ISRAEL JOURNAL OF MATHEMATICS **207** (2015), 793–834
DOI: 10.1007/s11856-015-1162-z

# LOCAL MEAN DIMENSION OF ASD MODULI SPACES OVER THE CYLINDER

BY

Shinichiroh Matsuo*

*Department of Mathematics, Osaka University, Toyonaka, Osaka 560-0043, Japan*
*e-mail: matsuo@math.sci.osaka-u.ac.jp*

AND

Masaki Tsukamoto*

*Department of Mathematics, Kyoto University, Kyoto 606-8502, Japan*
*e-mail: tukamoto@math.kyoto-u.ac.jp*

ABSTRACT

We study an infinite-dimensional ASD moduli space over the cylinder. Our main result is the formula of its local mean dimension. A key ingredient of the argument is the notion of non-degenerate ASD connections. We develop their deformation theory and show that there exist sufficiently many non-degenerate ASD connections by using the method of gluing infinitely many instantons.

## 1. Introduction

1.1. Main result. A fundamental object in geometric analysis is the space of solutions of a partial differential equation over a manifold. When the underlying manifold is compact and the equation is linear, a milestone result is the Atiyah–Singer index theorem; it gives a topological formula of the index of a linear elliptic equation. When the equation is non-linear, the problem is much more involved. General non-linear equations are probably too

* Shinichiroh Matsuo and Masaki Tsukamoto were supported by Grant-in-Aid for Young Scientists (B) 2580045 and 25870334, respectively, from JSPS
Received March 22, 2013 and in revised form May 13, 2014

difficult for our present technologies. But there exist some tractable non-linear equations. Most notable ones are the Yang–Mills anti-self-dual equation, (pseudo-)holomorphic curves and the Seiberg–Witten monopole equations. When the underlying manifold is compact, the moduli spaces of solutions of these equations become finite-dimensional, and it is possible to study their topological and geometric properties in detail. A lot of surprising facts on these moduli spaces have been obtained for more than 30 years. But if the underlying manifold is non-compact, the situation drastically changes. The moduli spaces can be infinite-dimensional, and traditional methods cannot be applied to such infinite-dimensional spaces. We know almost nothing about them.

When we consider linear elliptic equations over non-compact manifolds, there already exist established techniques of operator algebra to study infinite-dimensional solution spaces (Atiyah [2]). Is there any analogous method for non-linear equations? Relatively recently Gromov [14] proposed one idea. He proposed the study of the dynamics on infinite-dimensional moduli spaces. Infinite-dimensional (non-linear) spaces are very wild objects, but if we look at dynamics on them, a new structure will emerge. In order to capture information on these infinite-dimensional dynamics, Gromov introduced a new topological invariant of dynamical systems called mean dimension, which is a non-linear/$L^\infty$ analogue of von Neumann dimension. The purpose of our present paper is to examine this idea in the context of Yang–Mills theory. We will show that the new dynamical viewpoint certainly reveals a non-trivial structure of infinite-dimensional Yang–Mills moduli spaces. Probably some of our techniques can be also applied to the Seiberg–Witten equations. For the case of holomorphic curves, see Gournay [11], Tsukamoto [22, 24] and Matsuo–Tsukamoto [19]. The paper [19] is the origin of the main techniques of the present paper.

To explain our motivation more concretely, we need to recall the pioneering work of Atiyah–Hitchin–Singer [3]. The paper [3] is a starting point of the mathematical study of Yang–Mills gauge theory. One of their main results [3, Theorem 6.1] is a calculation of the dimension of a Yang–Mills moduli space by using the Atiyah–Singer index theorem. Their result can be stated as follows: Let $A$ be an $SU(2)$ connection over a compact oriented Riemannian 4-manifold. We denote its curvature by $F_A$. The connection $A$ is said to be **ASD** (**anti-self-dual**) if $F_A$ is anti-self-dual ($*F_A = -F_A$) with respect to the Hodge star $*$. Suppose that $A$ is an irreducible ASD connection and that the underlying 4-manifold is anti-self-dual (i.e., its Weyl conformal curvature is anti-self-dual) and

of positive scalar curvature. Then the number of the deformation parameters of $A$ is equal to

$$8(\text{instanton number of } A) - 3(1 - b_1).$$

Here $b_1$ is the first Betti number of the underlying 4-manifold. The "instanton number" means the second Chern number of the bundle which the connection $A$ belongs to, and it is equal to the Yang–Mills functional

$$\frac{1}{8\pi^2}\int |F_A|^2 d\text{vol}.$$

This dimension formula is the target of our work. Our main result (Theorem 1.2) is an infinite-dimensional analogue of the above formula. Although there is still much work to be done, probably our theorem is the first satisfactory result in this direction.

Let $S^3 := \{x_1^2+x_2^2+x_3^2+x_4^2 = 1\} \subset \mathbb{R}^4$ be the unit 3-sphere with the standard Riemannian metric induced by the Euclidean metric on $\mathbb{R}^4$. Let $X := \mathbb{R}\times S^3$ be the cylinder with the product metric. We will study infinite-dimensional Yang–Mills moduli spaces over $X$. The reason why we consider the cylinder $\mathbb{R}\times S^3$ is as follows: In Atiyah–Hitchin–Singer [3, Theorem 6.1] they needed the assumption that the underling 4-manifold is anti-self-dual and of positive scalar curvature. These metrical conditions were used via a certain Weitzenböck formula. In the present paper we also need to use the Weitzenböck formula. The cylinder $\mathbb{R} \times S^3$ is one of the simplest non-compact 4-manifolds which are anti-self-dual and of uniformly positive scalar curvature. So this is the first non-trivial case to examine our ideas.

Let $E := X \times SU(2)$ be the product principal $SU(2)$ bundle. (Every principal $SU(2)$ bundle on $X$ is isomorphic to the product bundle $E$.) Let $A$ be a connection on $E$. Its curvature $F_A$ is a 2-form valued in the adjoint bundle $\mathrm{ad}E = X \times su(2)$. So it gives a linear map:

$$F_A(p) : \Lambda^2(T_pX) \to su(2) \quad (\forall p \in X).$$

Let $|F_A(p)|_{\mathrm{op}}$ be the operator norm of this linear map, and let $\|F_A\|_{\mathrm{op}}$ be the supremum of $|F_A(p)|_{\mathrm{op}}$ over $p \in X$. The explicit formula is as follows: Let $p\in X$, and let $(x_1, x_2, x_3, x_4)$ be the normal coordinate system of $X$ centered at $p$. We suppose that the curvature $F_A$ is expressed by $F_A = \sum_{1\le i<j\le 4} F_{ij}dx_i \wedge dx_j$

around the point $p$. Then the norm $|F_A(p)|_{\mathrm{op}}$ is equal to

$$\sup\left\{\left|\sum_{1\le i<j\le 4} a_{ij}F_{ij}(p)\right| \,|\, a_{ij}\in\mathbb{R},\ \sum_{1\le i<j\le 4} a_{ij}^2 = 1\right\}.$$

Here the Lie algebra $su(2) = \{X \in M_2(\mathbb{C}) | X + X^* = 0,\ \mathrm{tr}(X) = 0\}$ is endowed with the inner product $\langle X, Y\rangle = -\mathrm{tr}(XY)$. In this paper we also use the Euclidean norm $|F_A(p)|$ defined by

$$|F_A(p)|^2 := \sum_{1\le i<j\le 4} |F_{ij}(p)|^2. \tag{1}$$

For a subset $U \subset X$ we denote by $\|F_A\|_{L^\infty(U)}$ the essential supremum of $|F_A(p)|$ over $p \in U$.

For a non-negative number $d$ we define $\mathcal{M}_d$ as the space of the gauge equivalence classes of ASD connections $A$ on $E$ satisfying

$$\|F_A\|_{\mathrm{op}} \le d.$$

This space is endowed with the topology of $C^\infty$-convergence over compact subsets: A sequence $[A_n]$ converges to $[A]$ in $\mathcal{M}_d$ if and only if there exists a sequence of gauge transformations $g_n : E \to E$ such that $g_n(A_n)$ converges to $A$ in $C^\infty$ over every compact subset. From the Uhlenbeck compactness ([26, 27]), the space $\mathcal{M}_d$ is compact and metrizable. The above condition $\|F_A\|_{\mathrm{op}} \le d$ is motivated by the notion of Brody curves (Brody [5]) in Nevanlinna theory. Note that the norm $\|F_A\|_{\mathrm{op}}$ does not dominate the $L^2$-norm of $F_A$. So the $L^2$-norm of the curvature of $[A] \in \mathcal{M}_d$ can be infinite.

