relative to the fixed reference connection A_{ref}. By elliptic regularity for $d^+_{A_{ref}} \oplus d^*_{A_{ref}}$, we have

$$\begin{aligned} \|A - A_\infty\|_{W^{k,2}} &\leq C\Big(\|d^+_{A_{ref},\mathbf{K}}(A - A_\infty)\|_{W^{k-1,2}} + \|d^*_{A_{ref},\mathbf{K}}(A - A_\infty)\|_{W^{k-1,2}} \\ &\qquad + \|A - A_\infty\|_{W^{k-1,2}}\Big) \\ &\leq C'\Big(\|d^+_{A_\infty,\mathbf{K}}(A - A_\infty)\|_{W^{k-1,2}} + \|d^*_{A_\infty,\mathbf{K}}(A - A_\infty)\|_{W^{k-1,2}} \\ &\qquad + \|A_\infty - A_{ref}\|_{W^{k-1,4}}\|A - A_\infty\|_{W^{k-1,4}} + \|A - A_\infty\|_{W^{k-1,2}}\Big) \end{aligned}$$

Our assumptions imply that the last two terms are going to zero in τ, so it suffices to show

$$\lim_{\tau\to\infty} \|d^+_{A_\infty,\mathbf{K}}(A(\tau) - A_\infty)\|_{W^{k-1,2}} + \|d^*_{A_\infty,\mathbf{K}}(A(\tau) - A_\infty)\|_{W^{k-1,2}} = 0.$$

We will work with the first limit; the other is similar (use the second Bianchi identity). Apply $d^+_{A_\infty,\mathbf{K}}$ to both sides of (55) to get

$$\begin{aligned} d^+_{A_\infty,\mathbf{K}}(A(\tau) - A_\infty) &= -2d^+_{A_\infty,\mathbf{K}} \int_\tau^\infty \mathcal{D}_A^1 F_{A,\mathbf{K}}\, d\tau \\ &= -2\int_\tau^\infty d^+_{A_\infty,\mathbf{K}} \mathcal{D}_A^1 F^+_{A,\mathbf{K}}\, d\tau \end{aligned}$$

Consider the identity

$$d^+_{A_\infty,\mathbf{K}} = d^+_{A,\mathbf{K}} + \frac{1}{2}(1+*)\Big(\,[A_\infty - A \wedge \cdot] + d\mathbf{K}_A - d\mathbf{K}_{A_\infty}\Big),$$

where $d\mathbf{K}_A$ is the linearization of the perturbation at A. Using this identity, taking the $W^{k-1,2}$-norm, and using Axiom 2 gives

$$\begin{aligned}\|d^+_{A_\infty,\mathbf{K}}(A(\tau)-A_\infty)\|_{W^{k-1,2}} \quad \leq \quad & C\int_\tau^\infty \|\mathcal{D}^2_A F^+_{A,\mathbf{K}}\|_{W^{k-1,2}} \\ & + \|A_\infty - A\|_{W^{k-1,4}}\|\mathcal{D}^1_A F^+_{A,\mathbf{K}}\|_{W^{k-1,4}}\, d\tau.\end{aligned}$$

Our assumptions imply a uniform bound on $\|A_\infty - A\|_{W^{k-1,4}}$. Now use elliptic regularity for the operator $\mathcal{D}^{k-1}_{A_{ref}}$ on self-dual 2-forms to continue this and get

$$\begin{aligned}\|d^+_{A_\infty,\mathbf{K}}(A(\tau)-A_\infty)\|_{W^{k-1,2}} \quad \leq \quad & C'\int_\tau^\infty \|\mathcal{D}^{k-1}_{A_{ref}}\mathcal{D}^2_A F^+_{A,\mathbf{K}}\|_{L^2} + \|\mathcal{D}^2_A F^+_{A,\mathbf{K}}\|_{L^2} \\ & + \|\mathcal{D}^{k}_{A_{ref}}\mathcal{D}^1_A F^+_{A,\mathbf{K}}\|_{L^2} + \|\mathcal{D}^1_A F^+_{A,\mathbf{K}}\|_{L^2}\, d\tau.\end{aligned}$$

