Geometry & Topology 12 (2008) 1265–1311

Growth of Casson handles and transversality for ASD moduli spaces

TSUYOSHI KATO

In this paper we study the growth of Casson handles which appear inside smooth four-manifolds. A simply-connected and smooth four-manifold admits decompositions of its intersection form. Casson handles appear around one side of the end of them, when the type is even. They are parameterized by signed infinite trees and their growth measures some of the complexity of the smooth structure near the end. We show that with respect to some decompositions of the forms on the K3 surface, the corresponding Casson handles cannot be of bounded type in our sense. In particular they cannot be periodic. The same holds for all logarithmic transforms which are homotopically equivalent to the K3 surface. We construct Yang–Mills gauge theory over Casson handles of bounded type, and verify that transversality works over them.

57M30, 57R57; 14J80

1 Introduction

Casson handles are open four-manifolds with boundary, which arise when simply-connected, oriented four-manifolds are decomposed topologically with respect to their intersection forms. Let us describe how Casson handles appear inside smooth four-manifolds. Let M be a simply-connected, smooth, closed and oriented four-manifold of even type. Let us fix an isomorphism of the form

$$\Phi\colon (H_2(M;\mathbb{Z}),(\ ,\ )) \mapsto (\mathbb{Z}^{8k} \oplus \mathbb{Z}^{2l}, -kE_8 \oplus lH)$$

where H is the hyperbolic matrix (the standard form of $S^2 \times S^2$). It is called a marking.

M Freedman constructed an oriented, topological and closed four-manifold $|-E_8|$ with $\pi_1(M)=1$ whose intersection form is isomorphic to $-E_8$. Moreover any Φ above is induced from a homeomorphism:

$$\varphi\colon M \cong k|-E_8|\#l(S^2 \times S^2)$$

as $\Phi = \varphi_*$, where $k|-E_8|$ is the k times connected sums of $|-E_8|$ (see Freedman [5]). There exists an open subset $S \subset M$ homeomorphic to $l(S^2 \times S^2)\setminus \mathrm{pt}$, which is

Published: 3 June 2008 DOI: 10.2140/gt.2008.12.1265

compatible with the decomposition of the forms. In particular S admits an induced smooth structure. This smoothing on the end is far from the standard product $S^3 \times [0, \infty)$ in general. Casson and Freedman constructed the end of S explicitly. Let D^4 be the zero handle, a standard open four ball with boundary. Then S admits a smooth decomposition:

$$S \cong D^4 \cup \bigcup_{i=1}^{2l} CH(T_i).$$

$CH(T_i)$ are called *Casson handles*, which are embedded inside M smoothly.

Each Casson handle can be represented by an infinite tree T_i with one end point, whose edges are assigned with signs $\pm$. We call T_i *signed infinite trees*, or just *signed trees*. When two signed trees admit an embedding $T \subset T'$, then correspondingly there is a smooth embedding of Casson handles $CH(T') \subset CH(T)$ preserving the attaching circles. Thus growth of the signed trees representing S above measures the complexity of the smooth structure of $S \subset M$ on the end.

There are sufficiently many Casson handles. In fact, for each signed infinite tree, one can associate a Casson handle. It is known that there are uncountably many Casson handles which are mutually nondiffeomorphic (see the references in Kato [10]). In fact even the simplest cases, periodic Casson handles, are known to be exotic (see Bižaca and Gompf [1]). On the other hand, they are all homeomorphic to the standard open 2 handle $(\bar{D}^2 \times D^2, S^1 \times D^2)$ (see Freedman [5]).

We say that an open four-manifold S has *tree-like end* if there is a finite family of signed trees $T_1, \dots, T_l$, such that S is diffeomorphic to $D^4 \cup \bigcup_{j=1}^{l} CH(T_j)$, where every $CH(T_j)$ is attached to the zero handle D^4 along the attaching S^1 of the first stage kinky handle (which corresponds to the end point in T_j). Notice that there is a natural smooth structure on it induced from the handle body.

In [10; 11], we defined a subclass of signed infinite trees which are called *homogeneous trees of bounded type*. These are constructed by iterating to attach infinitely many half periodic trees. To be precise, let $\mathbb{R}_+$ be the nonnegative real line equipped with the same signs on each edge $+$ or $-$. The corresponding $CH(\mathbb{R}_+)$ is the periodic Casson handles. There are two types of them with respect to the signs $\pm$. The set of vertices in $\mathbb{R}_+$ are $\mathbb{R}_+ \supset \mathbb{N} = \{0, 1, 2, \dots\}$. Attach another infinite number of the same $\mathbb{R}_+$ on each vertex in $\mathbb{R}_+$ as

$$T_2^0 = \mathbb{R}_+ \cup \bigcup_{n \in \mathbb{N}} \mathbb{R}_+.$$

Again let us take infinite number of the same T_2^0, and attach the base vertex of each T_2^0 with each $n \in \mathbb{N} \subset \mathbb{R}_+$. Then one obtains

$$T_3^0 = \mathbb{R}_+ \cup \bigcup_{n \in \mathbb{N}} T_2^0.$$

One can iterate the same process and obtain T_m^0, $m = 1, 2, \ldots$. There are 2^m numbers of type T_m^0 with respect to choices of signs, and these are the homogeneous trees of bounded type.

Definition 1.1 A connected signed infinite subtree of a homogeneous tree of bounded type is called a *tree of bounded type*.

A Casson handle of a tree of bounded type is also said to be of bounded type.

Thus a Casson handle of bounded type contains one of homogeneously bounded type preserving the attaching circle. It is known that they are all exotic (see [10]). Notice that any tree of bounded type grows polynomially, and so for example regular infinite trees of valency more than two, like the binary trees are not included in this class.

Let M be as above, and take an $SO(3)$ bundle $E \to M$. With respect to any marking $H_2(M:\mathbb{Z}) \cong \mathbb{Z}^{8k} \oplus \mathbb{Z}^{2l}$, the second Stiefel–Whitney class w_2 of E decomposes as $w_2 = w_2^1 \oplus w_2^2$. Let us say that a marking is *generic* with respect to the $SO(3)$ bundle E, if both $w_2^1 \neq 0$ and $w_2^2 \neq 0$ do not vanish.

Let us recall that any marking gives an open four-manifold $S \subset M$. In this paper we show the following Theorem.

Theorem 1.2 *Let M be K3 surface or its logarithmic transforms X_p for odd p. Then there is an $SO(3)$ bundle $E \to M$ with $w_2(E) \neq 0$ which admits nonempty generic markings, so that for any generic marking, the corresponding embedded Casson handles in S can not be all of bounded type.*

There are many logarithmic transforms which are homotopically equivalent to K3, but nondiffeomorphic (see Friedman and Morgan [6], Gompf and Mrowka [9] and Kronheimer [12]).

This bundle admits many generic markings. In fact the construction of E above shows that with respect to the decomposition of the form over the Kummer surface, there is some $\mathbb{Z}^{12} \subset \mathbb{Z}^8 \oplus \mathbb{Z}^8 \subset H_2(M:\mathbb{Z})$ on the $-2E_8$ side such that w_2 does not vanish on each generating element in $\mathbb{Z}^{12}$. By the results on the unimodular forms, there are automorphisms of the form so that any generator in the $-2E_8$ side can be transformed into the $3H$ side.

Diffeomorphisms φ: $M \cong M$ induce the automorphisms on the lattices. Let Φ be a generic marking, and denote the induced one by $\varphi_*(\Phi)$. Then the open subset S' corresponding to $\varphi_*(\Phi)$ also satisfies the conclusion of Theorem 1.2, and so the Casson handles in S' cannot be all of bounded type. This leads to a question whether one can remove the condition of genericity in Theorem 1.2, by use of diffeomorphisms. The situation drastically depends whether M is K3 or homotopy K3.

For the K3 surface, many of automorphisms of lattices are induced by diffeomorphisms. Let $O(-2E_8 \oplus 3H)$ be the group of automorphisms of the lattices.

Proposition 1.3 (Matumoto [13]) *There is an index 2 subgroup $O_+ \subset O(-2E_8 \oplus 3H)$ so that any element in O_+ is induced by a diffeomorphism on the K3 surface.*

Corollary 1.4 *For any marking on the K3 surface, the corresponding embedded Casson handles in S cannot all be of bounded type.*

Proof The following argument is due to Furuta. Let Φ_0 be a generic marking on K3. Then the conclusion follows for Φ_0 by Theorem 1.2. Let us take any marking Φ. When an element g in O_+ transforms Φ_0 to Φ as $\Phi = g(\Phi_0)$, then the conclusion follows for Φ by Proposition 1.3.

Let us consider the remaining case. With respect to the splitting $-2E_8 \oplus 3H = H \oplus (-2E_8 \oplus 2H)$, let us denote the basis of the first factor H by x, y. Let R' be an element in $O(-2E_8 \oplus 3H)$ such that

$$x \to -y, \quad y \to -x, \quad \xi \to \xi \quad (\xi \in (\mathbb{Z}^{20}, -2E_8 \oplus 2H)).$$

Then it is known that $O(-2E_8 \oplus 3H)$ is generated by R' and elements of O_+ (Wall [18], see Matumoto [13]).

Now suppose a marking Φ satisfies $\Phi = gR'(\Phi_0)$ for some $g \in O_+$. Since R' transforms one factor H into itself and on the other factors acts as the identity, it preserves genericity of markings. So $R'(\Phi_0)$ is also generic. Since an element $g \in O_+$ transforms Φ to a generic marking, it follows by the above argument that the conclusion also holds for Φ. This verifies the remaining case. □

Let us consider the case of homotopy K3, including X_p for p odd. The Seiberg–Witten invariant SW_M gives strong constraints. It is a functional on the set of spin^c structures $\text{Spin}^c = \{\beta \in H^2(M : \mathbb{Z}) : \beta \equiv w_2(TM) \text{mod } 2\}$. The set of basic classes is given by spin^c structures with nonvanishing SW invariants. It is known that the set is finite. More strongly for any symplectic homotopy K3 surface, the lattice spanned by the basic classes L_M is known to be isotropic, and so its rank is less than 3 (see Chen and Kwasik [2]). Since SW_M is a diffeomorphism invariant, if $\beta \in L_M$, then $\varphi^*(\beta) \in L_M$ must be also satisfied for any diffeomorphism φ. This observation gives several constraints on the sizes of diffeomorphism groups, when L_M is nontrivial. On the other hand, by the work on symplectic four-manifolds by Taubes [14; 15], L_M is known to be trivial (zero) for the standard K3, since in this case the canonical bundle is trivial.

Thus for homotopy K3 surfaces, it is hopeless to approach the above question by use of diffeomorphisms, however still it seems reasonable to expect the conclusion of Theorem 1.2 for any markings on homotopy K3 surfaces.

The analytic tools to study such structure of Casson handles are Yang–Mills gauge theory and the Donaldson invariant. In the application of gauge theory, there are two ingredients from functional analytic view points: one is Fredholm theory of linearized equations and the other is transversality of instanton moduli spaces. For the former we have treated this in [10]. The main theme in this paper is the construction for the latter. For application of the gauge theory to growth of Casson handles, the basic idea is described also in [11].

Theorem 1.2 follows from Theorem 1.9 and Proposition 1.5 below. Let Φ be any marking on the K3 surface, and $w_2 = w_1^1 + w_2^2$ be the corresponding decomposition of the second Stiefel–Whitney class $w_2 \in H^2(M:\mathbb{Z}_2)$.

Proposition 1.5 (Kronheimer [12]) *Let M be K3 surface and X_p be the logarithmic transforms. Then over both M and X_p, there are $SO(3)$ bundles E with $|p_1(E)| = 6$ and generic markings satisfying $w_2^1, w_2^2 \neq 0$ so that the Donaldson invariants satisfy:*

$$Q(M) = \pm 1, \quad Q(X_p) = \pm p.$$

We have Fredholm property for the Atiyah–Hitchin–Singer (AHS) complex.

Theorem 1.6 (Kato [10]) *Let $S = D^4 \cup \bigcup_{j=1}^{2l} CH(T_j)$ have a tree-like end whose trees are homogeneous of bounded type. Then there is a pair, a complete Riemannian metric g and a weight function w both of bounded geometry on S so that the bounded complex*

$$0 \longrightarrow W_w^{k+1}(S,g) \xrightarrow{d} W_w^k((S,g);\Lambda^1) \xrightarrow{d^+} W_w^{k-1}((S,g);\Lambda_+^2) \longrightarrow 0$$

is Fredholm with the cohomology groups $H^0(AHS) = 0$, $H^1(AHS) = 0$ and $l = \dim H_+^2(S:\mathbb{R}) \leq \dim H^2(AHS) \leq 2l = \dim H^2(S:\mathbb{R})$.

One can construct Riemannian metrics explicitly on each building block and kinky handle, and attach them isometrically.

Let S be an oriented open four-manifold which is simply-connected and also simply-connected at infinity. A pair of a complete Riemannian metric g and a weight function

w, both of bounded geometry on S, is called *admissible* if the triplet (S, g, w) satisfies the conclusion of Theorem 1.6.

The proof of Theorem 1.2 relies only on admissibility of metrics and weight functions, and do not use other special properties. Thus we have the following Theorem.

Theorem 1.7 *Let $S \subset M$ be the open four-manifolds with tree-like ends in the Theorem 1.2. Then there are no admissible pairs of metrics and weight functions (g, w) on S.*

Thus the Fredholm theory for the AHS complex cannot work on such Casson handles in the K3 surface and its logarithmic transforms.

This complex is the linearization of the *Yang–Mills moduli space* (ASD moduli space) which is one of the most important subjects in four dimensional geometry. Let $E \to S$ be an $SO(3)$ bundle whose trivialization near the end is fixed. Notice that both the end of S and S itself are simply-connected. A connection A over E is called *anti self dual* (ASD), if its curvature F_A satisfies the equation

$$F_A^+ \equiv F_A + *F_A = 0.$$

Let A_0 be an ASD connection over (E, S, g) whose curvature is L^2 finite, that is, $\|F_{A_0}\|_{L^2(S,g)} < \infty$. Then the ASD moduli space $\mathfrak{M}(A_0)_w$ is given by the set of the ASD solutions in $A_0 + W_w^k(S; E \otimes T^*S)$ modulo gauge transformation, where W_w^k are the weighted Sobolev k spaces. It is a smooth finite dimensional manifold near $A \in A_0 + W_w^k(S; E \otimes T^*S)$, when d_A^+ is surjective. Notice that the base connection A_0 has a meaning only when the space S is open. Let $\Sigma_1, \dots, \Sigma_n \subset S$ be embedded surfaces. Then as in the construction of the Donaldson invariant, one obtains the restricted ASD moduli space $\mathfrak{M}(A_0)_w(\Sigma_1, \dots, \Sigma_n)$. Its dimension is dim $\mathfrak{M}(A_0)_w - 2n$.

In order to realize transversality we study the perturbed ASD equations by use of holonomy. In this paper it will play one of the most important roles. The holonomy perturbation is local, and it has metrics when one considers families of solutions over an infinite sequences of spaces. Also in Section 3.2.8 we will use a property that it gives enough supplementary sections. As far as these properties are satisfied, basically our argument can be applicable for other local perturbations.

Let M be a simply-connected, oriented, closed, smooth four-manifold of even type, equipped with a marking

$$\Phi\colon (H_2(M;\mathbb{Z}), \langle , \rangle) \cong (\oplus^{8k}\mathbb{Z} \oplus^{2l} \mathbb{Z}, \quad k(-E_8) \oplus lH)$$

where $k \geq 2$ and $l \geq 3$. In this paper we always assume l is odd. By Casson–Freedman [5], one finds an open four-manifold S with tree-like end, homeomorphic to the interior of $l(S^2 \times S^2)\backslash D^4$, and finds a smooth embedding $S \hookrightarrow M$ which induces an embedding of the form:

$$(H^2(S:\mathbb{Z}), \langle\ ,\ \rangle) \cong (\oplus^{2l}\mathbb{Z}, lH) \hookrightarrow (H^2(M:\mathbb{Z}), \langle\ ,\ \rangle).$$

Definition 1.8 Let $E \to M$ be an $SO(3)$ bundle and take a generic marking Φ. The triplet (M, E, Φ) is nondegenerate, if there are exhausting compact subsets $K_0 \subset K_1 \subset \cdots \subset S$ and a family of generic Riemannian metrics $\{g_i\}_i$ so that:

1. the $SO(3)$ Donaldson invariant $Q(E; [\Sigma_1], \dots, [\Sigma_n])$ is nonzero for some classes $[\Sigma_1], \dots, [\Sigma_n] \in H_2(S; \mathbb{Z})$,
2. $\big||g_i|K_i - g|K_i|\big|_{C^l} \to 0$ as $i \to \infty$ for all $l = 0, 1, \dots$
3. for any family of (perturbed) ASD connections $A_i \in \mathfrak{M}(E; \Sigma_1, \dots, \Sigma_n)$ over (M, g_i), a subsequence converges after gauge transformations, to an ASD connection A over (S, g) without bubbling on S on each compact subset in C^∞.

By Uhlenbeck compactness (see Lemma 2.1) any family of such connections contains a convergent subsequence, but in general bubbling may occur.

The main Theorem follows.

Theorem 1.9 *Suppose M has a nondegenerate triplet (M, E, Φ). Then with respect to Φ, the embedded Casson handles cannot be all of bounded type.*

This gives an answer to the continuity problem addressed in [11] under the nondegeneracy condition.