The space $\mathcal{M}_d$ becomes a dynamical system with respect to the following natural $\mathbb{R}$-action: $\mathbb{R}$ acts on $X = \mathbb{R} \times S^3$ by $s(t, \theta) := (t + s, \theta)$. This action is lifted to the action on $E = X \times SU(2)$ by $s(t, \theta, u) := (t + s, \theta, u)$. Then $\mathbb{R}$ continuously acts on $\mathcal{M}_d$ by $s[A] := [s^*(A)]$ where $s^*(A)$ is the pull-back of $A$ by $s : E \to E$. Our main subject is the study of this dynamics on $\mathcal{M}_d$. Let's look at the example:

*Example 1.1:* If $d < 1$ then $\mathcal{M}_d$ is equal to the one-point space. The only one element is the gauge equivalence class of the trivial flat connection. (This fact is proved in [25]. The threshold value $d = 1$ is different from the value given in [25]. This is because the norm on $su(2)$ in the present paper is different from the norm in [25] by the multiple factor $\sqrt{2}$.) If $d = 1$ then the space $\mathcal{M}_1$ contains a non-trivial element: We define an $SU(2)$ ASD connection $A$ (discovered by

Belavin–Polyakov–Schwartz–Tyupkin [4]) over the Euclidean space $\mathbb{R}^4$ by

$$A(x) := \frac{1}{1+|x|^2}\left\{ \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix} (x_1dx_2 - x_2dx_1 - x_3dx_4 + x_4dx_3) \right.$$
$$+ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_1dx_3 - x_3dx_1 + x_2dx_4 - x_4dx_2)$$
$$\left. + \begin{pmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix} (x_1dx_4 - x_4dx_1 - x_2dx_3 + x_3dx_2) \right\}.$$

Let $I$ be the pull-back of $A$ by the conformal map

$$\mathbb{R} \times S^3 \to \mathbb{R}^4 \setminus \{0\}, \quad (t,\theta) \mapsto e^t\theta.$$

Then $I$ is an ASD connection on $E = X \times SU(2)$ with

$$|F_I(t,\theta)|_{\mathrm{op}} = \frac{4}{(e^t + e^{-t})^2}, \quad \|F_I\|_{\mathrm{op}} = 1.$$

Hence $[I]$ is contained in $\mathcal{M}_1$. So $\mathcal{M}_1$ contains a flat connection and the $\mathbb{R}$-orbit of $[I]$. The authors don't know whether it contains other elements or not.

Therefore $\mathcal{M}_d$ is trivial for $d < 1$, and possibly a simple space for $d = 1$. On the other hand, we will see later that the space $\mathcal{M}_d$ is infinite-dimensional for $d > 1$ (Remark 1.12). Moreover, its topological entropy (as a topological dynamical system) is also infinite. So $\mathcal{M}_d$ ($d > 1$) is a very large dynamical system. Mean dimension is exactly an invariant introduced for the study of this kind of infinite-dimensional and infinite entropy systems. So we want to calculate the mean dimension of $\mathcal{M}_d$. But our present technology is a little inadequate for the complete calculation of the mean dimension of $\mathcal{M}_d$. So we calculate the local mean dimension of $\mathcal{M}_d$ instead of mean dimension itself. Local mean dimension is a variant of mean dimension introduced by [18]. For each point $[A] \in \mathcal{M}_d$ we have the non-negative number $\dim_{[A]}(\mathcal{M}_d : \mathbb{R})$ called the local mean dimension of $\mathcal{M}_d$ at $[A]$. We define $\dim_{loc}(\mathcal{M}_d : \mathbb{R})$ as the supremum of $\dim_{[A]}(\mathcal{M}_d : \mathbb{R})$ over $[A] \in \mathcal{M}_d$. Mean dimension and local mean dimension are topological invariants of dynamical systems which count "dimension averaged by a group action" in certain ways. We review their definitions in Section 2.

Let $A$ be a connection on $E$. We define the **energy density** $\rho(A)$ by

$$\rho(A) := \lim_{T\to+\infty} \left( \frac{1}{8\pi^2 T} \sup_{t\in\mathbb{R}} \int_{(t,t+T)\times S^3} |F_A|^2 d\mathrm{vol} \right).$$

Here $|F_A|$ is the Euclidean norm defined in (1). This limit always exists because we have the natural sub-additivity:

$$\sup_{t\in\mathbb{R}} \int_{(t,t+T_1+T_2)\times S^3} |F_A|^2 d\mathrm{vol} \leq \sup_{t\in\mathbb{R}} \int_{(t,t+T_1)\times S^3} |F_A|^2 d\mathrm{vol} + \sup_{t\in\mathbb{R}} \int_{(t,t+T_2)\times S^3} |F_A|^2 d\mathrm{vol}.$$

The energy density $\rho(A)$ is an "averaged" instanton number, which was first introduced in [18]. It is zero for finite energy ASD connections. So it becomes meaningful only for infinite energy ones. We define $\rho(d)$ as the supremum of $\rho(A)$ over all $[A] \in \mathcal{M}_d$; $\rho(d)$ is a non-decreasing function in $d$. It is zero for $d < 1$ (Example 1.1). We will see later that $\rho(d)$ is positive for $d > 1$ (Remark 1.12) and that it goes to infinity as $d \to \infty$ (Example 1.6).

Let $\mathcal{D} \subset [0, +\infty)$ be the set of left-discontinuous points of $\rho(d)$:

$$\mathcal{D} = \{d \in [0, +\infty) | \lim_{\varepsilon\to+0} \rho(d-\varepsilon) \neq \rho(d)\}.$$

Since $\rho$ is monotone, the set $\mathcal{D}$ is at most countable. (Indeed we don't know whether it is empty or not.) Our main result is the following theorem.

Theorem 1.2: *For any $d \in [0, +\infty) \setminus \mathcal{D}$,*

$$\dim_{loc}(\mathcal{M}_d : \mathbb{R}) = 8\rho(d).$$

Since $\mathcal{D}$ is at most countable, we get the formula of the local mean dimension of $\mathcal{M}_d$ for almost every $d \geq 0$. Our calculation of the local mean dimension $\dim_{loc}(\mathcal{M}_d : \mathbb{R})$ started in [18]. Theorem 1.2 is an almost satisfactory answer (at least for the authors) to this problem. A next important problem is to calculate the mean dimension (not the local one) of $\mathcal{M}_d$. Another interesting problem is a detailed study of the function $\rho(d)$. Is it continuous? The authors hope to return to these questions in the future.

Theorem 1.2 is strongly analogous to the main theorem of Matsuo–Tsukamoto [19] concerning the mean dimension of a system of holomorphic curves. Gromov [14, p. 332] wrote that one of his motivations is to "set some branches of the Nevanlinna theory into a dynamical casting". Hopefully our paper is a step toward the realization of his vision. Indeed our definition of the energy density $\rho(A)$ is motivated by the idea of the Nevanlinna–Shimizu–Ahlfors characteristic function in the holomorphic curve theory; for a holomorphic curve $f : \mathbb{C} \to \mathbb{C}P^N$

its characteristic function $T(r, f)$ is defined by

$$T(r, f) = \int_1^r \frac{dt}{t} \int_{|z|<t} |df|^2 dxdy \quad (z = x + \sqrt{-1}y \in \mathbb{C}).$$

The main theme of Nevanlinna theory is the asymptotics of $T(r, f)$ as $r \to \infty$. So, roughly speaking, the study of the asymptotic energy distribution is one of the most important ingredients. The authors have been trying to find a corresponding idea in Yang–Mills theory. The energy density $\rho(A)$ came from this attempt.

*Remark 1.3:* Some readers might feel that the operator norm $\|F_A\|_{\mathrm{op}}$ used in the definition of $\mathcal{M}_d$ seems strange. This choice leads us to a very satisfactory result, but we will briefly discuss another possibility in the Appendix.