As in the proof of Lemma 6.7, we can convert from $\mathcal{D}^\ell_{A_{ref}}$ to $\mathcal{D}^\ell_A$ at the cost of picking up lower order terms. These lower order terms can be controlled uniformly using Lemma 6.5. In summary, we have

$$\|d^+_{A_\infty,\mathbf{K}}(A(\tau)-A_\infty)\|_{W^{k-1,2}} \quad \leq \quad C''\int_\tau^\infty \|\mathcal{D}^{k+1}_A F^+_{A,\mathbf{K}}\|_{L^2}\, d\tau.$$

It follows from Theorem 6.3 (for $k = 1,2$) and Theorem 6.4 (for $k \geq 3$) that there are constants C, κ (depending on k) so that

$$\int_{\tau+j}^{\tau+j+1} \|\mathcal{D}^{k+1}_A F^+_{A,\mathbf{K}}\|_{L^2}\, d\tau \leq Ce^{-\kappa(\tau+j)}$$

for all j, τ. This gives

$$\begin{aligned}\|d^+_{A_\infty,\mathbf{K}}(A(\tau)-A_\infty)\|_{W^{k-1,2}} \quad & \leq \quad C''\sum_{j=0}^\infty \int_{\tau+j}^{\tau+j+1} \|\mathcal{D}^{k+1}_A F^+_{A,\mathbf{K}}\|_{L^2}\, d\tau \\ & \leq \quad C''C\sum_{j=0}^\infty e^{-\kappa(\tau+j)} \\ & = \quad C''Ce^{-\kappa\tau}(1-e^{-\kappa})^{-1}.\end{aligned}$$

Step 3 follows from this estimate, and a similar one for $\|d^*_{A_\infty,\mathbf{K}}(A(\tau)-A_\infty)\|_{W^{k-1,2}}$.

Step 4. *For each $k \geq 2$, if the $A(\tau)$ converge exponentially in $W^{k-1,2}$, then the $A(\tau)$ converge exponentially in $W^{k,2}$.*

By Step 3, it suffices to show that show that the $A(\tau)$ converge exponentially in $W^{k-1,4}$. For this, take the $W^{k-1,4}$-norm of (55) to get

$$\|A(\tau+j) - A(\tau+j+1)\|_{W^{k-1,4}} \leq 2\int_{\tau+j}^{\tau+j+1} \|\mathcal{D}^1_A F^+_{A,\mathbf{K}}\|_{W^{k-1,4}}\, d\tau.$$

As in Step 3, we can bound this by

$$C\int_{\tau+j}^{\tau+j+1} \|\mathcal{D}^{k+1}_A F^+_{A,\mathbf{K}}\|_{L^2}\, d\tau$$

for some uniform constant C. When $k = 2$, we can apply Theorem 6.3 to bound this by a constant times $e^{-\kappa(\tau+j)}$. Then the exponential convergence follows by summing over j, as we did at the end of Step 3. When $k \geq 3$ the hypothesis that the $A(\tau)$ converge in $W^{k-1,2}$ implies that Theorem 6.4 applies, and so we can repeat the same argument we gave for $k = 2$.

Theorem 6.1 now follows from an induction argument using Step 4, with base case given by the $W^{1,2}$-convergence established in Step 3. □

A Existence of ASD-regular perturbations

To avoid a vacuous discussion of long-time existence and infinite-time convergence, we would like some sort of existence statement for perturbations satisfying the axioms of Section 2.1. This is supplied by the following theorem.

Theorem A.1. *Assume $Q \to Y$ is such that all flat connections are irreducible. Then there exists a perturbation $\mathbf{K}$ that is ASD-regular and satisfies Axioms 1, 2, 3, and 4.*

Sketch of Proof. We will consider perturbations having the form described in Example 2.3, since these automatically satisfy Axioms 1 and 3. Below, we refer to the notation of Example 2.3 (and hence Example 2.1 (b)). Restrict further to the family $\mathcal{F}$ of perturbations where the function h is defined by considering holonomy over thickened loops in the surface Σ; perturbations of this type were considered by Dostoglou and Salamon [8, Section 7]. Then each $\mathbf{K} \in \mathcal{F}$ satisfies Axioms 2 and 4; see [17, Prop. 7] and [17, Lem. 10], respectively.