When $|p_1(E)|$ is small, bubbling does not occur as below (see the proof of Lemma 2.4).

Lemma 1.10 *Let E be an $SO(3)$ bundle with $|p_1| \leq 6$ and equip with a generic marking. Suppose the $SO(3)$ Donaldson invariant $Q(E)$ does not vanish. Then the condition in Definition 1.8(3) is satisfied.*

Proof Since $Q(E) \neq 0$, families of (perturbed) ASD solutions exist. Thus the conclusion follows except for the bubbling statement. By Uhlenbeck compactness, a subsequence of the family $\{A_i\}_i$ converges to A on each compact subset in C^∞ except bubbling points. One can assume that the family $\{A_i\}_i$ itself converges by replacing the indices and $\{K_i\}_i$ by $\{K_{l_i}\}_i$ if necessarily.

As the space M is deformed by the family of Riemannian metrics, E splits as E_1 and E_2 over S and S^c respectively. By genericity, both bundles E_1 and E_2 are nontrivial. Since the base spaces S and S^c are both simply-connected, they are not flat bundles.

If bubbling could occur, then the estimate

$$|p_1(E_1)| \leq 6-4-2=0$$

would hold, and hence E_1 would be flat. This is a contradiction. □

Let us describe the idea of the proof of the main Theorem roughly. For simplicity, assume the generic ASD moduli space $\mathfrak{M}(E)$ over M is zero-dimensional. Suppose the embedded Casson handles are all of bounded type with respect to the generic marking Φ. Then by Theorem 1.6 there exists an admissible triplet (S, g, w). We induce a contradiction by the following argument. Let us take an exhaustion on S by compact subsets as

$$K_0 \subset K_1 \subset \cdots \subset S \subset M.$$

Choose a family of generic Riemannian metrics $\{g_i\}_i$ such that $g_i|K_i \sim g|K_i$ are sufficiently near each other in C^∞. By the condition in Definition 1.8(3), there exists a family of ASD connections A_i with respect to (M, g_i) which converges to an L^2 ASD connection A with respect to (S, g). By taking A as a base connection, one can construct the ASD moduli space $\mathfrak{M}(A, g)_w$, where we use weighted Sobolev spaces. Its formal dimension is finite, and so if g could be generic with respect to (S, A), then it should be a nonempty finite dimensional smooth manifold. But its dimension is in fact negative, which gives a contradiction.

This argument implies that these Casson handles should grow much more than bounded type so that the Fredholm theory breaks under nondegeneracy condition. In this paper we show that in fact this argument essentially works. Because we use infinitely many spaces at the same time, the perturbation is required to be local. Here we use holonomy perturbation by Floer [4].

In order to see the difficulty of the transversality condition, let us try to follow a standard argument of generic perturbation of metrics. The Freed–Uhlenbeck generic perturbation method of Riemannian metrics also works over the open manifold S. Let $\mathfrak{C}$ be the set of sufficiently small perturbations of g by automorphisms of the tangent spaces. Then there exists a Baire set $\mathfrak{C}(A) \subset \mathfrak{C}$ so that for any $g' \in \mathfrak{C}(A)$, the corresponding weighted ASD moduli spaces $\mathfrak{M}(A, g')_w$ will be a smooth finite dimensional manifold, if it is nonempty. Now let us choose any g' as above. One can choose another generic family $\{g'_i\}_i$ with $g'_i \to g'$ as $i \to \infty$ and ASD connections $\{A'_i\}_i$ over (M, g'_i). Then again a subsequence converges to another A' which is ASD with respect to (S, g'). The point is that A' may not be an element in $\mathfrak{M}(A, g')_w$, since a priori, their difference

$$A - A' \in L^2((S, g'); \Lambda^1 \otimes \mathrm{Ad}\, P)$$

is just in L^2. In order for A' in $\mathfrak{M}(A, g')_w$, the difference should be in the weighted L^2 space $L^2_w((S, g'); \Lambda^1 \otimes \mathrm{Ad}\, P)$. On the other hand if one chooses A' as a base connection, then the moduli space $\mathfrak{M}(A', g')_w$ may not be a smooth manifold, since $g' \in \mathfrak{C}(A)$ depends on A and is generic only for $\mathfrak{M}(A, \ )_w$.

In this paper we overcome this difficulty by using the projection method of function spaces so that irrelevant parts of connections which might affect the estimate norms of connections are eliminated, and we pick up only restrictions of connections on $S \subset M$. More precisely even though $A_j | K_j^c$ will degenerate very badly, we show that the $A_j | K_j$ part behaves very much like the ordinary convergence so that one can apply the above idea. For this we use the quotient Sobolev spaces $\overline{W}^*_w(K_i, g) \equiv W^*_w(S, g)/W^*_w(K_i^c, g)_0$. By use of holonomy perturbation, essentially we show that if we project perturbed connections $\overline{A}'_j$, then they converge to some A' which is uniformly bounded from A in $W^{k+1}_w(S, g)$. Following the contradiction argument above, if embedded Casson handles could be of bounded type, then using the Fredholm property of the AHS complex with respect to the admissible triplet (S, g, w), one would be able to obtain the nonempty perturbed ASD moduli space $\mathfrak{M}_b(A)_w$ over S for a generic element b of the perturbation. Thus it becomes a regular manifold, but its formal dimension should be negative, which gives a contradiction.

In order to follow the above idea using these function spaces, we will use mostly more general arguments of functional analysis which are not special to ASD equations, and will work also for general operators of Fredholm type. In fact in order to find out A' above in the proof of Theorem 1.9 in Section 3.2, we use a particular property of ASD only in Section 3.2.9 and Section 3.2.10, that the Donaldson invariant does not vanish.

We believe that detailed analysis of bubbling phenomenon will allow the extension of our method to the case of higher dimensional moduli spaces. It would have some applications to our contents as follows. Our bounded type Casson handles are required to have some conditions on signs on the corresponding trees. In Theorem 1.2, one would expect to exclude Casson handles whose signed trees can be changed to bounded type by blowing up, since it can eliminate negative double points from immersed surfaces without changing their homology classes. Then the blowing up formula of Friedman and Morgan [7] gives nonzero Donaldson invariants on the blow up of the K3 surface or its log transforms, where higher dimensional moduli spaces are used.

We conjecture the following: Let M be a simply-connected, smooth, closed and oriented four-manifold of even type. Suppose an $SO(3)$ bundle over M has nonvanishing Donaldson invariant on $H_2(S;\mathbb{Z})$ as above. Then for any generic marking, the corresponding embedded Casson handles will not be all of bounded type.

Our basic analytic tool is Yang–Mills gauge theory. On the other hand, Seiberg–Witten monopoles have contributed much to recent research in low dimensional topology and the study of smooth 4–manifolds, by the topological methods of Ozsváth and Szabó. It would be natural to try to construct a parallel method by SW monopoles to the content of this paper. For the moment we have already encountered some difficulty in the framework of SW monopoles, to obtain Riemannian metrics and weight functions on open four-manifolds which should give Fredholm linearized complexes. In fact we have used de Rham cohomology computation to obtain the conclusion in Theorem 1.6, where linearized Yang–Mills equation involves Laplace operators.

Acknowledgements

The author would like to thank the referees for numerous suggestions and comments. He also would like to thank to Professor M Furuta for valuable comments.

2 ASD moduli spaces

2.1 Bundles over noncompact spaces

Let (S,g) be a complete Riemannian open four-manifold of bounded geometry. Bounded geometry requires two properties on Riemannian manifolds. One is uniformity of injectivity radius from below $\epsilon_x > c > 0$ independently of $x \in S$. The other is that there are constants $C_k \geq 0$ and local trivializations on some neighbourhoods of any points $x \in S$ so that the corresponding connection matrices (ω^i_j) satisfy uniformity $\|(\omega^i_j)\|_{C^k} \leq C_k$ for all $k = 0, 1, \ldots$, independently of $x \in S$. We denote the Sobolev k norms as $W^k(S)$ which are given by $\|u\|^2_{W^k(S)} = \sum_{i \leq k} \|\nabla^i(u)\|^2_{L^2(S)}$.

We always assume that both the end of S and S itself are simply-connected. For our application later, it is enough to assume that the end of S is homeomorphic to $(1,\infty) \times S^3$.

Let $E \to S$ be an $SO(3)$ vector bundle over S, and denote the corresponding frame bundle by P. Let us assume that E admits an L^2 ASD connection A with

$$F_A + *F_A = 0 \quad \text{and} \quad \|F_A\|_{L^2(S,g)} < \infty.$$

Let us fix a small $\mu_0 > 0$, and $U \subset S$ be the μ_0 neighbourhood of $x \in S$ so that E admits a trivialization over U, $E|U \cong U \times \mathbb{R}^3$. Since (S, g) is of bounded geometry, the following Lemma holds.

Lemma 2.1 (Uhlenbeck [16; 17]) *There are constants $\epsilon > 0$ and C_k, $k = 0, 1, \dots$, independent of $x \in S$, such that if the curvature F_A satisfies $\|F_A\|_{L^2(U)} \leq \epsilon$, then there is a gauge transformation g over $E|U$ so that the following estimates hold:*

$$\|g^*(A)\|_{W^k(U)} \leq C_k \|F_A\|_{L^2(U)}$$

By the Sobolev embeddings, there are estimates $\|g^*(A)\|_{C^k(U)} \leq C'_k \|F_A\|_{L^2(U)}$, where C'_k again are independent of $x \in S$. Let $\mathrm{Ad}\, P$ be the adjoint bundle.

Proposition 2.2 *For a large compact subset $K \subset S$, there is a bundle trivialization of E over $S \backslash K$ so that $A = d + u$ satisfies the estimates $\|u\|_{W^k(S \backslash K)} < \infty$ on $S \backslash K$ for all k, where u is a section of $\mathrm{Ad}\, P \otimes T^*S|S \backslash K$.*

Proof Let us choose a sufficiently large $K \subset S$ so that $\|F_A\|_{L^2(S \backslash K)} \leq \epsilon$ holds. Let $x_0 \in S \backslash K$ be a fixed point and $U \subset S \backslash K$ be a small neighbourhood of x_0. Let us choose a trivialization of E over U by use of the parallel transform. Then Lemma 2.1 tells us that such a bundle trivialization gives the estimates $\|u\|_{W^k(U)} \leq C_k \|F_A\|_{L^2(U)}$ for some constant C_k independent of x_0 and U since the parallel transform is defined by use of the connection matrix.

Let us choose another $x \in S \backslash K$. For any smooth loop l in $S \backslash K$ between x_0 and x, the parallel transform is denoted by τ_l. Let us choose two such loops l and l' between x_0 and x. Then there is some $g_{l,l'} \in \mathrm{Aut}\ E_x$ so that $\tau_l = g_{l,l'} \tau_{l'}$ holds. We claim that for any small $\delta > 0$, there is a sufficiently large K independent of x, such that $g_{l,l'}$ lies in a δ neighbourhood N of the identity in $\mathrm{Aut}\ E_x \cong SO(3)$ In particular N admits a contraction to the identity. This holds by Lemma 2.1 when both l and l' are contained inside a small neighbourhood of x_0. For the general case this also holds, since S is simply-connected at infinity as follows. Let $c = l^{-1} \circ l' \colon [0,1] \to S \backslash K$ be the closed loop and $H \colon [0,1] \times [0,1] \to S \backslash K$ be a contracting homotopy so that $H(0, \ ) = c$ and $H(1, \ ) \equiv x_0$. One can cover the image of H by small balls. Then successive applications of Lemma 2.1 on each ball again give the same conclusion, as desired.

The rest follows by a general argument. Let us choose a triangulation on $S \backslash K$ covered by sufficiently small simplices. Take any geodesics from x_0 to vertices. They give a trivialization of E on the set of vertices by use of parallel transport. Let Δ be any simplex and take a vertex $x \in \Delta$. Again by use of parallel transport from x, one obtains a trivialization of E on Δ. For any other vertices $y \in \Delta$, this trivialization on E_y will be certainly different from the previous one which used parallel transport from x_0. However the difference is small, where by the above argument, there is some $g \in \text{Aut } E_y$ which gives difference of these two trivializations, and it lies within a δ neighbourhood of the identity as above. Thus by this property, one can extend these local trivializations on simplices so that they give a global trivialization of E on $S \backslash K$. Moreover such a bundle trivialization gives the estimate $\|u\|_{W^k(U)} < C_k \|F_A\|_{L^2(U)}$ for any small open subset $U \subset S \backslash K$, since it is constructed by extending local trivializations by small perturbations of parallel transforms.

This gives a trivialization $E|S \backslash K \cong (S \backslash K) \times \mathbb{R}^3$ and ensures that the desired estimate $\|u\|_{W^k(S \backslash K)} < \infty$ on $S \backslash K$ for all k. □

Later on we fix a trivialization of E over $S \backslash K$. E is determined by $w_2(E) \in H^2(S; \mathbb{Z}_2)$ and $p_1(E)$.

2.2 ASD moduli spaces

Let us denote the weighted Sobolev k spaces by $W^k_w(S, g)$, where their norms are defined by

$$\|u\|^2_{W^k_w(S,g)} = \sum_{i \leq k} \int_S \exp(w) |\nabla^i u|^2 \text{vol} .$$

Let A be as in Section 2.1, and define an affine Hilbert space as:

$$\mathfrak{A}_k(A) = \{A + a | a \in W^k_w(S; \text{Ad } P \otimes T^* S)\}, \quad k \geq 3.$$

Let us take $g \in C^1_{\text{loc}}(S; \text{Aut } E)$. By embedding as $g \in C^1_{\text{loc}}(S; \text{Hom}(E, E))$, one may consider $\nabla_A g \in C^0(S; Hom(E, E) \otimes T^* S)$. Notice that if g is locally W^4, then it is of C^1 class. Now one defines the weighted Sobolev gauge groups ($l \geq 4$):

$$\mathfrak{G}_l(P) = \{h \in W^l_{\text{loc}}(S; \text{Aut}(E)) | \nabla_A h \in W^{l-1}_w(S; \text{Hom}(E, E) \otimes T^* S)\},$$
$$\mathfrak{G}_l(P)_0 = \{h \in W^l_{\text{loc}}(S; \text{Aut}(E)) | h - \text{id} \in W^l_w(S; \text{Hom}(E, E) \otimes T^* S)\}.$$

The Lie algebras of $\mathfrak{G}_l(P)$ and $\mathfrak{G}_l(P)_0$ are correspondingly as follows:

$$\mathfrak{g}_l(P) = \{h \in W^l_{\text{loc}}(S; \text{Ad}(P)) | \nabla_A h \in W^{l-1}_w(S; \text{Ad}(P) \otimes T^* S)\},$$
$$\mathfrak{g}_l(P)_0 = W^l_w(S; \text{Ad}(P)).$$

These spaces all admit structure of Banach manifolds. $\mathfrak{G}_{k+1}(P)_0$ acts on $\mathfrak{A}_k(P)_0$ by $g^*(A + a) = g^{-1} \nabla_A g + g^{-1} a g$. Notice that $\mathfrak{G}_{k+1}(P)$ does not act on $\mathfrak{A}_k(P)_0$, since A is only in $L^2(S, g)$.

Let us put

$$\widehat{\mathfrak{M}}_k(A)_w = \{A' \in A + W^k_w((S,g); T^*S \otimes \mathrm{Ad}\, P) : F^+(A') = 0\}.$$

For any $A' \in \widehat{\mathfrak{M}}_k(A)_w$, one has the following AHS complex with coefficient A':

$$0 \longrightarrow W^{k+1}_w((S,g); \Lambda^0 \otimes \mathrm{Ad}\, P) \xrightarrow{d_{A'}} W^k_w((S,g); \Lambda^1 \otimes \mathrm{Ad}\, P) \\ \xrightarrow{d^+_{A'}} W^{k-1}_w((S,g); \Lambda^2_+ \otimes \mathrm{Ad}\, P) \longrightarrow 0$$

Proposition 2.3

(1) $\mathfrak{G}_{k+1}(P)_0$ *acts on both* $\hat{\mathfrak{M}}_k(A)_w$ *and* $\mathfrak{A}_k(A)$ *freely.*

(2) *If the AHS complex without coefficient is Fredholm, then so is the above complex.*

(3) *Suppose the above complex is Fredholm, and* $d^+_{A'}\colon W^k_w(S;\mathrm{Ad}\,P\otimes T^*S)\mapsto W^{k-1}_w(S;\mathrm{Ad}\,P\otimes\Lambda^2_+)$ *is a surjection for* $A'\in\hat{\mathfrak{M}}_k(A)_w$. *Then* $\mathfrak{M}_k(A)_w=\hat{\mathfrak{M}}_k(A)_w/\mathfrak{G}_{k+1}(P)_0$ *is a finite dimensional smooth manifold at* $[A']$. *Its dimension dim* $\ker d^+_{A'}/\operatorname{im} d_{A'}$ *is equal to* $-2p_1(P)+3(\dim H^1(AHS)-\dim H^2(AHS))$.

Proof For details of the proof, see [10, 1.B]. □

In the case when S is an open four-manifold with tree-like ends whose trees are homogeneous of bounded type, then by Theorem 1.6 S admits a pair (g,w) so that the assumption in (2) is satisfied, and the dimension in (3) is less than $-2p_1(P)-3b^2_+(S)$.