1.2. Non-degenerate ASD connections. The following notion is very important in the argument of the paper:

*Definition 1.4:* Let $[A] \in \mathcal{M}_d$ $(d \geq 0)$. $A$ is said to be **non-degenerate** if the closure of the $\mathbb{R}$-orbit of $[A]$ in $\mathcal{M}_d$ does not contain a gauge equivalence class of a flat connection.

This definition is motivated by the classical work of Yosida [28] in complex analysis. Yosida studied a similar non-degeneracy condition for meromorphic functions $f : \mathbb{C} \to \mathbb{C}P^1$. (He used the terminology "meromorphic functions of first category".) Eremenko [8, Section 4] discussed it for holomorphic curves $f : \mathbb{C} \to \mathbb{C}P^N$, and Gromov [14, p. 399] studied a similar condition for more general holomorphic maps.

*Example 1.5:* Let $A$ be an **instanton**, i.e., an ASD connection on $E$ with finite energy

$$\int_X |F_A|^2 d\mathrm{vol} < +\infty.$$

Then $s[A]$ converges to a gauge equivalence class of a flat connection when $s \to \pm\infty$ (Donaldson [6, Chapter 4, Proposition 4.3]). So $A$ is a degenerate (i.e., not non-degenerate) ASD connection.

*Example 1.6:* An ASD connection $A$ on $E$ is said to be **periodic** ([18]) if there exist $T > 0$, a principal $SU(2)$ bundle $F$ over $(\mathbb{R}/T\mathbb{Z}) \times S^3$ and an ASD connection $B$ on $F$ such that $(E, A)$ is isomorphic to the pull-back $(\pi^*(F), \pi^*(B))$.

Here $\pi : \mathbb{R} \times S^3 \to (\mathbb{R}/T\mathbb{Z}) \times S^3$ is the natural projection. If $A$ is periodic, then the energy density $\rho(A)$ is given by

$$\rho(A) = c_2(F)/T.$$

If $A$ is periodic and non-flat, then $A$ is non-degenerate. By Taubes [20], every principal $SU(2)$ bundle $F$ on $(\mathbb{R}/T\mathbb{Z}) \times S^3$ with $c_2(F) \geq 0$ admits an ASD connection. Therefore we have a lot of periodic ASD connections. From this fact we can easily see that the function $\rho(d)$ introduced in the previous subsection goes to infinity as $d \to \infty$.

LEMMA 1.7: *Let $[A] \in \mathcal{M}_d$. $A$ is non-degenerate if and only if there exist $\delta > 0$ and $T > 0$ such that for any interval $(\alpha, \beta) \subset \mathbb{R}$ of length $T$ we have*

$$\|F_A\|_{L^\infty((\alpha,\beta)\times S^3)} \geq \delta. \tag{2}$$

*Proof.* This is a Yang–Mills analogue of the result of Yosida [28, Theorem 4]. Suppose that $A$ does not satisfy (2) for $T = 1$. Then there exist $\{\alpha_n\}_{n\geq 1} \subset \mathbb{R}$ such that $\|F_A\|_{L^\infty((\alpha_n,\alpha_n+1)\times S^3)} < 1/n$. By choosing a subsequence we can assume that $\alpha_n[A]$ converges to $[B]$ in $\mathcal{M}_d$. Then $F_B = 0$ over $(0,1) \times S^3$. By the unique continuation, $F_B = 0$ all over $X$. Hence $B$ is flat and $A$ is degenerate.

Suppose the above condition (2) holds for some $\delta > 0$ and $T > 0$. Then any element $[B]$ in the closure of the $\mathbb{R}$-orbit of $[A]$ satisfies $\|F_B\|_{L^\infty((\alpha,\beta)\times S^3)} \geq \delta$ for every interval $(\alpha, \beta) \subset \mathbb{R}$ of length $T$. Hence $B$ is not flat.

Note that the above argument also proves the following: $[A]$ is non-degenerate if and only if for any $T > 0$ there exists $\delta > 0$ such that for any interval $(\alpha, \beta) \subset \mathbb{R}$ of length $T$ we have $\|F_A\|_{L^\infty((\alpha,\beta)\times S^3)} \geq \delta$. ∎

*Remark 1.8:* By the same argument we can prove the following: $[A] \in \mathcal{M}_d$ is non-degenerate if and only if there exist $\delta > 0$ and $T > 0$ such that for any interval $(\alpha, \beta) \subset \mathbb{R}$ of length $T$ we have

$$\|F_A\|_{L^2((\alpha,\beta)\times S^3)} \geq \delta.$$

In particular if $[A] \in \mathcal{M}_d$ is non-degenerate then its energy density $\rho(A)$ is positive.

The following Theorem 1.9 is proved in [18, Theorem 1.2]. (The paper [18] adopts a little different setting. So we explain how to deduce this result from [18] in the Appendix.)

THEOREM 1.9: *For any* $[A] \in \mathcal{M}_d$,

$$\dim_{[A]}(\mathcal{M}_d : \mathbb{R}) \leq 8\rho(A).$$

*Hence*

$$\dim_{loc}(\mathcal{M}_d : \mathbb{R}) = \sup_{[A]\in\mathcal{M}_d} \dim_{[A]}(\mathcal{M}_d : \mathbb{R}) \leq 8\rho(d).$$

The lower bound on the local mean dimension is given by using the next two theorems.

THEOREM 1.10: *Let* $A$ *be a non-degenerate ASD connection on* $E$ *with* $\|F_A\|_{\mathrm{op}} < d$. *Then*

$$\dim_{[A]}(\mathcal{M}_d : \mathbb{R}) = 8\rho(A).$$

In this theorem the strict inequality condition $\|F_A\|_{\mathrm{op}} < d$ is purely technical. The point is the non-degeneracy assumption. This makes the situation simpler. It is more difficult to study the local structure of $\mathcal{M}_d$ around degenerate ASD connections. We postpone it to a future paper. In the present paper we bypass it by using the following theorem.

THEOREM 1.11: *Suppose* $d > 1$, *and let* $A$ *be an ASD connection on* $E$ *with* $\|F_A\|_{\mathrm{op}} < d$. *For any* $\varepsilon > 0$ *there exists a non-degenerate ASD connection* $\tilde{A}$ *on* $E$ *satisfying*

$$\|F(\tilde{A})\|_{\mathrm{op}} < d, \quad \rho(\tilde{A}) > \rho(A) - \varepsilon.$$

Roughly speaking, this theorem means that we can replace a degenerate ASD connection by a non-degenerate one without losing energy. In the above statement we supposed $d > 1$ because there does not exist a non-flat ASD connection $A$ on $E$ satisfying $\|F_A\|_{\mathrm{op}} < 1$ (Example 1.1).

The main task of the paper is to prove Theorems 1.10 and 1.11. Here we prove the main theorem by assuming them:

*Proof of Theorem 1.2 (assuming Theorems 1.10 and 1.11).* We always have the upper bound $\dim_{loc}(\mathcal{M}_d : \mathbb{R}) \leq 8\rho(d)$ by Theorem 1.9. So the problem is the lower bound.

Let $\rho_0(d)$ be the supremum of $\rho(A)$ over $[A] \in \mathcal{M}_d$ satisfying $\|F_A\|_{\mathrm{op}} < d$. Obviously $\rho_0(d) \leq \rho(d)$. Then

$$\dim_{loc}(\mathcal{M}_d : \mathbb{R}) \geq 8\rho_0(d). \tag{3}$$

This is proved as follows:

Case 1: Suppose $d \leq 1$. Then the condition $\|F_A\|_{\mathrm{op}} < d$ implies $F_A \equiv 0$. (See Example 1.1.) Hence $\rho_0(d) = 0$ and the above (3) trivially holds.

Case 2: Suppose $d > 1$. Take $[A] \in \mathcal{M}_d$ with $\|F_A\|_{\mathrm{op}} < d$. For any $\varepsilon > 0$ there exists a non-degenerate ASD connection $\tilde{A}$ on $E$ satisfying $\|F(\tilde{A})\|_{\mathrm{op}} < d$ and $\rho(\tilde{A}) > \rho(A) - \varepsilon$ (Theorem 1.11). By applying Theorem 1.10 to $\tilde{A}$,

$$\dim_{loc}(\mathcal{M}_d : \mathbb{R}) \geq \dim_{[\tilde{A}]}(\mathcal{M}_d : \mathbb{R}) = 8\rho(\tilde{A}) > 8(\rho(A) - \varepsilon).$$

Since $\varepsilon > 0$ is arbitrary, $\dim_{loc}(\mathcal{M}_d : \mathbb{R}) \geq 8\rho(A)$. Taking the supremum over $A$, we get the above (3).