Now we turn to ASD-regularity. When all flat connections are irreducible, the same is true for all K-flat connections, provided K is sufficiently small. The key point then is that this family $\mathcal{F}$ is large enough to contain a comeager set of perturbations satisfying the remaining conditions of ASD-regularity. Indeed, the existence of this comeager set follows from a Sard-Smale argument that is now fairly standard in gauge theory; we refer the reader to Donaldson's book [6, Section 5.5] for a nice general treatment. This argument ultimately comes down to the idea that, modulo gauge, connections are distinguished by their holonomy. That it suffices to consider holonomy only in the Σ-directions (e.g., as opposed to all four directions in $U \times \Sigma$) follows because, modulo gauge, the $\mathbf{K}$-ASD connections on neighborhoods $U \times \Sigma$ are determined by their Σ-component $\alpha(x) := A|_{\{x\}\times\Sigma}$, provided the $\alpha(x)$ are irreducible for all $x \in U$. This irreducibility condition can be arranged by further refining the choice of $\mathbf{K}$. □

The next example shows that bundles Q satisfying the hypotheses of Theorem A.1 are fairly abundant.

Example A.2. *(a) Suppose $G = \mathrm{SO}(3)$ and Y has positive first Betti number. Fix any non-torsion class $\gamma \in H_1(Y, \mathbb{Z}_2)$, and define $Q \to Y$ to be the principal $\mathrm{SO}(3)$-bundle whose Stiefel-Whitney class $w_2(Q) \in H^2(Y, \mathbb{Z}_2)$ is Poincaré dual to γ. Then all flat connections on Q are irreducible.*

This strategy generalizes to $G = \mathrm{PU}(r)$ for $r \geq 2$.

(b) Suppose Y is any 3-manifold. Then taking the connect sum with the torus $Y \# T^3$ produces a 3-manifold with positive first Betti number. In particular, for each $r \geq 2$, the manifold $Y \# T^3$ admits a $\mathrm{PU}(r)$-bundle with no reducible flat connections. This strategy is due to Kronheimer–Mrowka [18].

References

[1] M. F. Atiyah, N. J. Hitchin and I. M. Singer, Self-duality in four-dimensional Riemannian geometry, *Proc. Roy. Soc. London Ser. A* **362**(1711) (1978) 425–461.

[2] P. J. Braam, S. K. Donaldson, Fukaya-Floer homology and gluing formulae for polynomial invariants, *The Floer Memorial Volume* (Birkäuser Verlag, 1995).

[3] J-P. Bourguignon, H. B. Lawson, Stability and isolation phenomena for Yang–Mills fields, *Comm. Math. Phys.* **79**(2) (1981) 189–230.

[4] G. Daskalopoulos, R. Wentworth, Convergence properties of the Yang–Mills flow on Kähler surfaces, *J. Reine Angew.* **575** (2004) 69–99.

[5] S. Donaldson, Anti-self dual Yang–Mills connections over complex algebraic surfaces and stable vector bundles, *Proc. London Math. Soc.* **50** (1985) 1–26.

[6] S. Donaldson, *Floer homology groups in Yang–Mills theory* (Cambridge University Press, 2002).

[7] S. Donaldson, P. Kronheimer, *The Geometry of Four-Manifolds* (Clarendon Press, 1997).

[8] S. Dostoglou, D. Salamon, Instanton homology and symplectic fixed points, *LMS Lecture Notes Series* 192 (Cambridge University Press, 1993).

[9] S. Dostoglou, D. Salamon, Self-dual instantons and holomorphic curves *Ann. of Math.* **139** (1994) 581–640.

[10] D. Duncan, *Compactness Results for the Quilted Atiyah-Floer Conjecture* Thesis, Rutgers University (2013).

[11] P. Feehan, Global Existence and Convergence of Smooth Solutions to Yang–Mills Gradient Flow over Compact Four-Manifolds (2014) arXiv.1409.1525v1.