The quotient space

$$\mathfrak{M}_k(A)_w=\hat{\mathfrak{M}}_k(A)_w/\mathfrak{G}_{k+1}(P)_0$$

is the (weighted) *Yang–Mills moduli space*, which plays the central role throughout the paper.

Now let M be an oriented, closed and smooth four-manifold with $\pi_1(M)=1$. Suppose the form is of even type with $b_2^+(M)=l$. Let us take an $SO(3)$ bundle E over M with $-p_1(E)=s$. Then the dimension of the generic ASD moduli space $\mathfrak{M}(E)$ is $d\equiv 2s-3(1+l)$.

Let us take a generic marking with respect to the $SO(3)$ bundle $E\to M$, where both $w_2^1(E), w_2^2(E)\neq 0$ do not vanish. Let $S\subset M$ be the corresponding open four-manifold with tree-like ends, and assume that the trees are homogeneous of bounded type. Let us choose an admissible pair (g,w) by Theorem 1.6, and suppose an L^2 ASD connection A over $E_1\to(S,g)$ is obtained as in Section 1, by choosing exhaustion $K_0\subset K_1\subset\cdots\subset S\subset M$.

Lemma 2.4 *The formal dimension* D *of* $\mathfrak{M}(A)_w$ *satisfies the inequality*

$$D\leq d-4+3=d-1.$$

Proof By Proposition 2.3(3) and the remark following it, the inequality $D\leq d+3+2p_1(E)-2p_1(E_1)$ holds, where the 3 comes from the constant gauge transformation. Let us see that $-p_1(E_1)$ is strictly smaller than $s=-p_1(E)$. In fact A is obtained as a limit of a sequence of ASD connections (A_i, g_i) restricted on S, where the convergence is in C^∞ on each compact subset. Moreover the Riemannian metrics g_i on M converge to g in C^∞ on each compact subset in S. By the ASD condition, the equalities

$$-p_1(E) = \frac{1}{4\pi^2}\int_M tr(F_{A_i} \wedge F_{A_i}) = \frac{1}{4\pi^2}\int_M |F_{A_i}|^2$$

hold, and so the inequalities

$$-p_1(E_1) = \frac{1}{4\pi^2}\int_S |F_A|^2 \leq \frac{1}{4\pi^2}\int_M |F_{A_i}|^2 = -p_1(E)$$

hold too.

Now suppose that equality holds in the above inequality. With respect to a topological decomposition $M \cong l(S^2 \times S^2)\#k|-E_8|$, the bundle E also splits topologically as $E_1 \# E_2$ by Proposition 2.2 for all large i. Then E_2 must be flat, in fact trivial, since $l(S^2 \times S^2)$ is simply-connected. However since w_2 does not change under such a limit, this cannot happen. This implies that the inequality must be strict.

Since the form is even, $p_1(E_1)$ takes an even value. □

In particular if $d = 0$, then D is negative and so the generic ASD moduli space $\mathfrak{M}(A)_w$ should be empty. This will be a crucial point in our argument of the proof of the main theorem.

2.2.1 Proof of Theorem 1.9 under a favorable case Even though the complete proof of Theorem 1.9 contains some complicated analysis which occupies the whole of this paper, the basic idea is relatively simple. Here we will describe it assuming some analytically favorable situation.

Let M be as above in Section 2.2, equipped with a generic marking Φ. Suppose the corresponding Casson handles are all of bounded type. Then there is an open subset $S \subset M$ homeomorphic to $l(S^2 \times S^2) \setminus \mathrm{pt}$, and by Theorem 1.6 there is an admissible pair (g, w) on S. Let $E \to M$ be an $SO(3)$ bundle and suppose (M, E, Φ) could consist of a nondegenerate triplet. We have a contradiction under some analytic condition as below. For simplicity of the argument, we consider the case without embedded surfaces. Thus the $SO(3)$ Donaldson invariant $Q(E)$ is nonzero.

By definition, Φ is generic so that $w_2^1, w_2^2 \neq 0$ do not vanish with respect to the decomposition $w_2 = w_2^1 \oplus w_2^2$ as in Section 1. Let us take exhausting compact subsets $K_0 \subset K_1 \subset \cdots \subset S$ and a family of generic Riemannian metrics $\{g_i\}_i$ so that $\big|| g_i | K_i - g | K_i |\big|_{C^l} \to 0$ as $i \to \infty$ for all $l = 0, 1, \ldots$. Choose a sequence of ASD connections $A_i \in \mathfrak{M}(E)$ over (M, g_i). After gauge transformations, a subsequence converges to an ASD connection A over $(E', (S, g))$ on each compact subset in C^∞.

Now by use of the triple (g, w, A) on S, one obtains the weighted Yang–Mills moduli space $\mathfrak{M}_k(A)_w$ for a large k, which is nonempty since $A \in \mathfrak{M}_k(A)_w$.

Suppose $d_A^+ \colon W_w^k(S; \mathrm{Ad}\, P \otimes T^*S) \mapsto W_w^{k-1}(S; \mathrm{Ad}\, P \otimes \Lambda_+^2)$ is a surjection. Then by Proposition 2.3(3), it is a finite dimensional smooth manifold. By the assumption, the dimension of the moduli space $\mathfrak{M}(E)$ over M is zero. Then by Lemma 2.4, the formal dimension of $\mathfrak{M}(A)_w$ becomes negative. It follows from surjectivity of the above differential that the dimension of $\mathfrak{M}(A)_w$ should coincide with the formal one. Since it is nonempty and the formal dimension is negative, this is a contradiction.

Remark 2.5 The main point in considering the general case is that one cannot assume surjectivity of the differential as above. Thus one has to perturb the ASD equation over S. By following the above procedure by use of a perturbed one, one will choose another family of connections on M. The most important analytic step in this paper is to control the sizes of Sobolev norms of differences of these families of connections. At present if we allow bubbling phenomena on S in Definition 1.8(3), then we will lose control of such sizes between connections, and will encounter difficulties in obtaining surjective differentials. In the above proof of the easiest case, we do not have to care about bubbling, since we have assumed surjectivity from the beginning.

2.3 Holonomy perturbation

Perturbation of equations often uses Banach spaces. In our analytic setting, we need technically to use Hilbert spaces B as perturbation spaces. In this paper we use a variant of the Floer holonomy perturbation, where our space B contains the Floer Banach space P densely. For our construction of perturbed connections this extension is auxiliary, and finally we will choose generic paths inside P.

Let us take a fixed point $p_0 \in K_0 \subset S \subset M$ and fix a trivialization of $E|U$ on some small neighbourhood U of p_0. Let $(l_1, \dots, l_K)$ be a family of smooth loops in K_0 with $l_i(0) = p_0$. For any $SO(3)$ connection A on M or on S let

$$\mathrm{Hol}(A) = (\mathrm{Hol}_{l_1}(A), \dots, \mathrm{Hol}_{l_K}(A)) \in SO(3)^K$$

be the set of holonomy around these loops.

Lemma 2.6 *There is a large K and a family of smooth loops $\{l_1, \dots, l_K\}$ so that for all large $i \geq i_0$, the corresponding holonomy maps*

$$\mathrm{Hol}(A_i), \mathrm{Hol}(A) \in SO(3)^K$$

have no isotropy subgroups with respect to the diagonally adjoint $SO(3)$ action respectively.

The same properties also hold near $[A_i]$ and $[A]$.

Proof By modifying Hol slightly still depending only on the restriction of connections to K_1, we construct the desired map as below. There is a small $\alpha \in so(3)$ with $|\alpha| << 1$ so that $\mathrm{Hol}'(A) \equiv (\mathrm{Hol}(A), \exp(\alpha)) \in SO(3)^K$ has no isotropy subgroup under $SO(3)$ adjoint action. We extend it to $\mathrm{Hol}'\colon \mathfrak{B}_{k+1}(A) \to SO(3)^K/SO(3)$ so that $\mathrm{Hol}'(A')$ still depends only on $A'|K_1$ as follows. Let $U \subset W_w^{k+1}$ be a small open subset with the local slice for the gauge group action $A+U \cong V \times W$, where $V \subset U$ and $W \subset \mathfrak{G}_{k+2}(P)_0$, and diam $V < \epsilon_0$. Choose two functions, $\varphi \in C_c^\infty(K_1)$ with $\varphi|K_0 \equiv 1$, and a decreasing function $f\colon [0,\epsilon_0] \to [0,1]$ with $f(0) = 1$ and $f(\epsilon_0) = 0$. Then for $v \in V$ and $A' = A+v$, we put $\mathrm{Hol}'(A') = (\mathrm{Hol}(A'), \exp(f(||\varphi v||)\alpha))$. One can extend Hol' equivariantly on $\mathfrak{A}_{k+1}(A)$, since gauge group action is free by Proposition 2.3(1). Hol' depends only on $A'|K_1$ and the restricted action $\mathfrak{G}_{k+2}(P)_0|K_1$, which is the desired map. The A_i case is similar. This completes the proof. □

Let us extend these embedded loops as $l_i\colon D^3 \times S^1 \hookrightarrow M$, and take a smooth compactly supported function $\psi \in C^\infty(D^3)_c$ with $\int_{D^3} \psi = 1$. Then we extend the holonomy maps to $\mathrm{Hol}'_i(A_i) \in SO(3)$ which are small perturbations of:

$$\int_{D^3} \psi(x)\, \mathrm{Hol}_{l_i(x,\ )}(A_i).$$

By this way one obtains a smooth map

$$\mathfrak{B}(M,E) = \mathfrak{A}(E)/\mathfrak{G}(E) \to SO(3)^K/SO(3)$$

which is an embedding on compact subsets near $[A_i]$ and $[A]$ for all large i.

Floer constructed a Banach space P as a perturbation space by use of a subspace of C^∞ functions $C^{\overline{\beta}}(U) \subset C^\infty(U)_0$. In this paper we use a Sobolev space B as the perturbation space, which contains P.

When one chooses $b \in B \backslash P$, then the corresponding perturbed moduli spaces will be C^l manifolds, rather than C^∞, for some l. This will still work sufficiently for our applications. But here in fact we see that at last one can take $b \in P$ and so the corresponding moduli spaces are smooth, since $P \subset B$ is dense.

Let us choose a sufficiently large $m \gg k$. For any $u \in C_c^0(U)^9$ and a connection A, let us assign

$$\Psi(u,A) \in C^0(M; \Lambda_+^2 \otimes \mathrm{Ad}\, P)$$

by identifying u in this space by use of A. Let us identify $\Lambda^2_+\mathbb{R}^4\otimes so(3)\cong\mathbb{R}^9$ and let

$$\gamma = SO(3)^K\times_{SO(3)}\mathbb{R}^9$$

be the vector bundle over $SO(3)^K/SO(3)$. Then we define B by

$$B = W^m(\gamma\times U)_0.$$

This embeds into $C^l\big(SO(3)^K\times\bar{U};\mathbb{R}^9\big)^{SO(3)}$ by the Sobolev embedding for some l, where the perturbation space P is a dense subspace of $C_0^\infty\big(SO(3)^K\times U;\mathbb{R}^9\big)^{SO(3)}$. In particular P is dense in our space B.

The map Ψ induces a perturbation $s_0\colon \mathfrak{B}_{k+1}(E)\times B\to W^k(M;\Lambda^2_+\otimes\mathrm{Ad}\,P)$ by

$$s_0\big([A],\sum_i u_i\otimes v_i\big) = \Psi\big(\sum_i u_i(\mathrm{Hol}\,A)\otimes v_i, A\big).$$

Let us fix a small $\epsilon_0>0$, and put

$$s=\epsilon_0 s_0\colon \mathfrak{B}_{k+1}(E)\times B\to W^k(M;\Lambda^2_+\otimes\mathrm{Ad}\,P).$$

Let us consider the perturbed map

$$F_s^+ = F^+ + s\colon \mathfrak{A}_{k+1}(E)\times B\to W^k(M;\Lambda^2_+\otimes\mathrm{Ad}\,P).$$

Let g_i be a Riemannian metric on M. For any choice $b\in B$, we denote the perturbed ASD moduli space by

$$\mathfrak{M}_b(M,g_i) = \{(A,b) : F^+(A)+s(A,b)=0\}/\mathfrak{G}_{k+2}(E).$$

Recall $\mathfrak{A}_{k+1}(A) = A + W_w^{k+1}((S,g);\Lambda^1\otimes\mathrm{Ad}\,P)$. Since the holonomy perturbation is local, it also induces another one:

$$F_s^+ = F^+ + s\colon \mathfrak{A}_{k+1}(A)\times B\to W_w^k((S,g);\Lambda^2_+\otimes\mathrm{Ad}\,P)$$

Let us denote its differential

$$D_A^+\colon W_w^{k+1}((S,g):\Lambda^1\otimes\mathrm{Ad}\,P)\oplus T_0B\to W_w^k((S,g):\Lambda^2_+\otimes\mathrm{Ad}\,P).$$

Proposition 2.7 *The linearization map*

$$D_A^+\equiv d_A^+ + ds\colon W_w^{k+1}((S,g);\Lambda^1\otimes\mathrm{Ad}\,P)\oplus T_0B\to W_w^k((S,g);\Lambda^2_+\otimes\mathrm{Ad}\,P)$$

is surjective.

The essential point for this is that the restrictions on U of any nonzero elements in the cokernel are still nonzero, which follows from unique continuity for solutions of elliptic operators $(d_A^+)^*_w(u) = 0$. Once one has this property, then it is immediate to see surjectivity: suppose not and choose a nonzero element $u \in \ker(D_A^+)^*_{w_i}$. As above, the restriction $u|U$ on the open subset is still nonzero. Thus one can choose some $v \in C_c^\infty(U)$ so that the $L^2_{w_i}$ inner product with u is still nonzero $\langle u, v\rangle_{L^2_{w_i}} \neq 0$. By Lemma 2.6, there are no isotropy subgroups for $\mathrm{Hol}(A) \in SO(3)^K$, and so one can choose some $\psi \in P$ so that $\psi(\mathrm{Hol}\, A, \ ) = v$. In particular $\langle ds_{(A,0)}(0,\psi), u\rangle_{L^2_{w_i}} \neq 0$ which cannot happen, since u lies in the cokernel of D_A^+.

Let us put

$$V = \mathrm{im}\, d_A, \quad W = V^\perp \subset W_w^{k+1}((S,g); \Lambda^1 \otimes \mathrm{Ad}\, P)$$

where $V^\perp$ is the $W_w^{k+1}(S,g)$ orthogonal complement. Let us consider the restriction

$$D_A^+ \colon W \oplus T_0 B \to W_w^k((S,g); \Lambda_+^2 \otimes \mathrm{Ad}\, P).$$

The same proof of the above gives the following Corollary.

Corollary 2.8 *The above restriction is still surjective.*

Later on we choose and fix a small $\epsilon > 0$ so that the ϵ ball $D_\epsilon \subset T_0 B$ is identified with an open subset in B near zero.

By the Sard–Smale transversality, there exist positive $\delta_1, \delta_2 > 0$ so that for the δ_i balls $D_{\delta_1} \subset W_w^{k+1}(S,g)$ and $D_{\delta_2} \subset B$, one obtains the following Proposition.

Proposition 2.9 *There is a Baire set $B(S,g) \subset D_{\delta_2}$ such that for any $b \in B(S,g)$ the perturbed moduli space $\mathfrak{M}_b(S,g)_w \cap (A + D_{\delta_1})/\mathfrak{G}$ is a finite dimensional manifold.*

One obtains similar statements by replacing (S,g) by (M,g_i).

Remark 2.10 In general for small perturbation, not necessarily compact, the conclusion of Proposition 2.9 still holds if the surjectivity condition in Proposition 2.7 holds and the operator norm $\|ds_{(A,0)}\|$ is sufficiently small.

For our case of holonomy perturbation, there is a small constant $c' > 0$ so that the operator norm $\|ds_{(A,0)}|T_0 B\| < c'$ is bounded by c'. Since s is compact, the spectral decomposition $\{(\psi_i, \lambda_i)\}_i$ of $(ds_{(A,0)}|T_0 B)^* ds_{(A,0)}|T_0 B$ for nonzero eigenvalues, shows that $\lambda_i \to 0$ goes to zero, and so the sets

$$L(s,c) \equiv \{v \in T_0 B : c\|v\| < \|ds_{(A,0)}(v)\|\}$$

are finite dimensional for any positive $c > 0$. Let us fix a small $c > 0$. Then one can modify s so that the result s' still satisfies $\|ds'_{(A,0)}|T_0 B\| < c'$, but $L(s', c)$ is infinite dimensional. In fact one can modify these perturbations on (M, g_j) so that the results s_j satisfy the estimates $\|d(s_j)_{(A,0)}(\psi_l)\| \geq c\|\psi_l\|$ for $l \leq j$ and $s_j(A, \psi_k) = s(A, \psi_k)$ for $k \geq j+1$. The s_j are still compact and $\lim_{j\to\infty} s_j = s'$ holds. By the above remark, if one replaces s with s', then Proposition 2.7 and Proposition 2.9 still hold on (S, g). We will use these modifications of perturbations in Section 3.2.8.

Since $P \subset B$ is dense, for any $b \in B$ and small $\epsilon > 0$, there is some $b' \in P$ so that $\|b - b'\|_B < \epsilon$. Since surjectivity is an open condition, one can choose generic elements $b \in P$. Later we will use this property.