For any $\varepsilon > 0$, we have $\rho(d - \varepsilon) \leq \rho_0(d) \leq \rho(d)$. Hence if $\rho$ is left-continuous at $d$ (i.e., $d \notin \mathcal{D}$), then we have $\rho_0(d) = \rho(d)$. Therefore

$$\dim_{loc}(\mathcal{M}_d : \mathbb{R}) \geq 8\rho(d) \quad (d \in [0, +\infty) \setminus \mathcal{D}). \qquad \blacksquare$$

*Remark 1.12:* Let $d > 1$. By applying Theorem 1.11 to a flat connection, we can conclude that $\mathcal{M}_d$ always contains a non-degenerate ASD connection. (Indeed $\mathcal{M}_d$ always contains a non-flat periodic ASD connection. See Remark 6.3.) Since the energy density of a non-degenerate ASD connection is positive (Remark 1.8), the function $\rho(d)$ is positive for $d > 1$. Moreover, by Theorem 1.10, the local mean dimension of $\mathcal{M}_d$ is also positive for $d > 1$. In particular, $\mathcal{M}_d$ is infinite-dimensional for $d > 1$.

1.3. Ideas of the proofs. We explain the ideas of the proofs of Theorems 1.10 and 1.11.

The basic idea of the proof of Theorem 1.10 is a deformation theory. Let $A$ be a non-degenerate ASD connection on $E$ satisfying $\|F_A\|_{\mathrm{op}} < d$. Let $H^1_A$ be the Banach space of $a \in \Omega^1(\mathrm{ad}E)$ satisfying

$$d_A^* a = d_A^+ a = 0, \quad \|a\|_{L^\infty(X)} < \infty.$$

Here $d_A^*$ is the formal adjoint of $d_A : \Omega^0(\mathrm{ad}E) \to \Omega^1(\mathrm{ad}E)$, and $d_A^+$ is the self-adjoint part of $d_A : \Omega^1(\mathrm{ad}E) \to \Omega^2(\mathrm{ad}E)$. For each $a \in H^1_A$ the connection $A + a$ is almost ASD: $F^+(A + a) = O(a^2)$. Therefore there exists a small $R > 0$ such that for each $a \in B_R(H^1_A)$ (the $R$-ball with respect to $\|\cdot\|_{L^\infty(X)}$) we can construct a small perturbation $a'$ of $a$ satisfying $F^+(A + a') = 0$. So we get a deformation map:

$$B_R(H^1_A) \to \mathcal{M}_d, \quad a \mapsto [A + a']. \tag{4}$$

We study the local mean dimension of $\mathcal{M}_d$ through this map.

A construction of the map (4) does not require the non-degeneracy condition of $A$. But a further study of (4) requires it. We need to compare the distances of both sides of (4). $\mathcal{M}_d$ is a quotient space by gauge transformations. Hence its metric structure is more complicated than that of $B_R(H^1_A)$. For example, even if $a, b \in B_R(H^1_A)$ are not close to each other, the points $[A+a']$ and $[A+b']$ might be very close to each other in $\mathcal{M}_d$. We need the non-degeneracy condition for addressing this problem. This is a technical issue. So here we do not go into detail but just point out that the above map (4) becomes injective if $R \ll 1$ and $A$ is non-degenerate. (This injectivity is not enough for our main purpose. The result we need is Lemma 5.5, and it is based on the study of the Coulomb gauge condition in Section 3.)

Assume that we have a good understanding of the deformation map (4). A next problem is the study of the Banach space $H^1_A$. We investigate a structure of finite-dimensional linear subspaces of $H^1_A$. ($H^1_A$ itself is infinite-dimensional.) We need the following result (Proposition 4.1): For any interval $(\alpha, \beta) \subset \mathbb{R}$ of length $> 2$ there exists a finite-dimensional linear subspace $V \subset H^1_A$ such that

$$\dim V \geq \frac{1}{\pi^2} \int_{(\alpha,\beta)\times S^3} |F_A|^2 d\mathrm{vol} - \mathrm{const}_A,$$
$$\forall a \in V : \|a\|_{L^\infty(X)} \leq 2\|a\|_{L^\infty((\alpha,\beta)\times S^3)}.$$

The energy density $\rho(A)$ comes into our argument through the first condition of $V$. The second condition means that essentially all the information of $a \in V$ is contained in the region $(\alpha, \beta) \times S^3$. A main ingredient of the proof of this result is the Atiyah–Singer index theorem. Combining this knowledge on $H^1_A$ with the study of the deformation map (4), we can prove Theorem 1.10. The proof is finished in Section 5.2.

Next we explain the idea of the proof of Theorem 1.11. Suppose $d > 1$ and that $A$ is a degenerate ASD connection on $E$ with $\|F_A\|_{\mathrm{op}} < d$. We want to replace $A$ with a non-degenerate one. The idea is gluing instantons. Lemma 1.7 implies that $A$ has a region where the curvature $F_A$ is very small. We glue an instanton $I$ (described in Example 1.1) to $A$ over such a "degenerate region". $A$ has infinitely many degenerate regions. So we need to glue infinitely many instantons to $A$.

More precisely the argument goes as follows: Let $0 < \delta \ll 1$ and $T \gg 1$. We define $J \subset \mathbb{Z}$ as the set of $n \in \mathbb{Z}$ such that $|F_A| < \delta$ over $[nT, (n+1)T] \times S^3$. Since $A$ is degenerate, the set $J$ is infinite. For each $n \in J$ we glue (an appropriate

translation of) the instanton $I$ to $A$ over the region $[nT,(n+1)T]\times S^3$. If we choose $\delta$ sufficiently small and $T$ sufficiently large, then the resulting new ASD connection $\tilde{A}$ becomes non-degenerate and satisfies $\|F(\tilde{A})\|_{\mathrm{op}}<d$. Moreover, roughly speaking, gluing instantons increases the energy of connections. So we have $\rho(\tilde{A})>\rho(A)-\varepsilon$.

The paper [19] is the origin of our idea to use the deformation theory of non-degenerate objects and gluing infinitely many instantons. In [19] we study the mean dimension of the system of Brody curves (holomorphic 1-Lipschitz maps) $f:\mathbb{C}\to\mathbb{C}P^N$ by developing the deformation theory of non-degenerate Brody curves and gluing technique of infinitely many rational curves. After the authors wrote the paper [19], they felt that the ideas of [19] have a wide applicability beyond the holomorphic curve theory. The second main purpose of the present paper is to show that a basic structure of the argument in [19] is certainly flexible and can be also applied to Yang–Mills theory. The authors are satisfied with the result.

The main difference between the case of Brody curves and Yang–Mills theory is the presence of gauge transformations. A substantial part of the present paper is devoted to the study of the method to deal with gauge transformations. (The technique of perturbing Hermitian metrics described in [19, Section 4.2] might have a flavor of gauge fixing. But it is much simpler.) At least for our present technology, the Yang–Mills case is more involved than Brody curves. Another, relatively minor, difference is the techniques of gluing. The gluing construction in [19] is more elementary than that of the present paper. The reason is that for meromorphic functions $f$ and $g$ in $\mathbb{C}$ we have a natural definition of their sum $f+g$. But we don't have such a definition for the "sum" of ASD connections.

1.4. Organization of the paper. Section 2 is a review of mean dimension and local mean dimension. Section 3 is devoted to the study of the Coulomb gauge condition. In Section 4 we study the Banach space $H^1_A$. In Section 5 we develop the deformation theory of non-degenerate ASD connections and prove Theorem 1.10. In Section 6 we study the gluing method and prove Theorem 1.11. In the Appendix we investigate another definition of the ASD moduli space.

1.5. Notations. • In most of the argument the variable $t$ means the natural projection $t:\mathbb{R}\times S^3\to\mathbb{R}$.