[12] A. Floer, An instanton-invariant for 3-manifolds, *Comm. Math. Phys.* **118** (1988) 215–240.

[13] D. Freed, K. Uhlenbeck, *Instantons and Four-Manifolds* (Springer, 2nd ed, 1991).

[14] G. Y. Guo, Yang–Mills fields on cylindrical manifolds and holomorphic bundles I, *Comm. Math. Phys.* **179** (1996) 737–775.

[15] R. Janner, Perturbed geodesics on the moduli space of flat connections and Yang–Mills theory (2010) arXiv:1006.5340v1.

[16] H. Kozono, Y. Maeda, H. Naito, Global solutions for the Yang–Mills gradient flow for 4-manifolds, *Nagoya Math. J.* **139** (1995) 93–128.

[17] P. Kronheimer, Four-manifold invariants from higher-rank bundles, *J. Diff. Geom.* **70**(1) (2005) 59–112.

[18] P. Kronheimer, T. Mrowka, Knot homology groups from instantons, *J. Topol.* **4**(4) (2011) 835–918.

[19] R. B. Lockhart, R. C. McOwen, Elliptic differential operators on noncompact manifolds, *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze* 4^e *série* **12**(3) (1985) 409–447.

[20] J. Morgan, T. Mrowka, D. Ruberman, *The L^2-moduli space and a vanishing theorem for Donaldson polynomial invariants*, Monographs in Geometry and Topology, II. (International Press, 1994).

[21] H. N. Sà Earp, G_2-instantons over asymptotically cylindrical manifolds, *G. & T.* **19**(1) (2015) 61–111.

[22] D. Salamon. Lectures on Floer homology, *Park City Series* **7** (1999) 145–229.

[23] D. Salamon, Quantum products for mapping tori and the Atiyah-Floer conjecture, *Amer. Math. Soc. Transl.* **196** (1999) 199-235.

[24] S. Sedlacek, A direct method for minimizing the Yang–Mills functional over 4-manifolds, *Comm. Math. Phys.* **86** (1982) 515–527.

[25] A. Schlatter, Global existence of the Yang–Mills flow in four dimensions, *J. Reine Angew.* **479** (1996) 133–148.

[26] A. Schlatter, Long-time behavior of the Yang–Mills flow in four-dimensions, *Annals of Global Analysis and Geometry* **15**(1) (1997) 1–25.

[27] A. Schlatter, M. Struwe, A. S. Tahvildar-Zadeh, Global existence of the equivariant Yang–Mills heat flow in four space dimensions, *American Journal of Mathematics* **120**(1) (1998) 117–128.

[28] M. Struwe, The Yang–Mills flow in four-dimensions, *Calc. Var.* **2**(9) (1994) 123–150.

[29] C. H. Taubes, Self-dual Yang–Mills connections on non-self-dual four-manifolds, *J. Diff. Geom.* **17** (1982) 139–170.

[30] K. Uhlenbeck, The Chern classes of Sobolev connections, *Comm. Math. Phys.* **101** (1985) 449–457.

[31] K. Uhlenbeck, Connections with L^p bounds on the curvature, *Comm. Math. Phys.* **83** (1982) 31–42.

[32] A. Waldron, Instantons and singularities in the Yang–Mills flow (2014) arXiv:1402.3224v1.

[33] A. Waldron, Long-time existence for Yang–Mills flow (2016) arXiv:1610.02324v1.

[34] K. Wehrheim, Floer Field philosophy (2016) arXiv:1602.04908.

[35] K. Wehrheim, C. Woodward Floer field theory (2018) arXiv:1601.04924v2.

[36] K. Wehrheim, C. Woodward, Quilted Floer cohomology, *Geo. & Top.*, **14** (2010) 833–902.

[37] K. Wehrheim, *Uhlenbeck compactness*, (EMS Series of Lectures in Mathematics, 2004).

[38] L.M. Woodward, The classification of principal PU_n-bundles over a 4-complex, *J. London Math. Soc.* **25** (1982) 513–524.