Now let g_i be a generic Riemannian metric on M. When one uses a smooth perturbation, then the smooth transversality follows by Floer [4]:

Lemma 2.11 (Floer [4]) *There is a Baire set $P(M, g_i) \subset P$, so that the perturbed moduli space $\mathfrak{M}_b(M, g_i)$ is a smooth manifold of finite dimension for any $b \in P(M, g_i)$. Moreover if g_i is already generic, then there is a smooth cobordism between $\mathfrak{M}(E, g_i)$ and $\mathfrak{M}_b(E, g_i)$.*

Let g_i be as in Section 1. Now the countable intersection of Baire sets is dense. Thus by putting $P' = \bigcap_i P(M, g_i) \subset P$, one obtains a dense subset in P so that for any $b \in P'$, the perturbed moduli spaces $\mathfrak{M}_b(M, g_i)$ and $\mathfrak{M}_b(A)_w$ are all finite dimensional manifolds if nonempty. Moreover in that case, there are cobordisms between $\mathfrak{M}(M, g_i)$ and $\mathfrak{M}_b(M, g_i)$ by a generic smooth path b_t.

For our application, we need families of moduli spaces by two dimensional parameterization in Section 3.2.10. Let $c\colon [0,1]^2 \to P$ be a smooth map and put $\mathfrak{M}_c(M, g_j) = \{([A], c(s,t)) : F_s^+([A], c(s,t)) = 0, 0 \leq s, t \leq 1\}$.

Let b_t and b'_t be two generic paths. The same argument as that with cobordisms gives the following Lemma.

Lemma 2.12 *Let $c\colon [0,1]^2 \to P$ be a smooth map with $c(0,t) = b_t$ and $c(1,t) = b'_t$. Then for a Baire set of paths c as above, $\mathfrak{M}_c(M, g_j)$ gives smooth finite dimensional manifolds with corners.*

3 Proof of Theorem 1.9

Perturbing Riemannian metrics produces families of moduli spaces. In order to analyse these families of spaces for our purpose, one has to study estimates of injectivity radii of regular moduli spaces. Such estimates are given by the infinite dimensional implicit function theorem, which we will recall for convenience.

Let V, W, G be Banach spaces, $A = B(x_0)_{2\delta_1} \times B(y_0)_{2\delta_2}$ be the open subset of $V \times W$ and $f\colon A \to G$ be a C^∞ map, where $B(x_0)_{\delta_1}$ is the δ_1 ball with the center x_0, and $B(y_0)_{\delta_2}$ is similar. Suppose at $(x_0, y_0) \in A$, $f(x_0, y_0) = 0$ and $D_2 f(x_0, y_0)\colon W \cong G$ is the linear isomorphism, where D_2 is the derivative with respect to the second variable. Then there are positive $\epsilon_1, \epsilon_2 > 0$, and a C^∞ map $g\colon B(x_0)_{\epsilon_1} \to B(y_0)_{\epsilon_2}$ such that $g(x_0) = y_0$, and any $(x, y) \in B(x_0)_{\epsilon_1} \times B(y_0)_{\epsilon_2}$ satisfies $f(x, y) = 0$ if and only if $y = g(x)$. Moreover $Dg(x) = -[D_2 f(x, g(x))]^{-1} Df(x, g(x))$ holds.

Lemma 3.1 *The positive constants ϵ_1, ϵ_2 depend only on the constants C_1, C_2 and C_3 of the norms*

$$C_1 \leq \|D_2 f(x, y)\|_{\inf}, \quad \|Df(x, y)\| \leq C_2, \quad \|D_2^2 f(x, y)\| \leq C_3$$

for all $(x, y) \in B(x_0)_{\delta_1} \times B(y_0)_{\delta_2}$, where $\|D_2 f(x, y)\|_{\inf} = \inf_{a \neq 0} \frac{\|D_2 f(x,y)(a)\|}{\|a\|}$.

Later we need uniformity of these ϵ_1 and ϵ_2 from below, when one constructs families of moduli spaces under the process of perturbing the metrics.

Let E be the $SO(3)$ bundle over M, and choose a generic marking Φ with respect to E. Below we assume that the triplet (M, E, Φ) is nondegenerate, and the corresponding $S \subset M$ has an admissible triple (S, g, w). Under these two conditions, we will induce a contradiction. Our final step of the proof is to construct a regular and perturbed moduli space whose formal dimension is negative.

We verify Theorem 1.9 when the generic ASD moduli space $\mathfrak{M}(E)$ is 0–dimensional. If it has positive dimension then by taking embedded surfaces $\Sigma_1, \dots, \Sigma_n \subset M$ one can consider the restricted 0–dimensional moduli space $\mathfrak{M}(E; \Sigma_1, \dots, \Sigma_n)$ as the construction of the Donaldson invariant. The argument is parallel for this case.

Let us choose an exhaustion by compact subsets $K_0 \subset K_1 \subset \cdots \subset S$. Choose a generic family of Riemannian metrics $\{g_i\}_i$ and ASD connections A_i with respect to (M, g_i) so that $\big||g_i|K_i - g|K_i|\big|_{C^k} \to 0$ for all $k = 0, 1, \dots$. One may assume after gauge transformation and by taking subsequence if necessarily, that A_i converge to an $L^2(S, g)$ ASD connection A over $E' \to S$ in C^∞ on each compact subset. By the assumption, bubbling does not occur on the family $\{A_i\}_i$ restricted on S. A priori the weighted Yang–Mills moduli space $\mathfrak{M}(A)_w$ will not be regular, and so its dimension may not coincide with the formal one.

Let us choose a generic element $b \in B$ with respect to all the family $\{A_i\}_i$ and A. We choose it so that its norm $\|b\|_B$ is small. We seek for a pair (A', b) with $F_s^+(A', b) = 0$

and $A - A' \in W^{k+1}_w((S,g);\Lambda^1 \otimes \mathrm{Ad}\, P)$. Then one will obtain a nonempty, regular, weighted and perturbed ASD moduli space

$$\mathfrak{M}_b(A)_w$$

over S, whose dimension should be negative by Lemma 2.4. This is a contradiction, and is enough to verify Theorem 1.9.

In order to guarantee existence of elements of the moduli spaces under small perturbations, we assume that the Donaldson invariant is nonvanishing. Let $\mathfrak{M}(M,g_i)$ be a regular ASD moduli space.

Lemma 3.2 *There is an element $[A_i] \in \mathfrak{M}(M,g_i)$ and a parameterization b_t, $b_0 = 0$ and $b_1 = b$ of small norms $\|b_t\| \ll 1$ so that for any $t \in [0,1]$, there are solutions (A^t_i, b_t) with $F^+_s(A^t_i, b_t) = 0$ over $E \to M$ and $A^0_i = A_i$.*

Proof Since both g_i and b are generic and the Donaldson invariant does not vanish, there is a cobordism between the ASD moduli space $\mathfrak{M}(M,g_i)$ and its perturbed form $\mathfrak{M}_b(M,g_i)$ in $\mathfrak{B}(E) \times B$. From this, the result follows. □

Let us take a family of C^∞ weight functions

$$w_i \colon M \to [0,\infty)$$

so that $w_i|K_i = w|K_i$ hold, where w is the fixed one on S. Then using weighted Sobolev spaces $W^k_{w_i}((M,g_i);\Lambda^*)$ instead of the usual ones $W^k((M,g_i);\Lambda^*)$, one can obtain the same ASD moduli spaces over (M,g_i), since M is compact.

3.1 Test case

Here we verify Theorem 1.9 assuming some analytic conditions on operators. This will clarify what causes a complicated situation for the proof of the general case.

Let (M,E,Φ) be the nondegenerate triplet, and assume that the corresponding $S \subset M$ has an admissible triple (S,g,w). Let $K_0 \subset \cdots \subset K_i \subset \cdots \subset S \subset M$ be an exhaustion, $\{g_i\}_i$ be a generic family of Riemannian metrics, as in Definition 1.8, and choose ASD connections $[A_i] \in \mathfrak{M}(M,g_i)$ from the regular moduli spaces. By definition of nondegeneracy, this family of ASD connections converges to an ASD connection A over $E' \to S$ in C^∞ on each compact subset of S, by taking a subsequence if necessary.

3.1.1 Uniformity of spectral radii Let us consider the surjective differentials

$$d^+_{A_i}\colon W^{k+1}_{w_i}((M,g_i);\Lambda^1\otimes \mathrm{Ad}\,P)\to W^k_{w_i}((M,g_i);\Lambda^2_+\otimes \mathrm{Ad}\,P)$$

and let

$$(d^+_{A_i})^*_{w_i}\colon W^{k+1}_{w_i}((M,g_i);\Lambda^2_+\otimes \mathrm{Ad}\,P)\to W^k_{w_i}((M,g_i);\Lambda^1\otimes \mathrm{Ad}\,P)$$

be the $L^2_{w_i}$ weighted adjoint operator. In the application of the implicit function theorem, surjectivity of the differentials is used, which is equivalent to lower bounds of the weighted adjoint operators $\|(d^+_{A_i})^*_{w_i}\|_{\inf}\geq c_i>0$.

In the case where uniform estimates of spectral radii

$$\|(d^+_{A_i})^*_{w_i}\|_{\inf}\geq c>0 \quad (*)$$

hold, where c is independent of i, then the proof of Theorem 1.9 goes very simply as follows. In fact in this quite restrictive situation, the connection A itself also becomes regular and satisfies the estimates $\|(d^+_A)^*_w\|_{\inf}\geq c$ as below.

For any $i\leq j$, let $W^l_{w_j}(K_i)_0$ be the closure by $W^l_{w_j}$ norms of $C^\infty(K_i)_c$, the set of functions compactly supported inside K_i. Since $w_j|K_i=w_i|K_i$ holds,

$$W^{k+1}_{w_j}(K_i;\Lambda^2_+\otimes \mathrm{Ad}\,P)_0$$

is contained in the domain of $(d^+_{A_j})^*_{w_j}$. Let $\varphi_i\in C^\infty(K_i)_0$ be cut off functions satisfying $\varphi_i(K_{i-1})\equiv 1$. Let $w\in W^{k+1}_w(S;\Lambda^1\otimes \mathrm{Ad}\,P)$ be any element with $\|w\|_{W^{k+1}_w}=1$. Then for any small $0<\epsilon\ll c$, there is a sufficiently large i_0 so that

$$\|\varphi_{i_0}w\|_{W^{k+1}_w(K_{i_0};\Lambda^1\otimes\mathrm{Ad}\,P)_0}\geq 1-\epsilon.$$

Since the family $\{A_i\}_i$ converges to A as above, for any small $\delta>0$, there is another large $i_1\geq i_0$ so that $\|(d^+_{A_{i_1}})^*_{w_{i_1}}-(d^+_A)^*_w\|\leq\delta$ holds on $W^{k+1}_{w_{i_0}}(K_{i_0};\Lambda^2_+\otimes \mathrm{Ad}\,P)_0$.

Now we have the estimates

$$\begin{aligned}\|(d^+_A)^*_w(w)\|_{W^k_w}&\geq \|(d^+_A)^*_w(\varphi_{i_0}w)\|_{W^k_w}-\|(d^+_A)^*_w((1-\varphi_{i_0})w)\|_{W^k_w}\\&\geq \|(d^+_{A_{i_1}})^*_{w_{i_1}}(\varphi_{i_0}w)\|_{W^k_{w_{i_0}}}-\|((d^+_{A_{i_1}})^*_{w_{i_1}}-(d^+_A)^*_w)(\varphi_{i_0}w)\|_{W^k_w}\\&\quad-\|(d^+_A)^*_w((1-\varphi_{i_0})w)\|_{W^k_w}\\&\geq c\|\varphi_{i_0}w\|_{W^{k+1}_{w_{i_0}}}-\delta-\epsilon.\end{aligned}$$

In particular if both ϵ and δ are sufficiently small with respect to c, then one has the estimate

$$\|(d^+_A)^*_w(w)\|_{W^k_w}\geq \tfrac{c}{2}\|\varphi_{i_0}w\|_{W^{k+1}_{w_{i_0}}}\geq \tfrac{c}{3}\|w\|_{W^{k+1}_w}$$

as desired.

Since the differential d_A^+ becomes surjective, the rest of the proof of the theorem is reduced to the argument in Section 2.2.1.

Remark 3.3 The above analytic method of cut and paste type is used in [10] in many places and detailed estimates are given there. For example, see [10, Lemma 1.5, Lemma 5.2].

3.1.2 Uniformity of injectivity radii of moduli spaces Here we describe the basic direction of the proof of Theorem 1.9 again by assuming some analytic conditions. In particular we see how to use uniform sizes of local charts of the moduli spaces in the implicit function theorem.

Let us choose $[A_i] \in \mathfrak{M}(M, g_i)$ and parameterization b_t as in Lemma 3.2. Assume uniform estimates $0 < c \le \|(d_{A_i}^+)_{w_i}^*\| \le c'$ and $\|Dd_{A_i}^+\| \le c''$ where D is the abstract differentials as in the implicit function theorem above. By Lemma 3.1, there are positive constants $\epsilon_1, \epsilon_2 > 0$ such that if we choose smooth path b_t with sufficiently small norms $\|b_t\| \ll 1$ with respect to these constants ϵ_i, then for any $t \in [0, 1]$, the solutions

$$F_s^+(A_i^t, b_t) = 0$$

exist with respect to (M, g_i) inside

$$B(x_0)_{\epsilon_1} \times B(y_0)_{\epsilon_2} \subset W_{w_i}^{k+1}((M, g_i); \Lambda^1 \otimes \operatorname{Ad} P) \oplus B.$$

In particular uniform estimates

$$\|A_i - A_i^t\|_{W_{w_i}^{k+1}(M, g_i: \Lambda^1 \otimes \operatorname{Ad} P)} \le C$$

hold. Then a subsequence $\{A_{i_m}^1\}_m$ converges to some A' over (S, g) which satisfies the perturbed ASD solution $F_s^+(A, b) = 0$. Moreover the estimate

$$\|A - A'\|_{W_w^{k+1}(S; \Lambda^1 \otimes \operatorname{Ad} P)} \le C$$

holds. In particular one obtains an element in the regular moduli space $\mathfrak{M}_b(A)_w$. This will lead a contradiction as we have already explained.

In both Section 3.1.1 and Section 3.1.2, it is of course too much to expect such situations, and we will need a refined argument below. The basic reason why these spectral radii collapse is that connections A_i on the regions $K_i^c \subset M$ (which contain $k|-E_8|$ parts) will behave quite badly. A fundamental idea below is to remove $A_i|K_i^c$ parts by use of quotient function spaces. In Section 3.2, we follow the direction of the proof of Theorem 1.9 described in Section 3.1.2 by changing function spaces as above.

3.2 Proof of Theorem 1.9

Now let us treat the general case, where we no longer have control of the spectrums with respect to i.

Let $H \subset W^k_w((S,g) : \Lambda^2_+ \otimes \mathrm{Ad}\, P)$ be the finite dimensional subspace of cokernel of d^+_A. Later on, we will assume that H has positive dimension. Otherwise we have nothing to do with, since the moduli space $\mathfrak{M}(A)_w$ would be regular when $H = 0$ (see Section 2.2.1).

The idea is to verify that the spectrum as above are certainly uniformly bounded for particular directions in Hilbert spaces. Firstly let us describe what are the situations we encounter. When $b_t \in B$ is a generic path between 0 and b, then since the Donaldson invariant does not vanish, both moduli spaces $\mathfrak{M}(M, g_i)$ and $\mathfrak{M}_b(M, g_i)$ admit nonempty cobordisms. In particular there are continuous paths A^t_i satisfying the equations $F^+_s(A^t_i, b_t) = 0$ by Lemma 3.2. Roughly speaking, one has to estimate the norms $\|A_i - A^t_i\|_{W^{k+1}_{w_i}}(K_i)$ from above. Recall that one has obtained an L^2 ASD connection A over (S, g) by taking limit of a subsequence of $\{A_i\}_i$. The final aim is to verify that $A - \lim_{i\to\infty} A^1_i|S$ is bounded in $W^{k+1}_w(S, g; \Lambda^1 \otimes \mathrm{Ad}\, P)$ for suitable choices of paths b_t.

The main point here is that connections on the complements of K_i will affect for our analysis. We will use a method to eliminate them on such domains by use of quotient function spaces. Let A_i be an ASD connection on (M, g_i). Usual analysis in gauge theory uses function spaces like $W^k_{w_i}((M, g_i); \Lambda^* \otimes \mathrm{Ad}\, P)$. Below we use quotients of such function spaces as $\overline{W}^k_{w_j}(K_i) \equiv W^k_{w_j}(M, g_j)/W^k_{w_j}(K^c_i)_0$ for $j \ge i$ (see Section 3.2.1). We collect basic properties of these new spaces. In particular we see that the self dual curvature operator still induces functional on the quotient ones (see Lemma 3.5). In Section 3.2.2, a merit of use of these spaces is given by uniform estimates of differentials (see Corollary 3.11), which will not be satisfied for the usual function spaces. Here we use some comparing method among Sobolev spaces with different base spaces by cut and paste type argument. They are the basic estimates which we will use iteratively.

3.2.1 Quotient spaces Recall that the family of weight functions $\{w_i\}_i$ on M is fixed, which satisfy the equalities $w_i|K_i = w|K_i$. Notice that for $j \ge i$, the equalities $w_j|K_i = w_i|K_i$ hold. Later on we fix a family of cut off functions $\varphi_i \in C^\infty_c(K_{i+1})$ satisfying $\varphi_i|K_i \equiv 1$.