• The value of $d$ (which is used to define $\mathcal{M}_d$) is fixed in the rest of this paper (except for the Appendix). So we usually omit to write the dependence on $d$. We adopt the following notation:

*Notation 1.13:* For two quantities $x$ and $y$ we write

$$x \lesssim y$$

if there exists a positive constant $C(d)$ which depends only on $d$ such that $x \leq C(d)y$. Let $A$ be a connection on $E$. We also use the following notation:

$$x \lesssim_A y.$$

This means that there exists a positive constant $C(d, A)$ which depends only on $d$ and $A$ such that $x \leq C(d, A)y$. The notation $x \lesssim_A y$ is used in Sections 3, 4 and 5 where we fix a connection $A$ in most of the argument.

• Let $A$ be a connection on $E$. Let $k \geq 0$ be an integer, and let $p \geq 1$. For $\xi \in \Omega^i(\mathrm{ad}E)$ $(0 \leq i \leq 4)$ and a subset $U \subset X$, we define a norm $\|\xi\|_{L^p_{k,A}(U)}$ by

$$\|\xi\|_{L^p_{k,A}(U)} := \left( \sum_{j=0}^{k} \|\nabla_A^j \xi\|^p_{L^p(U)} \right)^{1/p}.$$

For $\alpha < \beta$ we often denote the norm $\|\xi\|_{L^p_{k,A}((\alpha,\beta)\times S^3)}$ by $\|\xi\|_{L^p_{k,A}(\alpha<t<\beta)}$.

## 2. Review of mean dimension and local mean dimension

In this section we review mean dimension and local mean dimension. Mean dimension was introduced by Gromov [14]. Lindenstrauss–Weiss [16] and Lindenstrauss [15] also gave fundamental contributions to the basics of this invariant. Local mean dimension was introduced in [18].

Let $(M, \mathrm{dist})$ be a compact metric space (dist is a distance function on $M$). Let $N$ be a topological space, and let $f : M \to N$ be a continuous map. For $\varepsilon > 0$, $f$ is called an $\varepsilon$-embedding if $\mathrm{Diam} f^{-1}(y) \leq \varepsilon$ for all $y \in N$. Here $\mathrm{Diam} f^{-1}(y)$ is the supremum of $\mathrm{dist}(x_1, x_2)$ over all $x_1$ and $x_2$ in the fiber $f^{-1}(y)$. Let $\mathrm{Widim}_\varepsilon(M, \mathrm{dist})$ be the minimum integer $n \geq 0$ such that there exist an $n$-dimensional polyhedron $P$ and an $\varepsilon$-embedding $f : M \to P$. The topological dimension $\dim M$ is equal to the limit of $\mathrm{Widim}_\varepsilon(M, \mathrm{dist})$ as $\varepsilon \to 0$.

The following important example was given in [14, p. 333]. This will be used in Section 5. The detailed proofs are given in Gournay [13, Lemma 2.5] and Tsukamoto [24, Appendix].

*Example 2.1:* Let $(V, \|\cdot\|)$ be a finite-dimensional Banach space. Let $B_r(V)$ be the closed ball of radius $r > 0$ centered at the origin. Then

$$\mathrm{Widim}_\varepsilon(B_r(V), \|\cdot\|) = \dim V \quad (0 < \varepsilon < r).$$

Suppose that the Lie group $\mathbb{R}$ continuously acts on a compact metric space $(M, \mathrm{dist})$. For a subset $\Omega \subset \mathbb{R}$, we define a new distance $\mathrm{dist}_\Omega$ on $M$ by $\mathrm{dist}_\Omega(x, y) := \sup_{a\in\Omega} \mathrm{dist}(a.x, a.y)$ $(x, y \in M)$. We define the **mean dimension** $\dim(M : \mathbb{R})$ by

$$\dim(M : \mathbb{R}) := \lim_{\varepsilon\to 0} \left( \lim_{T\to+\infty} \frac{\mathrm{Widim}_\varepsilon(M, \mathrm{dist}_{(0,T)})}{T} \right).$$

This limit always exists because we have the following sub-additivity:

$$\mathrm{Widim}_\varepsilon(M, \mathrm{dist}_{(0,T_1+T_2)}) \le \mathrm{Widim}_\varepsilon(M, \mathrm{dist}_{(0,T_1)}) + \mathrm{Widim}_\varepsilon(M, \mathrm{dist}_{(0,T_2)}).$$

The mean dimension $\dim(M : \mathbb{R})$ is a topological invariant. (This means that its value is independent of the choice of a distance function compatible with the topology.) If $M$ is finite-dimensional, then the mean dimension $\dim(M : \mathbb{R})$ is equal to 0.

Let $N \subset M$ be a closed subset. The function

$$T \mapsto \sup_{a\in\mathbb{R}} \mathrm{Widim}_\varepsilon(N, \mathrm{dist}_{(a,a+T)})$$

is also sub-additive. So we can define the following quantity:

$$\dim(N : \mathbb{R}) := \lim_{\varepsilon\to 0} \left( \lim_{T\to+\infty} \frac{\sup_{a\in\mathbb{R}} \mathrm{Widim}_\varepsilon(N, \mathrm{dist}_{(a,a+T)})}{T} \right).$$

For $r > 0$ and $p \in M$ we define $B_r(p)_\mathbb{R}$ as the set of points $x \in M$ satisfying $\mathrm{dist}_\mathbb{R}(p, x) \le r$. (Note that $\mathrm{dist}_\mathbb{R}(p, x) \le r$ means $\mathrm{dist}(a.p, a.x) \le r$ for all $a \in \mathbb{R}$.) We define the **local mean dimension** $\dim_p(M : \mathbb{R})$ **at** $p$ by

$$\dim_p(M : \mathbb{R}) := \lim_{r\to 0} \dim(B_r(p)_\mathbb{R} : \mathbb{R}).$$

We define the **local mean dimension** $\dim_{loc}(M : \mathbb{R})$ by

$$\dim_{loc}(M : \mathbb{R}) := \sup_{p\in M} \dim_p(M : \mathbb{R}).$$

$\dim_p(M : \mathbb{R})$ and $\dim_{loc}(M : \mathbb{R})$ are topological invariants of the dynamical system $M$. We always have

$$\dim_p(M : \mathbb{R}) \leq \dim_{loc}(M : \mathbb{R}) \leq \dim(M : \mathbb{R}).$$

In this paper we define mean dimension only for $\mathbb{R}$-actions. But we can define it for more general group actions. Gromov [14] defined mean dimension for actions of amenable groups. The most basic example is the natural $\mathbb{Z}$-action (shift action) on the infinite-dimensional cube

$$[0,1]^{\mathbb{Z}} := \cdots \times [0,1] \times [0,1] \times [0,1] \times \cdots .$$

Its mean dimension and local mean dimension are given by

$$\dim_0([0,1]^{\mathbb{Z}} : \mathbb{Z}) = \dim_{loc}([0,1]^{\mathbb{Z}} : \mathbb{Z}) = \dim([0,1]^{\mathbb{Z}} : \mathbb{Z}) = 1.$$

Here $0 = (\dots, 0, 0, 0, \dots) \in [0,1]^{\mathbb{Z}}$. We don't need this result in this paper. So we omit the detail. The detailed explanations can be found in Lindenstrauss–Weiss [16, Proposition 3.3] and [18, Example 2.9].

## 3. Coulomb gauge

In this section we study a gauge fixing condition. This is a technical step toward the proof of Theorem 1.10. The ASD equation is not elliptic and admits a large symmetry of gauge transformations. So in the standard Yang–Mills theory we introduce the Coulomb gauge condition in order to break the gauge symmetry and get the ellipticity of the equation. In our situation the gauge fixing seems more involved than in the standard argument. A difficulty lies in the point that we need to consider all gauge transformations $g : E \to E$ (without any asymptotic condition at the end) and that they don't form a Banach Lie group. The main result of this section is Proposition 3.6. But its statement is not simple. Probably Corollary 3.7 is easier to understand. So it might be helpful for some readers to look at Corollary 3.7 before reading the proof of Proposition 3.6.

The next lemma is proved in [18, Corollary 6.3]. This is crucial for our argument.

Lemma 3.1: *If $A$ is a non-flat ASD connection on $E$ satisfying $\|F_A\|_{\mathrm{op}} < \infty$, then $A$ is irreducible. (Recall that $A$ is said to be reducible if there is a gauge*

*transformation $g \neq \pm 1$ satisfying $g(A) = A$; $A$ is said to be irreducible if $A$ is not reducible.)*

In the rest of this section we always suppose that $A$ is a non-degenerate ASD connection on $E$ satisfying $\|F_A\|_{\mathrm{op}} \leq d$. The next lemma shows crucial properties of non-degenerate ASD connections.