Let us fix i and $j \ge i$. Let $K^c_i \subset M$ be the complement of K_i. We denote $W^k_{w_j}((K^c_i, g_j); \Lambda^* \otimes \mathrm{Ad}\, P)_0 \subset W^k_{w_j}((M, g_j); \Lambda^* \otimes \mathrm{Ad}\, P)$ as the completion of $C^\infty_c(K^c_i; \Lambda^* \otimes \mathrm{Ad}\, P)$ by the Sobolev $W^k_{w_j}$ norms. Thus they are closed subspaces.

Let us decompose

$$\begin{aligned}&W^k_{w_j}((M,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)\\ =&W^k_{w_j}((K_i^c,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)_0^{\perp}\oplus W^k_{w_j}((K_i^c,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)_0\\ \equiv&W^k_{w_j}((K_i,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)\oplus W^k_{w_j}((K_i^c,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)_0\end{aligned}$$

with respect to the $W^k_{w_j}$ inner products. We denote the quotient spaces as

$$\begin{aligned}&\overline{W}^k_{w_j}((K_i,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)\\ =&W^k_{w_j}((M,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)/\,W^k_{w_j}((K_i^c,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)_0.\end{aligned}$$

On (S,g), one also has similar spaces for $j\geq i$:

$$W^k_w((S,g);\Lambda^*\otimes \mathrm{Ad}\,P)=W^k_w((K_i^c,g);\Lambda^*\otimes \mathrm{Ad}\,P)_0^{\perp}\oplus W^k_w((K_i^c,g);\Lambda^*\otimes \mathrm{Ad}\,P)_0$$
$$\overline{W}^k_w((K_i,g);\Lambda^*\otimes \mathrm{Ad}\,P)=W^k_w((S,g);\Lambda^*\otimes \mathrm{Ad}\,P)/\,W^k_w((K_i^c,g);\Lambda^*\otimes \mathrm{Ad}\,P)_0$$

We equip with the metrics on these quotient spaces by use of orthogonal decomposition. Thus $\overline{W}^*_w(K_i,g)$ and $\overline{W}^*_{w_j}(K_i,g_j)$ are the same spaces, but they are equipped with different metrics. The projections $\bar{\ }$: $W^*_{w_j}((M,g_j);\Lambda^*\otimes\mathrm{Ad}\,P)\to \overline{W}^*_{w_j}((K_i,g_j);\Lambda^*\otimes \mathrm{Ad}\,P)$ are all distance decreasing maps.

We show that $\overline{W}^k_{w_j}((K_i,g_j);\Lambda^*\otimes\mathrm{Ad}\,P)$ and $\overline{W}^k_w((K_i,g);\Lambda^*\otimes\mathrm{Ad}\,P)$ are uniformly equivalent as metric spaces. In fact they are asymptotically isometric in the following sense.

Lemma 3.4 *For any small $\epsilon>0$ there is a constant $N\gg 0$ independent of choice of i,j, so that for $j\geq i+N$, the metrics on them are $(1-\epsilon,1+\epsilon)$–equivalent:*

$$(1-\epsilon)\|\bar{u}\|_{\overline{W}^k_{w_j}(K_i,g_j)}\leq \|\bar{u}\|_{\overline{W}^k_w(K_i,g)}\leq (1+\epsilon)\|\bar{u}\|_{\overline{W}^k_{w_j}(K_i,g_j)}$$

hold for all u.

Proof Let $u\in W^k_{w_j}((K_i,g_j);\Lambda^*\otimes\mathrm{Ad}\,P)$ and $u'\in W^k_w((K_i,g);\Lambda^*\otimes\mathrm{Ad}\,P)$ be representatives of $\bar{u}$. Then for some $i\leq i_0\leq j-1$, the estimate:

$$\|u\|_{W^k_{w_j}(K_{i_0+1}\setminus K_{i_0},g_j)}\leq \epsilon\|u\|_{W^k_{w_j}(M,g_j)}$$

should hold, where $[1/\epsilon]=N$. Let $\varphi_{i_0}\in C_c^{\infty}(K_{i_0+1})$ be the cut off function at the beginning of Section 3.2.1. Then $\|\varphi_{i_0}u\|_{W^k_{w_j}(M,g_j)}\leq (1+C\epsilon)\|u\|_{W^k_{w_j}(M,g_j)}$ holds where C is independent of i,j and ϵ. At the same time one may regard

$\varphi_{i_0} u \in W^k_w(S, g)$. Since $\overline{\varphi_{i_0} u} = \overline{u}$ have the same quotient element, this implies the estimates:

$$\begin{aligned}\|\overline{u}\|_{\overline{W}^k_w(K_i,g)} &= \|u'\|_{W^k_w(S,g)} \le \|\varphi_{i_0} u\|_{W^k_w(S,g)} = \|\varphi_{i_0} u\|_{W^k_{w_j}(M,g_j)} \\ &\le (1+C\epsilon)\|u\|_{W^k_{w_j}(M,g_j)} \\ &= (1+C\epsilon)\|\overline{u}\|_{\overline{W}^k_{w_j}(K_i,g_j)}.\end{aligned}$$

By the same argument, the converse estimates $\|\overline{u}\|_{\overline{W}^k_{w_j}(K_i,g_j)} \le (1+C\epsilon)\|\overline{u}\|_{\overline{W}^k_w(K_i,g)}$ hold. Since ϵ can be chosen arbitrarily small by taking large $j \gg i$, this completes the proof. □

Since metrics are all uniformly equivalent, later on one can choose any of them.

Let us consider the perturbed self dual curvature operators

$$F^+_s \colon A_j + W^{k+1}_{w_j}((M, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus B \to W^k_{w_j}((M, g_j) : \Lambda^2_+ \otimes \mathrm{Ad}\, P)$$

by $F^+_s(A_j + \alpha, b) \equiv F^+(A_j + \alpha) + s(A_j + \alpha, b)$.

Lemma 3.5 *Suppose $F^+_s(A_j, b) = 0$ holds over (M, g_j). Then the above operator induces the functional*

$$\overline{F}^+_s(\ , b) \colon A_j + \overline{W}^{k+1}_{w_j}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus B \to \overline{W}^k_{w_j}((K_i, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P).$$

Proof It is enough to check that for $\alpha \in W^{k+1}_{w_j}((K^c_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P)_0$, the element $F^+_s(A_j + \alpha, b)$ lies in $W^k_{w_j}((K^c_i, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P)_0$. By the condition,

$$F^+_s(A_j + \alpha, b) = d^+_{A_j}(\alpha) + P_+(\alpha \wedge \alpha)$$

since the support of s lies in K_0, where P_+ is the projection to the self dual part. From the Sobolev inequality, the result follows. □

Let $^-$ be the projection onto $\overline{W}^*_{w_j}(K_i, g_j)$.

Corollary 3.6 *Let $\alpha \in W^{k+1}_{w_j}((M, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P)$. If $F^+_s(A_j + \alpha, b) = 0$ then still $\overline{F}^+_s(A_j + \overline{\alpha}, b) = 0$.*

Lemma 3.7 *Let us take an ASD connection A_j on (M, g_j) with $F^+(A_j) = 0$. Then the projection induces maps*

$$\begin{aligned} d_{\overline{A}_j} &\colon \overline{W}^{k+2}_{w_j}((K_i, g_j); \mathrm{Ad}\, P) \to \overline{W}^{k+1}_{w_j}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P), \\ d^+_{\overline{A}_j} &\colon \overline{W}^{k+1}_{w_j}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P) \to \overline{W}^k_{w_j}((K_i, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P).\end{aligned}$$

In particular, if $d^+_{A_j}$ is surjective, then $d^+_{A_j}$ is still the same.

Proof This follows since

$$d^+_{A_j}(W^{k+1}_{w_j}((K^c_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P)_0) \subset W^k_{w_j}((K^c_i, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P)_0. \qquad \square$$

Let us recall the three constants in Lemma 3.1 of the abstract implicit function theorem, where $f\colon A \to G$,

$$C_1 \le \|D_2 f(x,y)\|_{\inf}, \quad \|Df(x,y)\| \le C_2 \quad \text{and} \quad \|D_2^2 f(x,y)\| \le C_3.$$

We apply it to our operators.

Lemma 3.8 *There are constants C_2, C_3 such that for any i, there is a large $j_0 \ge i$ so that for all $j \ge j_0$, the differentials of*

$$\overline{F}^+_s\colon A_j + \overline{W}^{k+1}_{w_j}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P) \times B \to \overline{W}^k_{w_j}((K_i, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P)$$

satisfy uniform estimates $\|D\overline{F}^+_s\| \le C_2$ and $\|D^2\overline{F}^+_s\| \le C_3$ near $(A_j, 0)$.

A uniform lower bound from below by C_1 is given in Corollary 3.11 below.

Proof Since we have assumed that the sequence $\{A_i\}_i$ converges to A in C^∞ on each compact subset in S, it follows that for any i and $j \ge j_0$ as above, the $A_j|K_i$ are near $A|K_i$ in C^l for a large l. Thus the uniform estimates of the former follow. The latter also follows since the curvature operator involves only twice multiplication of the connection coefficients, and so the equality $D^2\overline{F}^+_s(\gamma, b)(\beta, \alpha) = D^2 s(\gamma, b)(\beta, \alpha) + P_+([\beta, \alpha])$ holds. $\square$

3.2.2 Comparison method on estimates Let us state a general functional analytic property for the linearized operator of the ASD equation. Let A_j be an ASD connection over (M, g_j) with $F^+(A_j) = 0$. Let us consider the differentials

$$d^+_{A_j}\colon W^{k+1}_{w_j}((M, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P) \to W^k_{w_j}((M, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P),$$

and their subspaces

$$\mathrm{im}\, d^+_{A_j} \subset W^k_{w_j}((M, g_j) : \Lambda^2_+ \otimes \mathrm{Ad}\, P)$$
$$\text{and} \qquad (\ker d^+_{A_j})^\perp \subset W^{k+1}_{w_j}((M, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P).$$

$d^+_{A_j}\colon (\ker d^+_{A_j})^\perp \cong \mathrm{im}\, d^+_{A_j}$ gives an isomorphism, and for all $u \in (\ker d^+_{A_j})^\perp$ the estimates $\|d^+_{A_j}(u)\|_{W^k_{w_j}} \ge C_j \|u\|_{W^{k+1}_{w_j}}$ hold for some positive constants $C_j > 0$.

Let

$$d^+_{A_j}\colon \overline{W}^{k+1}_{w_j}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P) \to \overline{W}^k_{w_j}((K_i, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P)$$

be the induced map, and consider the subspace $(\ker d^+_{A_j})^\perp$. Below we have a weak version of the estimates.

Lemma 3.9 *Suppose $d^+_{A_j}$ is surjective. Then for $\overline{u} \in (\ker d^+_{A_j})^\perp$, the estimate*

$$\| d^+_{A_j}(\overline{u}) \|_{\overline{W}^k_{w_j}} \geq C_j \|\overline{u}\|_{\overline{W}^{k+1}_{w_j}}$$

still holds, where C_j is the same constant as above.

Proof $d^+_{A_j}\colon (\ker d^+_{A_j})^\perp \cong \mathrm{im}\, d^+_{A_j}$ with the constant C_j above is equivalent to that the image of the unit ball contains the $\epsilon = C_j$ ball

$$d^+_{A_j}(B_1) \supset D_\epsilon$$

where $B_1 \subset (\ker d^+_{A_j})^\perp$ and $D_\epsilon \subset \mathrm{im}\, d^+_{A_j}$.

Let us take any $\overline{W} \in \overline{W}^k_{w_j}((K_i, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P)$, and choose the unique representative $w \in W^k_{w_j}((M, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P)$. Thus

$$\| \overline{W} \|_{\overline{W}^k_{w_j}} = \|w\|_{W^k_{w_j}}.$$

Let $\|w\|_{W^k_{w_j}} < \epsilon$. Then by the assumption, there exists $u \in W^{k+1}_{w_j}((M, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P)$ with $\|u\|_{W^{k+1}_{w_j}} < 1$ satisfying $d^+_{A_j}(u) = w$. Now decompose $u = u_1 + u_2$ where

$$u_1 \in W^{k+1}_{w_j}((K_i^c, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P)_0^\perp \quad \text{and} \quad u_2 \in W^{k+1}_{w_j}((K_i^c, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P)_0.$$

Since $d^+_{A_j}(u_2) \in W^k_{w_j}((K_i^c, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P)_0$, the equality $d^+_{A_j}(\overline{u}_1) = \overline{W}$ holds. Since $\|\overline{u}_1\|_{\overline{W}^{k+1}_{w_j}} \leq \|u\|_{W^{k+1}_{w_j}} < 1$ the result follows. □

If one follows the above proof, then the constants $C_j > 0$ above will a priori depend on i and j. Below we have a *comparison method* which induces a uniform lower bound of constants, where we will compare the K_i part of functions between M and S.

Let A be an L^2 ASD connection over (S, g), obtained as $\lim_i A_i$ as before. Recall that in Section 2.3 we have a perturbed map

$$F^+_s\colon \mathfrak{A}_{k+1}(A) \times B \to W^k_w((S, g) : \Lambda^2_+ \otimes \mathrm{Ad}\, P)$$

and its surjective differential

$$D^+_A\colon W^{k+1}_w((S, g) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B \to W^k_w((S, g) : \Lambda^2_+ \otimes \mathrm{Ad}\, P).$$

We put $U = (\ker D_A^+)^\perp \subset W_w^{k+1}((S,g);\Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B$. As above, there is a constant $C > 0$ so that $\|D_A^+(u)\|_{W_w^k(S,g)} \geq C\|u\|_{W_w^{k+1}(S,g)}$ hold for all $u \in U$.
Let $\overline{W}_w^*(K_i, g) = W_w^*(S,g)/W_w^*(K_i^c, g)_0$ be the quotient spaces, and

$$D_{\bar{A}}^+ \colon \overline{W}_w^{k+1}((S,g) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B \to \overline{W}_w^k((S,g) : \Lambda_+^2 \otimes \mathrm{Ad}\, P)$$

be the induced map.

Corollary 3.10 *There is a uniform bound*

$$\|D_{\bar{A}}^+(\bar{u})\|_{\overline{W}_w^k(K_i,g)} \geq C\|\bar{u}\|_{\overline{W}_w^{k+1}(K_i,g)}$$

for all $\bar{u} \in \bar{U} \equiv (\ker D_{\bar{A}}^+)^\perp$, where $C > 0$ is the same constant as above.

Proof By Proposition 2.7, D_A^+ is surjective. Then the result follows by the same argument as Lemma 3.9. □

On (M, g_j), one also has perturbed maps

$$F_s^+ \colon \mathfrak{A}_{k+1}(M, g_j) \times B \to W_{w_j}^k((M,g_j) : \Lambda_+^2 \otimes \mathrm{Ad}\, P)$$

and their differentials

$$D_{A_j}^+ \colon W_{w_j}^{k+1}((M,g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B \to W_{w_j}^k((M,g_j) : \Lambda_+^2 \otimes \mathrm{Ad}\, P).$$

By Lemma 3.4, if j is sufficiently large, then the metrics are uniformly equivalent

$$\tfrac{1}{2}\|\ \|_{\overline{W}_w^*(K_i,g)} \leq \|\ \|_{\overline{W}_{w_j}^*(K_i,g_j)} \leq 2\|\ \|_{\overline{W}_w^*(K_i,g)}.$$

Thus one obtains the main uniform estimates.

Corollary 3.11 *There are constants $C > 0$ and N independent of i, j so that for $|i - j| \geq N$, the uniform estimates*

$$\|D_{\bar{A}_j}^+(\bar{u})\|_{\overline{W}_{w_j}^k(K_i,g_j)} \geq C\|\bar{u}\|_{\overline{W}_{w_j}^{k+1}(K_i,g_j)}$$

hold for all $\bar{u} \in \bar{U}^j = \big(\ker D_{\bar{A}_j}^+\big)^\perp$.

3.2.3 Convergence of image spaces So far we have treated function spaces like $\overline{W}_{w_j}^k(K_i)$ for $j \geq i$. Roughly speaking these are functions with support on K_i equipped with metrics induced from g_j. Our basic idea of the paper is to induce some uniform estimates of functions on $\overline{W}_{w_j}^k(K_i)$ by use of that on $W_w^k(S)$. One may think that in some sense, $\overline{W}_{w_j}^k(K_i)$ approach $W_w^k(S)$ as i goes to infinity. Let $\bar{V}_i^j \subset \overline{W}_{w_j}^k(K_i)$ be a family of subspaces and $V \subset W_w^k(S)$ be another closed subspace. Below we formulate notions of convergence of the family $\{\bar{V}_i^j\}_i$ to V.

3.2.4 Notions of convergence Let us consider a family of subspaces

$$\bar{V}_i^j \subset \bar{W}_{w_j}^{k+1}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P), \quad j \geq i$$

and also a closed subspace $V \subset W_w^{k+1}((S,g); \Lambda^1 \otimes \mathrm{Ad}\, P)$. For each i, we denote its quotient by $\bar{V} \subset \bar{W}_w^{k+1}((K_i, g) : \Lambda^1 \otimes \mathrm{Ad}\, P)$.

Definition 3.12 A bounded sequence $\{\bar{u}_i^j\}_{j \geq i}$ in $\{\bar{V}_i^j\}_{j \geq i}$ lies inside V at infinity, if there are some $\bar{v}_i^j \in \bar{V}$ so that for any small $\epsilon > 0$, there exist an arbitrarily large i_0 and $i' \geq i_0$ so that for all $j \geq i \geq i'$, the estimates

$$\|\bar{u}_i^j - \bar{v}_i^j\|_{\bar{W}_w^{k+1}((K_{i_0}, g); \Lambda^1 \otimes \mathrm{Ad}\, P)} < \epsilon$$

hold, where we regard $\bar{u}_i^j$ as elements in $\bar{W}_w^{k+1}((K_{i_0}, g); \Lambda^1 \otimes \mathrm{Ad}\, P)$ by taking further distance decreasing quotients, $\bar{\ }$: $\bar{W}_w^*(K_i, g) \to \bar{W}_w^*(K_{i_0}, g)$.

We write $\lim_{i,j\to\infty}\{\bar{u}_i^j\} \in V$, when a bounded sequence $\{\bar{u}_i^j\}_{j\geq i}$ in $\{\bar{V}_i^j\}_{j\geq i}$ lies inside V at infinity. We say that $\{\bar{V}_i^j\}$ lies inside V at infinity, if any bounded sequence $\{\bar{u}_i^j\}$ lies inside V as above.

The family of spaces $\{\bar{V}_i^j\}$ is said to *converge* to V, if they lie inside V at infinity, and for any $v \in V$, there is a bounded sequence $\{\bar{u}_i^j\}$ so that for any small $\epsilon > 0$, there exist an arbitrarily large i_0 and $i' \geq i_0$ so that for all $j \geq i \geq i'$, the estimates

$$\left\|\bar{u}_i^j - \bar{v}\right\|_{\bar{W}_w^{k+1}((K_{i_0}, g); \Lambda^1 \otimes \mathrm{Ad}\, P)} < \epsilon$$

hold. A closed space $T \subset W_w^{k+1}((S,g) : \Lambda^1 \otimes \mathrm{Ad}\, P)$ is said to be *orthogonal* to $\lim_{i,j\to\infty} \bar{V}_i^j$, if for any bounded sequence $\bar{u}_i^j \in \bar{V}_i^j$, another bounded sequence $v_i^j \in T$ and any small $\epsilon > 0$, there exist an arbitrarily large i_0 and $i' \geq i_0$ so that for all $j \geq i \geq i'$, the estimates

$$\left\|\bar{u}_i^j - \bar{v}_i^j\right\|_{\bar{W}_w^{k+1}((K_{i_0}, g); \Lambda^1 \otimes \mathrm{Ad}\, P)} \geq (1-\epsilon)\left\|\bar{u}_i^j\right\|_{\bar{W}_w^{k+1}((K_{i_0}, g); \Lambda^1 \otimes \mathrm{Ad}\, P)}$$

hold. Let us introduce an abstract notion in function spaces. Let φ_i be the cut off functions as before. A linear subspace $V \subset W_w^{k+1}((S,g) : \Lambda^1 \otimes \mathrm{Ad}\, P)$ is said to be *quasi local*, if for any $\epsilon > 0$, there is some $N \gg 0$ so that for any bounded sequence $\{v_i\}_i \subset V$ with $\|v_i\| \leq 1$, and $k > 0$, there are some $k \leq n_{k,i} \leq k + N$ so that $\|\operatorname{pr}(\varphi_{n_{k,i}} v_i)\| < \epsilon$, where pr: $W_w^{k+1}((S,g) : \Lambda^1 \otimes \mathrm{Ad}\, P) \to V^{\perp}$ is the orthogonal projection.

Any finite dimensional subspaces are quasi local: $\operatorname{im} d_A \subset W_w^{k+1}((S,g) : \Lambda^1 \otimes \mathrm{Ad}\, P)$ and $\ker d_A^+$ are both quasi local, where for the former closedness of $\operatorname{im} d_A$ is used (see the proof of Lemma 3.16).

Lemma 3.13 *Let $V \subset U \subset W_w^{k+1}((S,g) : \Lambda^1 \otimes \mathrm{Ad}\, P)$ be embeddings of closed subspaces. Assume that both V and $T = V^\perp \cap U$ are quasi local. If a bounded sequence $\{\overline{\alpha}_i^j\}$ lies inside U and is orthogonal to V at infinity, then $\{\overline{\alpha}_i^j\}$ converges to T at infinity.*

Proof Notice that in general, for any bounded sequences $\{a_i\}_i \subset V$ and $\{b_i\}_i \subset T$, if

$$\|\overline{a}_i - \overline{b}_i\|_{\overline{W}_w^{k+1}(K_{i_0},g)} \to 0$$

as $i \to \infty$ for all i_0, then both $\|\overline{a}_i\|_{\overline{W}_w^{k+1}(K_{i_0},g)}$ and $\|\overline{b}_i\|_{\overline{W}_w^{k+1}(K_{i_0},g)}$ must go to zero.

By definition, there is a sequence $\{u_i^j\} \subset U$ so that

$$\|\overline{\alpha}_i^j - \overline{u}_i^j\|_{\overline{W}_w^{k+1}(K_{i_0},g)} < \epsilon$$

for all $i' \leq i \leq j$. Let us decompose $u_i^j = u_j^i(1) \oplus u_j^i(2)$ and $\alpha_i^j = \alpha_i^j(1) + \alpha_i^j(2) + \alpha_i^j(3)$ with respect to $W_w^{k+1} = V \oplus T \oplus U^\perp$ respectively. Firstly by convergence to U,

$$\|\overline{\alpha}_i^j(3)\|_{\overline{W}_w^{k+1}(K_{i_0},g)} \to 0.$$

Thus

$$\|(\overline{\alpha}_i^j(1) - \overline{u}_i^j(1)) + (\overline{\alpha}_i^j(2) - \overline{u}_i^j(2))\|_{\overline{W}_w^{k+1}(K_{i_0},g)} \to 0,$$

and so in particular $\|\overline{\alpha}_i^j(1) - \overline{u}_i^j(1)\|_{\overline{W}_w^{k+1}(K_{i_0},g)} \to 0$.

We show the estimates

$$\|\overline{\alpha}_i^j - \overline{u}_i^j(2)\|_{\overline{W}_w^{k+1}(K_{i_0},g)} < \epsilon$$

for all $i'' \leq i \leq j$ for some $i' \leq i''$. Suppose not, and assume the estimates

$$\|\overline{\alpha}_i^j - \overline{u}_i^j(2)\|_{\overline{W}_w^{k+1}(K_{i_0},g)} \geq \epsilon_0 > 0$$

for some ϵ_0 for $i'' \leq i \leq j$. Then the lower bounds

$$\|\overline{\alpha}_i^j(1)\|_{\overline{W}_w^{k+1}(K_{i_0},g)} \geq \epsilon_0$$

hold. This contradicts to the assumption of orthogonality of the sequence to V. □

Remark 3.14 Our notion of convergence does not care on behaviour at infinity of individual functions, since we concern norms on compact subsets.

3.2.5 Finite dimensionality of cokernels at infinity Let us fix i and take sufficiently large $j \gg i$. We consider the quotient spaces $\overline{W}^*_{w_j}((K_i, g_j); \Lambda^* \otimes \mathrm{Ad}\, P)$. Let

$$d_{\bar{A}_j} \colon \overline{W}^{k+2}_{w_j}((K_i, g_j); \mathrm{Ad}\, P) \to \overline{W}^{k+1}_{w_j}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P)$$

be the differentials, and put $\bar{V}^j_i = \mathrm{im}\, d_{\bar{A}_j}$. Let us consider the closed subspaces

$$V = \mathrm{im}\, d_A, \quad U = \ker d^+_A \subset W^{k+1}_w((S, g); \Lambda^1 \otimes \mathrm{Ad}\, P)$$

and denote its quotient by $\bar{V} \subset \overline{W}^{k+1}_w((K_i, g); \Lambda^1 \otimes \mathrm{Ad}\, P)$.

The next Proposition is crucial for our later discussion.

Proposition 3.15 *The orthogonal subspace $T \subset U$ to $\lim_{i,j \to \infty} \bar{V}^j_i$ is finite dimensional.*

For this, we have the following Lemma.

Lemma 3.16 *The family $\{\bar{V}^j_i\}_{j \geq i}$ lies inside U and contains V at infinity.*

Proof Let $\bar{u}^j_i \in \bar{V}^j_i$ be a bounded sequence with $\|\bar{u}^j_i\|_{\overline{W}^{k+1}_{w_j}(K_i, g_j)} \leq C$. We choose representatives $u^j_i \in W^{k+1}_{w_j}(M, g_j)$ with

$$\|\bar{u}^j_i\|_{\overline{W}^{k+1}_{w_j}(K_i, g_j)} = \|u^j_i\|_{W^{k+1}_{w_j}(M, g_j)}.$$

Let us take any i_0. Then for any $\epsilon > 0$, there is a larger i' such that for all $i \geq i'$, there are some $i_0 \leq i'' \leq i$ such that

$$\|u^j_i\|_{W^{k+1}_{w_j}(K_{i''+1} \backslash K_{i''}, g_j)} \leq \epsilon.$$

Let $\varphi_{i''} \in C^\infty_c(K_{i''+1})$ be the cut off function in Section 3.2.1. Then one has the estimates

$$\|d^+_{A_j}(\varphi_{i''} u^j_i)\|_{W^k_{w_j}(M, g_j)} \leq C\epsilon$$

where C is independent of i, j.

Now regard $\varphi_{i''} u^j_i \in W^{k+1}_w(S, g)$ and decompose

$$\varphi_{i''} u^j_i = w^j_i + z^j_i$$

with respect to $W^{k+1}_w(S, g) = U \oplus U^\perp$, where $\|z^j_i\|_{W^{k+1}_w}(S, g) \leq C\epsilon$ hold. Since the metrics

$$\|\ \|_{\overline{W}^*_{w_j}(K_{i_0}, g_j)} \quad \text{and} \quad \|\ \|_{\overline{W}^*_w(K_{i_0}, g)}$$

are uniformly equivalent by Lemma 3.4, it follows that w_i^j is the desired family. Thus one has shown that $\lim_{i,j} \overline{V}_i^j \subset U$.

Next we show surjectivity, that the orthogonal subspace to $\lim_{i,j\to\infty} \overline{V}_i^j$ in V is equal to zero.

Let $u = d_A(f) \in W_w^{k+1}((S,g); \Lambda^1 \otimes \mathrm{Ad}\, P)$ be a unit element. Then since d_A has closed range with zero kernel [10], there is a positive constant $C > 0$ so that the estimate

$$\|f\|_{W_w^{k+2}} \leq C\|d_A(f)\|_{W_w^{k+1}} = C$$

holds. Thus for any small $\epsilon > 0$, there are sufficiently large i so that the estimate:

$$\|u - d_A(\varphi_i f)\|_{W_w^{k+1}} < \epsilon$$

holds. Since one can assume $\varphi_i f \in W_{w_j}^{k+2}((M, g_j) : \mathrm{Ad}\, P)$, the assertion holds. □

Proof of Proposition 3.15 Since $H^1(A) = \ker d_A^+ / d_A(W_w^{k+2}((S,g); \mathrm{Ad}\, P)$ is of finite dimension by Theorem 1.6, the above lemma already implies that $T \subset U$ is of finite dimension. □

Corollary 3.17 *Suppose a bounded sequence $\{\overline{v}_i^j\}_{i,j}$ is orthogonal to $V = \mathrm{im}\, d_A$ at infinity. If*

$$\lim_{i,j\to\infty} d^+_{\overline{A}_j}(\overline{v}_i^j) = 0$$

holds, then the above sequence lies inside a finite dimensional subspace T at infinity.

The first condition of orthogonality is realized by use of Coulomb gauge representatives. As before suppose ASD connections A_j over (M, g_j) converge to A over (S, g). Let us consider connections on the quotients $\overline{A}'_j = A_j + \overline{\alpha}_i^j \in A_j + \overline{W}_{w_j}^{k+1}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P)$.

Proposition 3.18 *There is some positive constant $c > 0$ independent of i, j so that after gauge transformation by elements*

$$\overline{u}_j \in \overline{W}_{w_j}^{k+2}(K_i, g_j) \cap \mathrm{Aut}(E|K_i),$$

and taking subindices of (i, j), the connections $\overline{\beta}_j \equiv \overline{u}_j^*(\overline{A}'_j) - A_j$ are orthogonal to $V = \mathrm{im}\, d_A$ at infinity, as far as uniform estimates

$$\|\overline{\alpha}_i^j\|_{\overline{W}_{w_j}^{k+1}} < c$$

hold. Moreover $\|\overline{\beta}_j\|_{\overline{W}_{w_j}^{k+1}}$ is bounded by $C\|\overline{\alpha}_i^j\|_{\overline{W}_{w_j}^{k+1}}$, where C is independent of i, j and c.

We combine the standard choice of Coulomb gauge representatives over (S,g) with the cut and paste method.

Sublemma 3.19 *Let $A' = A+\alpha \in A+W^{k+1}_w((S,g):\Lambda^1 \otimes \mathrm{Ad}\, P)$ satisfy $\|\alpha\|_{W^{k+1}_w} \leq c$ for some positive constant $c>0$. Then there is a gauge transformation $u \in \mathfrak{G}_{k+2}(P)_0$ (see Section 2.2), so that $(d_A)^*_w(u^*(A') - A) = 0$ holds. Thus $\beta = u^*(A') - A$ is orthogonal to im d_A.*

Proof This is standard for compact base Donaldson–Kronheimer [3, page 56]. For our case, it also works since the proof uses only the implicit function theorem. We seek for a solution to the equation $G(u) \equiv (d_A)^*_w(d_A(u)u^{-1} - u\alpha u^{-1}) = 0$, where $u \in \mathfrak{G}_{k+2}(P)_0$. Its derivative is given by $DG(v) = (d_A)^*_w(d_A(v) + [\alpha, v])$, where $DG\colon W^{k+2}_w((S,g):\mathrm{Ad}\, P) \to W^k_w((S,g):\mathrm{Ad}\, P)$. By Theorem 1.6 and Proposition 2.3(2), d_A has closed range, and so DG is a surjection, if $\|\alpha\|_{W^{k+1}_w}$ is sufficiently small. Thus by the implicit function theorem, the result follows. □

Proof of Proposition 3.18 The rest uses cut and paste method in Lemma 3.16 We choose representatives $\alpha^j_i \in W^{k+1}_{w_j}((M,g_j):\Lambda^1 \otimes \mathrm{Ad}\, P)$ with

$$\|\overline{\alpha}^j_i\|_{\overline{W}^{k+1}_{w_j}(K_i,g_j)} = \|\alpha^j_i\|_{W^{k+1}_{w_j}(M,g_j)}.$$

Let us take a small $\epsilon>0$ and $i \geq [\epsilon^{-1}]+1$, where [ ] is the integer part. Then there is some $i'' \leq i$ and infinite subindices of j so that

$$\|\alpha^j_i\|_{W^{k+1}_{w_j}(K_{i''}\backslash K_{i''-1},g_j)} \leq \epsilon.$$

Regarding $\varphi_{i''-1}\alpha^j_i \in W^{k+1}_w(S,g)$, one can apply Sublemma 3.19 since $\varphi_{i''-1}\alpha^j_i$ still have small norms. Thus there are gauge transformations u_j such that $\beta_j = u^*_j(B'_j) - A$ is orthogonal to im d_A, where $B'_j = A + \varphi_{i''-1}\alpha^i_j$. Then the quotient of the family $\overline{u}_j \in \overline{W}^{k+2}_{w_j}(K_i,g_j)$ is the desired one, as follows. By a similar argument, taking other subindices if necessary, one may assume that

$$\|\beta_j\|_{W^{k+1}_{w_j}(K_{i''}\backslash K_{i''-1},g_j)} \leq \epsilon.$$

Then $\varphi_{i''-1}\beta_j$ satisfy the estimates

$$\left|\langle \varphi_{i''-1}\beta^j_i, u\rangle_{W^{k+1}_{w_j}(M,g_j)}\right| \leq \epsilon\|u\|$$

for any $u \in$ im d_A. Since A_i converge to A, the family

$$\overline{\beta}'_j \equiv \overline{u}^*_j(A_j + \overline{\alpha}^i_j) - A_j \in \overline{W}^{k+1}_{w_j}((K_i,g_j):\Lambda^1 \otimes \mathrm{Ad}\, P)$$

is orthogonal to im d_A at infinity. □

3.2.6 Uniformity in implicit function theorems In Section 3.1, we have seen that for our purpose some uniformity of spectral radii are seriously related. At the first part of Section 3.2, we have explained that K_i^c regions of connections $A_i|K_i^c$ will affect to control the behaviour of the spectrum, and this is the reason why we are using quotient function spaces $\overline{W}_w^k$. On the other hand we have seen in Section 3.1.2 that once one has obtained some uniformity of injectivity radii from below, then there are families of balls with uniform sizes where one can apply the implicit function theorem and perturbation for families of spaces.

Now we use quotient function spaces $\overline{W}_{w_j}^k(K_i, g_j; \Lambda^*)$ to apply the abstract implicit function theorem for families in Lemma 3.1, and induce some spectral information on $W_w^k(S, g; \Lambda^* \otimes \mathrm{Ad}\, P)$.