LEMMA 3.2: (i) *For any $s \in \mathbb{R}$ and any $u \in \Omega^0(\mathrm{ad}E)$,*

$$\int_{s<t<s+1} |u|^2 d\mathrm{vol} \leq C_1(A) \int_{s<t<s+1} |d_A u|^2 d\mathrm{vol}.$$

(ii) *For any $s \in \mathbb{R}$ and any gauge transformation $g : E \to E$,*

$$\min(\|g-1\|_{L^\infty(s<t<s+1)}, \|g+1\|_{L^\infty(s<t<s+1)}) \leq C_2(A)\|d_A g\|_{L^\infty(s<t<s+1)}.$$

*We will abbreviate the left-hand-side to $\min_\pm \|g \pm 1\|_{L^\infty(s<t<s+1)}$.*

*Proof.* (i) Suppose that the statement is false. Then there exist $s_n \in \mathbb{R}$ and $u_n \in \Omega^0(\mathrm{ad}E)$ satisfying

$$1 = \int_{s_n<t<s_n+1} |u_n|^2 d\mathrm{vol} > n \int_{s_n<t<s_n+1} |d_A u_n|^2 d\mathrm{vol}.$$

Set $v_n := s_n^*(u_n)$ and $A_n := s_n^*(A)$ (the pull-backs by $s_n : E \to E$). Then

$$1 = \int_{0<t<1} |v_n|^2 d\mathrm{vol} > n \int_{0<t<1} |d_{A_n} v_n|^2 d\mathrm{vol}.$$

Since $\mathcal{M}_d$ is compact, there exist a sequence of natural numbers $n_1 < n_2 < n_3 < \cdots$ and gauge transformations $g_k : E \to E$ $(k \geq 1)$ such that $B_k := g_k(A_{n_k})$ converges to some $B$ in $C^\infty$ over every compact subset of $X$. Since $A$ is non-degenerate, $B$ is not flat and hence irreducible by Lemma 3.1. Set $w_k := g_k(v_{n_k})$. Then

$$1 = \int_{0<t<1} |w_k|^2 d\mathrm{vol} > n_k \int_{0<t<1} |d_{B_k} w_k|^2 d\mathrm{vol}.$$

Since

$$d_B w_k = d_{B_k} w_k + [B - B_k, w_k],$$

the sequence $\{w_k\}$ is bounded in $L^2_{1,B}((0,1) \times S^3)$. Hence, by choosing a subsequence, we can assume that $w_k$ weakly converges to some $w$ in $L^2_{1,B}((0,1) \times S^3)$. We have $\|w\|_{L^2(0<t<1)} = 1$ and $d_B w = 0$ over $(0,1) \times S^3$. This means that the connection $B$ is reducible over $(0,1) \times S^3$. By the unique continuation theorem [7, p. 150], $B$ is reducible over $X$. This is a contradiction.

(ii) Fix $4 < p < \infty$. (Note that the Sobolev embedding $L^p_1 \hookrightarrow C^0$ is compact.) By an argument similar to the above (i), we can prove the following statement: For any $s \in \mathbb{R}$ and any $u \in \Omega^0(\mathrm{ad}E)$

$$\|u\|_{L^\infty(s<t<s+1)} \lesssim_A C(p)\|d_A u\|_{L^p(s<t<s+1)}. \tag{5}$$

We prove (ii) by using this statement. Suppose (ii) is false. Then, as in the proof of (i), there exist connections $A_n$ (which are translations of $A$) and gauge transformations $g_n : E \to E$ satisfying

$$\min_{\pm} \|g_n \pm 1\|_{L^\infty(0<t<1)} > n\|d_{A_n} g_n\|_{L^\infty(0<t<1)}.$$

We can choose a sequence of natural numbers $n_1 < n_2 < n_3 < \cdots$ and gauge transformations $h_k : E \to E$ $(k \geq 1)$ such that $B_k := h_k(A_{n_k})$ converges to some $B$ in $C^\infty$ over every compact subset. $B$ is irreducible. Set $g'_k := h_k g_{n_k} h_k^{-1}$. Then

$$\min_{\pm} \|g'_k \pm 1\|_{L^\infty(0<t<1)} > n_k\|d_{B_k} g'_k\|_{L^\infty(0<t<1)}. \tag{6}$$

$\{g'_k\}$ is bounded in $L^p_{1,B}((0,1)\times S^3)$. By choosing a subsequence, $g'_k$ converges to some $g'$ weakly in $L^p_{1,B}((0,1)\times S^3)$ and strongly in $L^\infty((0,1)\times S^3)$. We have $d_B g' = 0$. Since $B$ is irreducible, $g' = \pm 1$. We can assume $g' = 1$ without loss of generality. Then there are $u_k \in L^p_{1,B}((0,1)\times S^3, \Lambda^0(\mathrm{ad}E))$ $(k \gg 1)$ satisfying $g'_k = e^{u_k}$ and $|u_k| \lesssim |g'_k - 1|$ over $0 < t < 1$. Then by (5)

$$\|g'_k - 1\|_{L^\infty(0<t<1)} \lesssim \|u_k\|_{L^\infty(0<t<1)} \lesssim_A C(p)\|d_{B_k} u_k\|_{L^p(0<t<1)}.$$

We have

$$\|d_{B_k} u_k\|_{L^p(0<t<1)} \leq 2\|d_{B_k} g'_k\|_{L^p(0<t<1)} \quad \text{for } k \gg 1.$$

Hence, for $k \gg 1$,

$$\|g'_k - 1\|_{L^\infty(0<t<1)} \lesssim_A \|d_{B_k} g'_k\|_{L^\infty(0<t<1)}.$$

This contradicts (6). ∎

Lemma 3.3: *There exists a positive number $\varepsilon_1 = \varepsilon_1(A)$ such that, for any integers $m < n$ and any gauge transformation $g : E \to E$, if $\|d_A g\|_{L^\infty(m<t<n)} \leq \varepsilon_1$, then*

$$\min_{\pm} \|g \pm 1\|_{L^\infty(m<t<n)} \leq C_2(A)\|d_A g\|_{L^\infty(m<t<n)}.$$

*This is also true for the case $(m,n) = (-\infty, \infty)$.*

*Proof.* For simplicity we suppose $m = 0$. By Lemma 3.2 (ii), for every $k \in \mathbb{Z}$,

$$\min_{\pm} \|g \pm 1\|_{L^\infty(k<t<k+1)} \leq C_2 \|d_A g\|_{L^\infty(k<t<k+1)}. \tag{7}$$

Take a positive number $\varepsilon_1 = \varepsilon_1(A)$ satisfying $(C_2+1)\varepsilon_1 < 1$. Suppose

$$\|d_A g\|_{L^\infty(0<t<n)} \leq \varepsilon_1.$$

We can also suppose

$$\|g-1\|_{L^\infty(0<t<1)} \leq \|g+1\|_{L^\infty(0<t<1)}$$

without loss of generality. Then $\|g-1\|_{L^\infty(0<t<1)} \leq C_2\varepsilon_1$. Since $|d_A g| \leq \varepsilon_1$ over $0 \leq t \leq 2$, we have

$$\|g-1\|_{L^\infty(1<t<2)} \leq (C_2+1)\varepsilon_1 < 1.$$

Then $\|g+1\|_{L^\infty(1<t<2)} \geq 2-(C_2+1)\varepsilon_1 > 1$. Hence

$$\|g-1\|_{L^\infty(1<t<2)} < \|g+1\|_{L^\infty(1<t<2)}.$$

In the same way, we can prove that for every $0 \leq k < n$

$$\|g-1\|_{L^\infty(k<t<k+1)} < \|g+1\|_{L^\infty(k<t<k+1)}.$$

By (7),

$$\|g-1\|_{L^\infty(k<t<k+1)} \leq C_2 \|d_A g\|_{L^\infty(k<t<k+1)}.$$

Thus $\|g-1\|_{L^\infty(0<t<n)} \leq C_2 \|d_A g\|_{L^\infty(0<t<n)}$. ∎

Fix a positive integer $T = T(A)$ satisfying

$$\frac{10C_1 + 20\sqrt{C_1}}{T} < \frac{1}{4}. \tag{8}$$

Here $C_1 = C_1(A)$ is the positive constant introduced in Lemma 3.2 (i). For later convenience (Lemma 5.5) we assume $T > 3$. For $\xi \in \Omega^i(\mathrm{ad}E)$ and integers $m \leq n$, we set

$$\|\xi\|_m^n := \max_{m \leq k \leq n} \|\xi\|_{L^2(kT<t<(k+1)T)}.$$

Let $d_A^* : \Omega^1(\mathrm{ad}E) \to \Omega^0(\mathrm{ad}E)$ be the formal adjoint of $d_A : \Omega^0(\mathrm{ad}E) \to \Omega^1(\mathrm{ad}E)$. We set $\Delta_A u := d_A^* d_A u$ for $u \in \Omega^0(\mathrm{ad}E)$.