Let A_j be ASD connections over (M, g_j) which converges to A over (S, g) as before. Let

$$F_s^+ \colon A_j + W_{w_j}^{k+1}((M, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \times B \to W_{w_j}^k((M, g_j) : \Lambda_+^2 \otimes \mathrm{Ad}\, P)$$

be the parameterized self dual curvature operators, and denote their differentials at A_j

$$D_{A_j}^+ \colon W_{w_j}^{k+1}((M, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B \to W_{w_j}^k((M, g_j) : \Lambda_+^2 \otimes \mathrm{Ad}\, P).$$

Similarly one obtains

$$D_A^+ \colon W_w^{k+1}((S, g) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B \to W_w^k((S, g) : \Lambda_+^2 \otimes \mathrm{Ad}\, P).$$

These are all surjective maps by Proposition 2.7. By putting $A \equiv A_\infty$, one obtains a family of operators $\{D_{A_j}^+\}_{0 \le j \le \infty}$.

Let us fix i, and for any $i \le j \le \infty$, one obtains projections

$$\overline{} \colon W_{w_j}^k((X, h) : \Lambda^* \otimes \mathrm{Ad}\, P) \to \overline{W}_{w_j}^k((K_i, h) : \Lambda^* \otimes \mathrm{Ad}\, P)$$

where $(X, h) = (M, g_j)$ or (S, g).

Let us take $\epsilon > 0$ so that the image of the unit ball contains the ϵ ball

$$D_A^+(B_1) \supset D_\epsilon$$

where $B_1 \subset W_w^{k+1}((S, g) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B$ and $D_\epsilon \subset W_w^k((S, g) : \Lambda_+^2 \otimes \mathrm{Ad}\, P)$. This is equivalent to the estimates

$$\left\| D_A^+(u, c) \right\|_{W_w^k} \ge \epsilon \|(u, c)\|_{W_w^{k+1} \oplus T_0 B}$$

for all $(u, c) \in (\ker D_A^+)^\perp$ (see proof of Lemma 3.9). Let us denote the induced map:

$$D_{\overline{A}}^+ \colon \overline{W}_w^{k+1}((K_i, g) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B \to \overline{W}_w^k((K_i, g) : \Lambda_+^2 \otimes \mathrm{Ad}\, P).$$

Similarly for sufficiently large $j \geq i$, there are the induced maps

$$D^+_{\bar{A}_j} \colon \overline{W}^{k+1}_{w_j}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B \to \overline{W}^k_{w_j}((K_i, g_j) : \Lambda^2_+ \otimes \mathrm{Ad}\, P).$$

By Proposition 2.7 and Lemma 3.7, these induced maps are all surjective. By Corollary 3.11, one has uniformity of spectral radii

$$\| D^+_{\bar{A}_j}(\bar{u}, c)\|_{\overline{W}^k_{w_j}} \geq C \|(\bar{u}, c)\|_{\overline{W}^{k+1}_{w_j} \oplus T_0 B}$$

for all $(\bar{u}, c) \in (\ker D^+_{\bar{A}_j})^{\perp}$ and for all large $i \ll j \leq \infty$, where C is independent of i and j.

For $i \leq j \leq \infty$, let us consider

$$F^+_s \colon A_j + \overline{W}^{k+1}_{w_j}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \times B \to \overline{W}^k_{w_j}((K_i, g_j) : \Lambda^+_2 \otimes \mathrm{Ad}\, P),$$

where at $i = \infty$, we regard $\overline{W}^k_{w_j}(K_i, g_j) = W^k_w(S, g)$. Let us denote the kernels of the surjective differentials $D^+_{\bar{A}_j}$ and their orthogonal complements by

$$\bar{R}^j = \ker D^+_{\bar{A}_j} \quad \text{and} \quad \bar{U}^j = \big(\ker D^+_{\bar{A}_j}\big)^{\perp}$$

in $\overline{W}^{k+1}_{w_j}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B$. respectively.

Let us identify open neighbourhoods of $A_j + \overline{W}^{k+1}_{w_j}\big((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P\big) \times B$ at $(A_j, 0)$ by $\bar{N}^j_1 \times \bar{N}^j_2 \subset \overline{W}^{k+1}_{w_j}\big((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P\big) \times T_0 B$, where $\bar{N}^j_1 \subset \bar{R}^j$ and $\bar{N}^j_2 \subset \bar{U}^j$. Thus one may regard the above map as $F^+_s \colon \bar{N}^j_1 \times \bar{N}^j_2 \to \overline{W}^k_{w_j}$. Let D_2 be the differential on the second factor with respect to the decomposition $\overline{W}^{k+1}_{w_j}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B = \bar{R}^j \oplus \bar{U}^j$, as in the abstract implicit function theorem.

Combining Lemma 3.8 with Corollary 3.11, one obtains the following conditions which are required to apply Lemma 3.1.

Corollary 3.20 *There exists $\epsilon_0 > 0$ such that for all $(a, b) \in \bar{N}^j_1 \times \bar{N}^j_2$ with $\|(a, b)\| \leq \epsilon_0$, there are constants C_1, C_2, C_3 such that for any i, there is a large $j_0 \geq i$ so that for all $\infty \geq j \geq j_0$, the differentials*

$$\bar{F}^+_s : A_j + \overline{W}^{k+1}_{w_j}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus B \to \overline{W}^k_{w_j}((K_i, g_j); \Lambda^2_+ \otimes \mathrm{Ad}\, P)$$

satisfy uniform estimates

$$0 < C_1 \leq \big\| D_2 \bar{F}^+_s \big\|, \quad \big\| D \bar{F}^+_s \big\|_{\inf} \leq C_2 \quad \text{and} \quad \big\| D^2_2 \bar{F}^+_s \big\| \leq C_3$$

at (a, b).

The implicit function theorem states that for all $i \ll j \le \infty$, there are smooth maps $G_j\colon \bar{B}^j_1 \to \bar{B}^j_2$, where $\bar{B}^j_2 \subset \bar{U}^j$ and $\bar{B}^j_1 \subset \bar{R}^j$ are open neighbourhoods, such that $F^+_s(\bar{A}', b') = 0$ for $(\bar{A}', b') \in \bar{B}^j_1 + \bar{B}^j_2$, if and only if $(\bar{A}', b') = (\bar{B}, c) + G_j(\bar{B}, c)$ for some $(\bar{B}, c) \in \bar{B}^j_1$. Moreover $d(G_j)_{(\bar{A}',b')} = -\big[\big(D^+_{(\bar{A}',b')}|\bar{U}^j\big)^{-1} \circ D^+_{(\bar{A}',b')}\big]$ holds.

Now we are in a position to apply Lemma 3.1.

Corollary 3.21 *For all $i \ll j \le \infty$, the following hold.*

1. *There exist $\delta, \mu > 0$, independent of i and j, such that for the δ–ball $\bar{D}^j_\delta \subset \bar{W}^{k+1}_{w_j}\big((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P\big)$ and the μ–ball $D_\mu \subset T_0 B$, $\bar{B}^j_1 + \bar{B}^j_2$ contains $\bar{D}^j_\delta \times D_\mu$.*
2. *There are uniform estimates of norms*

$$\|dG_j\| \le C.$$

Proof

1. By Lemma 3.1, there exist $\delta_1, \delta_2 > 0$, independent of $j \gg i$ such that $\bar{B}^j_m$ contain δ_m balls $B(\delta_m)$ for $m = 1, 2$. Since they are open subsets, there exist $\delta, \mu > 0$ such that for $\bar{D}^j_\delta$ and D_μ as above, the inclusion $\bar{D}^j_\delta \times D_\mu \subset B(\delta_1) + B(\delta_2)$ holds.
2. By Corollary 3.11, one obtains uniform estimates $\big\|(D^+_{\bar{A}_j}|\bar{U}^j)^{-1}\big\| \le C^{-1}$ of operator norms.

Thus the result follows from Corollary 3.20 and the formula for dG_j above. □

3.2.7 Restriction of the differentials Let us consider $d^+_A\colon W^{k+1}_w((S,g) : \Lambda^1 \otimes \mathrm{Ad}\, P) \to W^k_w((S,g) : \Lambda^2_+ \otimes \mathrm{Ad}\, P)$, and denote

$$V = \mathrm{im}\, d_A, \quad W = V^\perp \subset W^{k+1}_w((S,g) : \Lambda^1 \otimes \mathrm{Ad}\, P).$$

By Corollary 2.8, the restriction of the differential of F^+_s at $(A, 0)$

$$D^+_A\colon W \oplus T_0 B \to W^k_w\big((S,g) : \Lambda^2_+ \otimes \mathrm{Ad}\, P\big)$$

is still surjective.

Later on, we will restrict on the closed subspace W. This becomes important in Section 3.2.8.

Let us put

$$U_1 \equiv (W \oplus T_0 B) \cap \ker D^+_A, \quad U_2 = (W \oplus T_0 B) \cap (\ker D^+_A)^\perp.$$

Then by the implicit function theorem again, there are balls $B_1 \subset U_1$ and $B_2 \subset U_2$, and a smooth function $G\colon B_1 \to B_2$ such that any solution $(A, b) \in B_1 + B_2$ is on the graph of G.

3.2.8 Dimension comparisons by (in)finiteness Here we carefully choose a generic path $b_t \in B$. We will assume that the space H of the cokernel of d_A^+ has positive dimension (see the second paragraph of Section 3.2).

Firstly let

$$T = \ker d_A^+ \cap (\operatorname{im} d_A)^\perp \subset W_w^{k+1}((S,g) : \Lambda^1 \otimes \operatorname{Ad} P)$$

be the finite dimensional subspace (see Proposition 3.15). Let $D \subset T$ and $B_1 \subset T_0 B$ be the unit balls. We define the projection

$$\mathrm{pr}\colon W_w^{k+1}((S,g) : \Lambda^1 \otimes \operatorname{Ad} P) \oplus T_0 B \to T_0 B,$$

and let Gr be the graph of G above. Let us put

$$\widetilde{G}r(D,0) = Gr \cap D \times B_1 \subset W \oplus T_0 B.$$

The following Lemma is a key to the proof of the Theorem. It says that a perturbation b_t can be chosen transverse to the projection of the tangent space at $[A]$ of the slice of the universal moduli space.

Lemma 3.22 *There is some constant $C > 0$ and smooth path $b_t \in B$ with $b_0 = 0$ so that the estimates*

$$\|b_t - \mathrm{pr} \circ \widetilde{G}r(D,0)\| \geq Ct$$

hold for all $0 \leq t \leq 1$.

Proof We start by explaining the idea of the proof in the special case $T = 0$. Consider the linear space

$$\widetilde{V} = \ker D_A^+ \cap T_0 B.$$

By the construction of the perturbation, $\widetilde{V} \subset T_0 B$ has infinite codimension. In fact the image $ds(0, T_0 B)$ is already infinite dimensional. We define the map

$$\widetilde{G}_0\colon \widetilde{V} \to T_0 B$$

by $c \to \mathrm{pr} \circ G(0,c)$. Let $\widetilde{G}_0 r$ be the graph of $\widetilde{G}_0$. Then since $\|dG\| \leq C$ is bounded, and $\widetilde{V}^\perp \subset T_0 B$ is infinite dimensional, there is a smooth path b_t which satisfies the lower bound

$$\left\| b_t - \widetilde{G}_0 r \cap B \right\| \geq Ct$$

for all t, as follows. Let $T\widetilde{G}_0 r = \{v + d\widetilde{G}_0(v)\} \subset T_0 B$ be the closed linear subspace. Then since $\widetilde{V}^\perp$ is infinite dimensional, its orthogonal complement $(T\widetilde{G}_0 r)^\perp$ is also infinite dimensional. In particular one can choose a smooth path $b_t \in B$ such that $\frac{d}{dt}b_t\big|_{t=0} \in (T\widetilde{G}_0 r)^{\perp}$. Since $\|dG\| \le C$ is bounded, this gives a path satisfying the above estimate.

The conclusion follows immediately for the general case, if we use positivity of dim $H > 0$ and a special property of the holonomy perturbation that, if we restrict to $W^{k+1}_w((S,g) : \Lambda^1 \otimes AdP) \subset W^{k+1}_w((S,g) : \Lambda^1 \otimes AdP) \oplus T_0 B$, then the restriction $ds_{(A,0)} | W^{k+1}_w((S,g) : \Lambda^1 \otimes AdP) \equiv 0$. On the other hand by the same proof as Lemma 2.6 one can modify s so that $ds_{(A,0)}|T$ is injective. However we shall in fact use a different method which is applicable to general local perturbations. The proof that follows implies that the conclusion also holds if we replace T by another finite dimensional linear subspace T' which is sufficiently near T.

We follow a parallel argument to the special case. Let $B_1 \subset T_0 B$ and $B'_1 \subset W$ be the unit balls. Let Gr be the graph of G. For each $c \in B_1$, let us define a set

$$S(c) = \{\mathrm{pr} \circ Gr(m,c) : (m,c) \in U_1 \cap (B'_1 \times B_1)\}.$$

This defines a multi-valued map

$$\widetilde{G}\colon\ B_1 \cap (\mathrm{pr} \circ U_1) \to T_0 B.$$

We claim that each $\widetilde{G}(c)$ has at most finite dimension bounded by $\dim T$. In order to verify this, it is enough to see that the set

$$dS(c) \equiv \{\mathrm{pr} \circ dGr(m,c) : (m,c) \in T_0 U_1\}$$

is finite dimensional.

For $(m,c) \in T_0 U_1$, $D^+_A(m,c) = 0$ holds, since $U_1 = \ker D^+_A \cap (W \oplus T_0 B)$. Let us put $dGr(m,c) = (m,c) + dG(m,c)$. Suppose $\mathrm{pr} \circ dGr(m,c) = \mathrm{pr} \circ dGr(m',c)$ hold for some $(m,c), (m',c) \in T_0 U_1$. Then $m - m' \in \ker D^+_A \cap W$. We see that this space is finite dimensional. Let us decompose $W = T \oplus (\ker d^+_A)^{\perp}$. Then there is a constant $C > 0$ independent of the perturbation such that for any $v \in (\ker d^+_A)^{\perp}$, $\|d^+_A(v)\| \ge C\|v\|$ holds. Since the perturbation is sufficiently small, one may assume the estimate $\|D^+_A(v)\| \ge C\|v\|$ still holds. In particular the projection

$$\ker D^+_A \cap W \hookrightarrow T$$

is injective, where the latter space is finite dimensional. This verifies the claim.

Let us verify that the closure of $pr \circ \widetilde{G}r(D,0) \subset T_0 B$ has infinite codimension. Then one can obtain the desired b_t by the same argument as the first paragraph of the proof. We next verify this infinite codimensionality.

Let us choose small and positive constants $1 >> c' > c > 0$, and denote the infinite dimensional linear space:

$$J = \{w \in T_0 B : c'\|w\| \geq \|ds(0,w)\| \geq c\|w\|\}$$

consisting of vectors of bounded spectrum (Remark 2.10).

We show that $(T \times J) \cap U_1$ is finite dimensional. Suppose not. By compactness of the unit ball of T, for any orthonormal sequence $\{(m_i, v_i)\}_i$ such that $\|(m_i, v_i)\| = b > 0$ with $\{(m_i, v_i)\}_i \in T \times J \cap U_1$, there is some $m \in T$ so that $\lim_{i\to\infty} m_i = m$ holds. By definition of J, the estimates $c\|v_i - v_j\| \leq \|ds(0, v_i - v_j)\|$ hold. On the other hand for any small $\epsilon > 0$,

$$\|ds(0, v_i - v_j)\| = \|D_A^+(0, v_i - v_j)\| < \epsilon$$

hold for all large i, j.

Now suppose $\|m\| = b$. Then $\lim_i \|v_i\| = 0$ and so $\{(m_i, v_i)\}_i$ are not orthogonal. Thus $0 \leq \|m\| < b$ holds. Then $0 < \lim_i \|v_i\| = a \leq b$, and so

$$\lim_{i,j} \|v_i - v_j\| = \sqrt{2a} > 0.$$

But this can not happen by the above estimates. This is a contradiction.

Finally we claim that the closure of $\mathrm{pr} \circ U_1 \subset T_0 B$ has infinite codimension. For simplicity of notation, we assume $\mathrm{pr} \circ U_1 \subset T_0 B$ is closed. The proof in this case carries over to the general case. We argue by contradiction. Assume it has finite codimension less than $l - 1 < \infty$, where l is sufficiently large more than $\dim T$. The argument is complicated by the fact that $\mathrm{pr}\, U_1 \cap \mathrm{pr}\, U_2$ may have positive dimension.

By infinite dimensionality, we can choose linearly independent vectors $(m_i, c_i') \in W \times M$ for $i = 1, \dots, l$. Let us put $c_i = w_0 + c_i'$ and consider $(m_i, c_i) \in W \times J$. Each (m_i, c_i) decomposes as $\left(m_i^j, c_i^j\right)$, $j = 1, 2$ with respect to $\ker D_A^+ \oplus (\ker D_A^+)^\perp$. By the finite codimension assumption, there are real numbers $\{a_i\}_i \neq 0$ such that the inclusion $\sum_i a_i c_i^2 \in \mathrm{pr}\, U_1$ should hold. One may assume $\sum_i a_i = 1$. Then there is some $m \in W$ such that $(m, 0) + \sum_i a_i(m_i^2, c_i^2) \in U_1$. Thus the inclusion

$$(m, 0) + \sum_i a_i(m_i, c_i) \in (W \times J) \cap \ker D_A^+$$

holds. This space is finite dimensional.