Lemma 3.4: *Let $n \in \mathbb{Z}$ and $K \in \mathbb{Z}_{>0}$, and let $u \in \Omega^0(\mathrm{ad}E)$. Then*

$$\int_{nT<t<(n+1)T} |d_A u|^2 d\mathrm{vol} \lesssim 2^{-K} (\|d_A u\|_{n-K}^{n+K})^2 + \|\Delta_A u\|_{n-K}^{n+K} \|u\|_{n-K}^{n+K}.$$

*Proof.* For simplicity, we suppose $n = 0$. Take any $m \in \mathbb{Z}$. Let $\varphi : \mathbb{R} \to [0,1]$ be a cut-off such that $\mathrm{supp}(\varphi) \subset [(m-1)T, (m+2)T]$, $\varphi = 1$ on $[mT, (m+1)T]$ and $|\varphi'|, |\varphi''| < 10/T$. Then

$$\int_{mT<t<(m+1)T} |d_A u|^2 \leq \int_X |d_A(\varphi u)|^2 = \int_X \langle \Delta_A(\varphi u), \varphi u \rangle.$$

We have $\Delta_A(\varphi u) = \varphi \Delta_A u + \Delta\varphi \cdot u + *(*d\varphi \wedge d_A u - d\varphi \wedge *d_A u)$,

$$|\Delta_A(\varphi u)| \leq (10/T)|u| + (20/T)|d_A u| + |\Delta_A u|.$$

Since $\Delta_A(\varphi u) = \Delta_A u$ over $mT \leq t \leq (m+1)T$,

$$\begin{aligned}\int_{mT<t<(m+1)T} &|d_A u|^2 \\ &\leq \int_{\{(m-1)T<t<mT \text{ or } (m+1)T<t<(m+2)T\}} (10/T)|u|^2 + (20/T)|d_A u||u| \\ &\quad + \int_{(m-1)T<t<(m+2)T} |\Delta_A u||u|.\end{aligned}$$

Using Lemma 3.2 (i), the right-hand side is bounded by

$$\begin{aligned}\frac{10C_1 + 20\sqrt{C_1}}{T} &\int_{\{(m-1)T<t<mT \text{ or } (m+1)T<t<(m+2)T\}} |d_A u|^2 \\ &+ \int_{(m-1)T<t<(m+2)T} |\Delta_A u||u|.\end{aligned}$$

From (8), this is bounded by

$$\frac{1}{4} \int_{\{(m-1)T<t<mT \text{ or } (m+1)T<t<(m+2)T\}} |d_A u|^2 d\mathrm{vol} + 3\|\Delta_A u\|_{m-1}^{m+1} \|u\|_{m-1}^{m+1}.$$

We define a sequence $a_m$ $(-K \leq m \leq K)$ by

$$a_m := \int_{mT<t<(m+1)T} |d_A u|^2 d\mathrm{vol}.$$

Then the above implies

$$a_m \leq \frac{a_{m-1} + a_{m+1}}{4} + 3\|\Delta_A u\|_{-K}^{K} \|u\|_{-K}^{K} \quad (-K+1 \leq m \leq K-1).$$

By applying Sublemma 3.5 below to this relation, we get

$$\begin{aligned}a_0 &\leq \frac{\max(a_K, a_{-K})}{2^{K-1}} + 18\|\Delta_A u\|_{-K}^{K} \|u\|_{-K}^{K} \\ &\leq \frac{1}{2^{K-1}} (\|d_A u\|_{-K}^{K})^2 + 18\|\Delta_A u\|_{-K}^{K} \|u\|_{-K}^{K}.\end{aligned}$$

SUBLEMMA 3.5: *Let $K$ be a positive integer, and let $b \geq 0$ be a real number. Let $\{a_m\}_{-K \leq m \leq K}$ be a sequence of non-negative real numbers satisfying*

$$a_m \leq \frac{a_{m-1} + a_{m+1}}{4} + b \quad (-K+1 \leq m \leq K-1).$$

*Then we have*

$$a_0 \leq \frac{\max(a_K, a_{-K})}{2^{K-1}} + 6b.$$

*Proof.* Set $b_m := \max(a_{-m}, a_m)$ $(0 \leq m \leq K)$. We have $b_0 \leq b_1/2 + b$. For $m \geq 1$, we have $b_m \leq (b_{m-1} + b_{m+1})/4 + b$, i.e., $4b_m \leq b_{m-1} + b_{m+1} + 4b$. Hence, for $m \geq 1$,

$$2(b_m - b_{m-1}) \leq 2b_m - b_{m-1} \leq -2b_m + b_{m+1} + 4b \leq b_{m+1} - b_m + 4b.$$

Thus $b_m - b_{m-1} \leq (b_{m+1} - b_m)/2 + 2b$ $(m \geq 1)$. Using this inequality recursively, we get

$$b_1 - b_0 \leq \frac{b_K - b_{K-1}}{2^{K-1}} + 2b\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{K-2}}\right) \leq \frac{b_K - b_{K-1}}{2^{K-1}} + 4b.$$

On the other hand, $2b_0 - b_1 \leq 2b$. Hence

$$a_0 = b_0 \leq \frac{b_K - b_{K-1}}{2^{K-1}} + 6b \leq \frac{b_K}{2^{K-1}} + 6b = \frac{\max(a_K, a_{-K})}{2^{K-1}} + 6b. \qquad \blacksquare$$

We have finished the proof of Lemma 3.4. $\blacksquare$

The next proposition is the main result of this section. Recall that we have supposed that $A$ is a non-degenerate ASD connection on $E$ with $\|F_A\|_{\mathrm{op}} \leq d$.

PROPOSITION 3.6: *For any $\tau > 0$, there exist*

$$\varepsilon_2 = \varepsilon_2(A, \tau) > 0 \quad \text{and} \quad K = K(A, \tau) \in \mathbb{Z}_{>0}$$

*satisfying the following statement.*

*Let $n \in \mathbb{Z}$. Let $a, b \in \Omega^1(\mathrm{ad}E)$ with $d_A^* a = d_A^* b = 0$, and let $g : E \to E$ be a gauge transformation. Set $\alpha := g(A + a) - (A + b)$. If the $L^\infty$-norms of $a$, $b$ and $\alpha$ over $(n-K)T < t < (n+K+1)T$ are all less than $\varepsilon_2$, then*

$$\|a - b\|_{L^2(nT<t<(n+1)T)} \leq \tau \|a - b\|_{n-K}^{n+K} + \sqrt{\|\alpha\|_{n-K}^{n+K} + \|d_A^* \alpha\|_{n-K}^{n+K}}, \tag{9}$$

$$\min_{\pm} \|g \pm 1\|_{L^2(nT<t<(n+1)T)} \lesssim_A \|\alpha\|_{L^2(nT<t<(n+1)T)} + \|a - b\|_{L^2(nT<t<(n+1)T)}. \tag{10}$$