This implies that if one takes linearly independent vectors $(m_i, c_i') \in W \times M$, $i = 1, \dots, l'$ for a sufficiently large $l' \geq l$, then for some real numbers $\{a_j\}_j \neq 0$, the linear combination $\sum_{j=1}^{l'} a_j(m_j, c_j) \in W \times \{w_0\}$. This contradicts the infinite dimensionality of J and completes the proof. □

Remark 3.23 (1) An ASD connection A over (S,g) is obtained as a limit of A_j over (M,g_j). Such choices of limits are not unique. Thus if one chooses another limit, then one will obtain another connection $A' \in A + L^2((S,g):\Lambda^1 \otimes \mathrm{Ad}\,P)$, and so the choice of path b_t in Lemma 3.22 depends on such limits. We want to choose b_t independently of choice of such limits, and for this we modify the perturbation space as follows. Let us take another Hilbert space C which is the closure of the infinite sums $\bigoplus^\infty B$ as another perturbation space. Let us choose a sufficiently decreasing sequence of constants $\epsilon_i > 0$, $\lim_i \epsilon_i = 0$, and put the perturbation $\bar{s}\colon \mathfrak{A}_k(E) \times C \to W^k_w$ by $\bar{s}\big(A', \bigoplus_i b_i\big) \equiv \sum \epsilon_i s_0(A', b_i)$. Then one can choose a path $b_t \in C$ independently of choices of limits and so of A, which satisfies the conclusion of Lemma 3.22 as follows. In fact by covering $L^2((S,g):\Lambda^1 \otimes \mathrm{Ad}\,P)$ by countably many small open subsets, there are infinite number of paths $b^i_t \in B$, $i = 0, 1, \dots$, for which, any A have some b^i_t which satisfy the conclusion of Lemma 3.22. Then one has the desired path in C by $(b^0_t, \frac{1}{2}b^1_t, \dots, \frac{1}{2^i}b^i_t, \dots) \in C$. Later on we will assume that a path b_t in Lemma 3.22 are chosen independent of A. Such property of independence of b_t is used in Section 3.2.9.

(2) When the tangent space of the ASD moduli space $T = T_{[A]}\mathfrak{M}(A)_w$ happens to be zero dimensional, then the proof of Lemma 3.22 becomes simpler as shown at the start of the proof above. Moreover the proof of Proposition 3.24 below also becomes much easier. It follows from Theorem 1.6 that such situations can occur if the limit of the ASD connections is trivial, since in this case the first cohomology of the AHS complex is zero. This can be seen by looking at the second Stiefel–Whitney classes of the bundles in detail (see Lemma 2.4). Notice also that in this case the gauge group actions are free, since they are assumed to be the identity at infinity. This situation will be considered elsewhere.

3.2.9 Asymptotic solutions

So far we have not used any special properties of ASD equations, rather we have worked within a general functional analytic framework. Here we use the assumption that the Donaldson invariant does not vanish over $E \to M$.

Let $A_j \in \mathfrak{M}(M,g_j)$ be a sequence of ASD connections on regular moduli spaces $j = 1, 2, \dots$, and suppose they converge to an L^2 ASD connection A over (S,g). We denote by $\mathfrak{M}(M,g_j)$ the usual ASD moduli spaces, and $\mathfrak{M}(A)_w$ be the weighted moduli space in section 1. Similarly we denote the perturbed moduli spaces $\mathfrak{M}_b(M,g_j)$ or $\mathfrak{M}_b(A)_w$ in Section 2.3. Concerning $\mathfrak{M}(M,g_j)$, we will choose the Floer element $b \in P$ for generic values of the perturbation. In particular these are all smooth and finite dimensional manifolds.

Let $B_j = A_j + \alpha_j$ be a connection on (M,g_j). Let us fix i and for $j \geq i$, denote by $\bar{B}_j = A_j + \bar{\alpha}_j$ the projection of B_j on $A_j + \overline{W}^{k+1}_{w_i}((K_i,g_j):\Lambda^1 \otimes \mathrm{Ad}\,P)$.

Let $\bar{D}^j_\delta \subset \bar{W}^{k+1}_{w_j}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P)$ be the δ ball, and $D_\mu \subset T_0 B$ be the μ ball. By Corollary 3.21, there are $\delta, \mu > 0$ independent of $j \gg i$ so that one has obtained local charts

$$\bar{D}^j_\delta \times D_\mu \subset \bar{W}^{k+1}_{w_j}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P) \oplus T_0 B$$

around $\bar{A}_j$, in which the implicit function theorem works for the perturbed ASD equations F^+_s.

We now prove the following proposition.

Proposition 3.24 *There is a regular point $b \in B$ with respect to $\mathfrak{M}_b(A)_w$ with a sufficiently small norm, such that there exist elements*

$$(\bar{A}_j - \bar{C}_j, b) \in (\bar{D}^j_\delta \times D_\mu)$$

satisfying the equations $F^+_s(\bar{C}_j, b) = 0$ with respect to (K_i, g_j).

In particular there is a uniform constant C independent of i and j so that the estimates

$$\| \bar{A}_j - \bar{C}_j \|_{\bar{W}^{k+1}_{w_j}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P)} \leq C$$

hold for all $i \ll j$.

Proof Here we use the property that the generic moduli spaces $\mathfrak{M}(M, g_i)$ are nonempty, since the Donaldson invariant is not zero. Even though there are cobordisms between $\mathfrak{M}(M, g_j)$ and $\mathfrak{M}_b(M, g_j)$, one has to check that the latter is nonempty inside $\bar{D}^j_\delta \times D_\mu$ after taking the quotient.

Now let $P \subset B$ be the Floer perturbation. Since $D^+_A | W \oplus T_0 P$ is already surjective by Corollary 2.8, and since the inclusion is dense, one can replace any small generic element $b \in B$ by another one $b' \in P$ with $\|b - b'\|$ sufficiently small. In particular for b_t in Lemma 3.22, one may assume that it is a generic path inside P.

Since each ASD connection A_j is regular and so if b is generic, then there exists one dimensional cobordism between $\mathfrak{M}(M, g_j)$ and $\mathfrak{M}_b(M, g_j)$ parameterized by A^t_j. Below we show that its image $\bar{A}^1_j$ under the projection gives the desired $\bar{C}_j$.

By the Sard–Smale transversality theorem, there is a generic path $b_t \in B$, $b_0 = 0$ and $b_1 = b$ with respect to all $\mathfrak{M}(M, g_j)$, $j = 1, 2, \ldots$. One may assume that b_t satisfies the conclusion of Lemma 3.22. Since the Donaldson invariant does not vanish, there are nonempty cobordisms between $\mathfrak{M}(M, g_j)$ and $\mathfrak{M}_b(M, g_j)$. In particular for each j, one can choose some A_j so that there is a parameterization A^t_j, $0 \leq t \leq 1$ with $A^0_j = A_j$ and satisfying the equations $F^+_s(A^t_j, b_t) = 0$.

Let us put $A_j - A_j^t = \alpha_j^t$ and denote the quotients

$$\overline{A}_j^t = A_j + \overline{\alpha}_j^t \in A_j + \overline{W}_{w_j}^{k+1}\big((K_i, g_j) : \Lambda^1 \otimes \operatorname{Ad} P\big).$$

Lemma 3.25 *One may assume that the family $\{\overline{\alpha}_j^t\}_j$ is orthogonal to im d_A at infinity, after gauge transformations.*

Proof This follows from Proposition 3.18. □

The next Lemma is the heart of the proof of the Theorem. We show that if the $\overline{\alpha}_j^t$ lose control of their norms, then they should contain some sequences, as $t \to 0$, which approach to the tangent to the slice of the moduli space $\mathfrak{M}(A)_w$, while keeping uniformly bounded norms from below. This contradicts Lemma 3.22.

Lemma 3.26 *There exists some $t_0 > 0$ such that $\big(\overline{\alpha}_j^{t_0}, b_{t_0}\big) \in \overline{D}_\delta^j \times D_\mu$ for all $j \gg i$.*

Proof Suppose the contrary, and let us fix a small $0 < \epsilon_0 \ll \delta, \mu \ll 1$. Then there are indices $\{t_m\}_m$, $t_m > 0$ and $\{i_m, j_m\}_m$, $t_m \to 0$ and $i_m, j_m \to \infty$ such that

$$\| \overline{A}_j - \overline{A}_{j_m}^{t_m} \|_{\overline{W}_{w_{j_m}}^{k+1}((K_{i_m}, g_{j_m}):\Lambda^1 \otimes \operatorname{Ad} P)} = \epsilon_0.$$

Let us put $\overline{\alpha}_j \equiv \overline{A}_j - \overline{A}_{j_m}^{t_m}$. By Lemma 3.25, the sequence $\{\overline{\alpha}_j\}_j$ are orthogonal to im d_A at infinity, and so it follows from Corollary 3.17 that the differentials of the family $\{\overline{\alpha}_j\}_j$ converge to the finite dimensional subspace T. Let us denote the projection on T by $\overline{T} \subset \overline{W}_{w_j}^{k+1}((K_i, g_j) : \Lambda^1 \otimes \operatorname{Ad} P)$, and let $\overline{\alpha}_j | \overline{T}$ be the orthogonal projection.

Then by Corollary 3.21 the estimates $\| \overline{\alpha}_j | \overline{T} \| \geq \frac{1}{2}\epsilon_0$ and $\| \overline{\alpha}_j - \overline{\alpha}_j | \overline{T} \| \leq C\epsilon_0^2$ hold, where C is independent of ϵ_0.

It follows from the property of b_t in Lemma 3.22 and positivity of $t_m > 0$ that $F_s^+\big(\overline{A}_{j_m}^{t_m}, b_{t_m}\big) = 0$ cannot happen, if we have chosen a sufficiently small $\epsilon_0 > 0$. This completes the proof of Lemma 3.26. □

Now the path may not be generic with respect to A, since one has chosen the sequence $\{A_j\}_j$ after the choice of b_t. Notice that originally possible number of A may have the one of the choices of sequences which is the Cantor set.

Thus let us choose another generic path $b'_t \in P$ which is sufficiently near b_t that it is generic with respect to all $\mathfrak{M}(M, g_j)$ for all j, and $\mathfrak{M}_{b'_t}(A)_w$ are regular for all $0 < t \leq t_0 \leq 1$ for some t_0.

Now by Lemma 2.12, let c: $[0,1]^2 \to P$ be a generic map with $c(0,t) = b_t$ and $c(1,t) = b'_t$. Thus $\mathfrak{M}_c(M, g_j)$ are smooth manifolds with corners. For each j, we have chosen A_j so that there are nonempty parameterizations A_j^t with respect to b_t. Since the moduli spaces are manifolds, this implies that their parameterizations can be extended as $A_j^{(s,t)}$ so that the equations $F_s^+\big(A_j^{(s,t)}, c(s,t)\big) = 0$ hold.

In particular there are solutions $B_j^t \equiv A_j^{(1,t)}$ with $F_s^+\big(B_j^t, b_t'\big) = 0$. By Lemma 3.26, one can find another solutions $\big(\bar{B}_j^{t_0}, b_{t_0}'\big)$ with $F_s^+\big(\bar{B}_j^{t_0}, b_{t_0}'\big) = 0$ which satisfy uniformity of norms

$$\|\bar{A}_j - \bar{B}_j^{t_0}\|_{\bar{W}_{w_j}^{k+1}(K_i, g_j)} \leq c$$

for a small constant $c > 0$. Then one can put $\bar{C}_j = \bar{B}_j^{t_0}$ and $b = b_{t_0}'$. This completes the proof of Proposition 3.24. □

3.2.10 Completion of the proof of Theorem 1.9 Now we replace b' by b and t_0 by 1 by reparameterization. We complete the proof of Theorem 1.9.

We have a Lemma concerning abstract functional analysis.

Lemma 3.27 *Let $W^{k+1}(K)$ be the Sobolev spaces on a compact $K \subset S$. Then any bounded sequence $\{u_j\}_j \subset W^{k+1}(K)$, $\|u_j\|_{W^{k+1}(K)} \leq C$ admits a convergent subsequence to u in $W^k(K)$ with $\|u\|_{W^{k+1}(K)} \leq C$.*

Proof This follows from the Rellich lemma, see Gilberg–Trudinger [8, pages 168–169]. □

In order to specify the indices i, let us write $\bar{C}_j^i$ for $\bar{C}_j$ above. Thus we have obtained a family of solutions $(\bar{C}_j^i, b)$ with $F_s^+(\bar{C}_j^i, b) = 0$ over (K_i, g_j) and $\bar{A}_j - \bar{C}_j^i \in \bar{W}_{w_j}^{k+1}((K_i, g_j) : \Lambda^1 \otimes \mathrm{Ad}\, P)$, which satisfy the uniform estimates:

$$\|\bar{A}_j - \bar{C}_j^i\|_{\bar{W}_{w_j}^{k+1}((K_i, g_j); \Lambda^1 \otimes \mathrm{Ad}\, P)} \leq C.$$

Recall that as $j \to \infty$, $g_j | K_i \to g | K_i$, where g is the metric on S.

Then for each i, using Lemma 3.27, one takes a convergent subsequence $\bar{C}_{j_m}^i$ to $\bar{C}^i$ over (K_i, g). This satisfies the equation $F_s^+(\bar{C}^i, b) = 0$ with respect to $g|K_i$, and the estimates hold

$$\|\bar{A} - \bar{C}^i\|_{\bar{W}_w^{k+1}((K_i, g); \Lambda^1 \otimes \mathrm{Ad}\, P)} \leq C.$$

Finally one takes another subsequence $\{\bar{C}^{i_m}\}_m$ which converges to a smooth connection A' over (S, g). It satisfies the equation $F_s^+(A', b) = 0$ over (S, g) and the estimate

$$\|A - A'\|_{W_w^{k+1}((S, g) : \Lambda^1 \otimes \mathrm{Ad}\, P)} \leq C$$

holds. This is the desired connection, and one has finished the proof of Theorem 1.9.

References

[1] **Ž Bižaca**, **R E Gompf**, *Elliptic surfaces and some simple exotic* $\mathbb{R}^4$*s*, J. Differential Geom. 43 (1996) 458–504 MR1412675

[2] **W Chen**, **S Kwasik**, *Symmetric homotopy K3 surfaces* arXiv:0709.2448v1

[3] **S K Donaldson**, **P B Kronheimer**, *The geometry of four-manifolds*, Oxford Mathematical Monographs, Oxford University Press (1990) MR1079726

[4] **A Floer**, *An instanton-invariant for 3–manifolds*, Comm. Math. Phys. 118 (1988) 215–240 MR956166

[5] **M H Freedman**, *The topology of four-dimensional manifolds*, J. Differential Geom. 17 (1982) 357–453 MR679066

[6] **R Friedman**, **J W Morgan**, *Complex versus differentiable classification of algebraic surfaces*, Topology Appl. 32 (1989) 135–139 MR1007985

[7] **R Friedman**, **J W Morgan**, *Smooth four-manifolds and complex surfaces*, Ergebnisse series 27, Springer, Berlin (1994) MR1288304

[8] **D Gilbarg**, **N S Trudinger**, *Elliptic partial differential equations of second order*, second edition, Grundlehren series 224, Springer, Berlin (1983) MR737190

[9] **R E Gompf**, **T S Mrowka**, *Irreducible 4–manifolds need not be complex*, Ann. of Math. (2) 138 (1993) 61–111 MR1230927

[10] **T Kato**, *ASD moduli spaces over four-manifolds with tree-like ends*, Geom. Topol. 8 (2004) 779–830 MR2087070

[11] **T Kato**, *Spectral analysis on tree like spaces from gauge theoretic view points*, from: "Discrete geometric analysis", Contemp. Math. 347, Amer. Math. Soc., Providence, RI (2004) 113–129 MR2077033

[12] **P B Kronheimer**, *Instanton invariants and flat connections on the Kummer surface*, Duke Math. J. 64 (1991) 229–241 MR1136374

[13] **T Matumoto**, *On diffeomorphisms of a* $K3$ *surface*, from: "Algebraic and topological theories (Kinosaki, 1984)", Kinokuniya, Tokyo (1986) 616–621 MR1102278

[14] **C H Taubes**, *The Seiberg–Witten invariants and symplectic forms*, Math. Res. Lett. 1 (1994) 809–822 MR1306023

[15] **C H Taubes**, *More constraints on symplectic forms from Seiberg–Witten invariants*, Math. Res. Lett. 2 (1995) 9–13 MR1312973

[16] **K K Uhlenbeck**, *Connections with* L^p *bounds on curvature*, Comm. Math. Phys. 83 (1982) 31–42 MR648356

[17] **K K Uhlenbeck**, *Removable singularities in Yang–Mills fields*, Comm. Math. Phys. 83 (1982) 11–29 MR648355

[18] **C T C Wall**, *On the orthogonal groups of unimodular quadratic forms II*, J. Reine Angew. Math. 213 (1963/1964) 122–136 MR0155798

Department of Mathematics, Faculty of Science, Kyoto University,
606-8502 Kyoto, Japan

tkato@math.kyoto-u.ac.jp

Proposed: Simon Donaldson
Seconded: Ron Stern, Rob Kirby

Received: 12 May 2006
Revised: 22 April 2008