*Proof.* For simplicity of the notations, we assume $n = 0$. Set

$$U := S^3 \times (-KT, KT + T).$$

We have $d_A g = -\alpha g + ga - bg$. Then $|d_A g| < 3\varepsilon_2$ over $U$. We choose $\varepsilon_2$ so that $3\varepsilon_2 \leq \varepsilon_1$ (the constant introduced in Lemma 3.3). Then by Lemma 3.3, we can suppose $\|g - 1\|_{L^\infty(U)} \lesssim_A \varepsilon_2 \ll 1$. So there is a section $u$ of $\Lambda^0(\mathrm{ad}E)$ over $U$ satisfying $g = e^u$ and $\|u\|_{L^\infty(U)} \lesssim_A \varepsilon_2 \ll 1$. Then $2^{-1}|d_A u| \leq |d_A g| \leq 2|d_A u|$ and $|g - 1| \leq 2|u|$ over $U$. By Lemma 3.2 (i),

$$\|g - 1\|_{-K}^K \leq 2\|u\|_{-K}^K \leq 2\sqrt{C_1}\|d_A u\|_{-K}^K \leq 4\sqrt{C_1}\|d_A g\|_{-K}^K. \tag{11}$$

We have

$$d_A g = -\alpha g + (g - 1)a - b(g - 1) + (a - b). \tag{12}$$

Then

$$\begin{aligned}\|d_A g\|_{-K}^K \leq & \|\alpha\|_{-K}^K + \|g - 1\|_{-K}^K(\|a\|_{L^\infty(U)} + \|b\|_{L^\infty(U)}) + \|a - b\|_{-K}^K \\ \leq & \|\alpha\|_{-K}^K + \|a - b\|_{-K}^K + 8\varepsilon_2\sqrt{C_1}\|d_A g\|_{-K}^K.\end{aligned}$$

We choose $\varepsilon_2 > 0$ so small that $8\varepsilon_2\sqrt{C_1} \leq 1/2$. Then

$$\|d_A g\|_{-K}^K \leq 2(\|\alpha\|_{-K}^K + \|a - b\|_{-K}^K). \tag{13}$$

This and (11) show that

$$\|g - 1\|_{-K}^K \leq 8\sqrt{C_1}(\|\alpha\|_{-K}^K + \|a - b\|_{-K}^K). \tag{14}$$

In the same way we get (10):

$$\|g - 1\|_{L^2(0<t<T)} \leq 8\sqrt{C_1}(\|\alpha\|_{L^2(0<t<T)} + \|a - b\|_{L^2(0<t<T)}).$$

From (12),

$$\begin{aligned}\|a - b\|_{L^2(0<t<T)} \leq & \|d_A g\|_{L^2(0<t<T)} + \|\alpha\|_{L^2(0<t<T)} \\ & + \|g - 1\|_{L^2(0<t<T)}(\|a\|_{L^\infty(U)} + \|b\|_{L^\infty(U)}) \\ \leq & \|d_A g\|_{L^2(0<t<T)} + \|\alpha\|_{L^2(0<t<T)} + 2\varepsilon_2\|g - 1\|_{L^2(0<t<T)} \\ \leq & \|d_A g\|_{L^2(0<t<T)} + (1 + 16\varepsilon_2\sqrt{C_1})\|\alpha\|_{L^2(0<t<T)} \\ & + 16\varepsilon_2\sqrt{C_1}\|a - b\|_{L^2(0<t<T)}.\end{aligned}$$

Since $|d_A g| \leq 2|d_A u|$ and $\varepsilon_2 \ll 1$,

$$\|a - b\|_{L^2(0<t<T)} \lesssim_A \|d_A u\|_{L^2(0<t<T)} + \|\alpha\|_{L^2(0<t<T)}. \tag{15}$$

We have the Coulomb gauge condition $d_A^* a = d_A^* b = 0$. Therefore

$$\begin{aligned}\Delta_A g &= - *d_A * d_A g = - * d_A(- * \alpha g + g * a - *bg)\\ &= - (d_A^*\alpha)g - *(*\alpha \wedge d_A g) - *(d_A g \wedge *a) - *(*b \wedge d_A g).\end{aligned}$$

By (13),

$$\begin{aligned}\|\Delta_A g\|_{-K}^K \le &\|d_A^*\alpha\|_{-K}^K + (\|\alpha\|_{L^\infty(U)} + \|a\|_{L^\infty(U)} + \|b\|_{L^\infty(U)})\|d_A g\|_{-K}^K\\ \le &\|d_A^*\alpha\|_{-K}^K + 6\varepsilon_2(\|\alpha\|_{-K}^K + \|a-b\|_{-K}^K);\end{aligned} \tag{16}$$

$\Delta_A g = \sum_{n=0}^\infty \Delta_A(u^n/n!)$ and $|\Delta_A(u^n)| \le n(n-1)|u|^{n-2}|d_A u|^2 + n|u|^{n-1}|\Delta_A u|$. Hence

$$|\Delta_A g - \Delta_A u| \le e^{|u|}|d_A u|^2 + (e^{|u|}-1)|\Delta_A u| \lesssim \varepsilon_2(|d_A g| + |\Delta_A u|)$$

over $U$. Here we have used $|u| \lesssim_A \varepsilon_2 \ll 1$ and $|d_A u| \le 2|d_A g| < 6\varepsilon_2$ over $U$. We choose $\varepsilon_2$ so small that $|\Delta_A u| \lesssim |\Delta_A g| + \varepsilon_2|d_A g|$ over $U$. By (13) and (16),

$$\|\Delta_A u\|_{-K}^K \lesssim \varepsilon_2\|\alpha\|_{-K}^K + \|d_A^*\alpha\|_{-K}^K + \varepsilon_2\|a-b\|_{-K}^K. \tag{17}$$

From (15), Lemma 3.4 and $\|\alpha\|_{L^\infty(U)} < \varepsilon_2$,

$$\begin{aligned}\|a-b\|_{L^2(0<t<T)}^2 &\lesssim_A (\|d_A u\|_{L^2(0<t<T)})^2 + (\|\alpha\|_{L^2(0<t<T)})^2\\ &\lesssim_A 2^{-K}(\|d_A u\|_{-K}^K)^2 + \|\Delta_A u\|_{-K}^K\|u\|_{-K}^K + \varepsilon_2\|\alpha\|_{-K}^K.\end{aligned}$$

From (13), $|d_A u| \le 2|d_A g|$ on $U$ and $\|\alpha\|_{L^\infty(U)} \le \varepsilon_2$,

$$(\|d_A u\|_{-K}^K)^2 \lesssim (\|\alpha\|_{-K}^K)^2 + (\|a-b\|_{-K}^K)^2 \lesssim_A \varepsilon_2\|\alpha\|_{-K}^K + (\|a-b\|_{-K}^K)^2.$$

From (14), $\|u\|_{-K}^K \lesssim_A \|\alpha\|_{-K}^K + \|a-b\|_{-K}^K$. From (17) and $\|a\|_{L^\infty(U)}$, $\|b\|_{L^\infty(U)}$, $\|\alpha\|_{L^\infty(U)} < \varepsilon_2$,

$$\begin{aligned}\|\Delta_A u\|_{-K}^K\|u\|_{-K}^K &\lesssim_A (\varepsilon_2\|\alpha\|_{-K}^K + \|d_A^*\alpha\|_{-K}^K + \varepsilon_2\|a-b\|_{-K}^K)(\|\alpha\|_{-K}^K + \|a-b\|_{-K}^K)\\ &\lesssim_A \varepsilon_2(\|\alpha\|_{-K}^K + \|d_A^*\alpha\|_{-K}^K) + \varepsilon_2(\|a-b\|_{-K}^K)^2.\end{aligned}$$

(The strange square root in (9) comes from the term $\|d_A^*\alpha\|_{-K}^K\|a-b\|_{-K}^K$ in this estimate.) Thus

$$\|a-b\|_{L^2(0<t<T)}^2 \lesssim_A (\varepsilon_2 + 2^{-K})(\|a-b\|_{-K}^K)^2 + \varepsilon_2(\|\alpha\|_{-K}^K + \|d_A^*\alpha\|_{-K}^K).$$

We choose $K > 0$ sufficiently large and $\varepsilon_2 > 0$ sufficiently small. Then we get

$$\|a-b\|_{L^2(0<t<T)}^2 \le \tau^2(\|a-b\|_{-K}^K)^2 + \|\alpha\|_{-K}^K + \|d_A^*\alpha\|_{-K}^K. \qquad \blacksquare$$

Corollary 3.7: *Suppose that $a, b \in \Omega^1(\mathrm{ad}E)$ satisfy $d_A^* a = d_A^* b = 0$ and $\|a\|_{L^\infty(X)}, \|b\|_{L^\infty(X)} \leq \varepsilon_2(A, 1/2)$ (the constant introduced in Proposition 3.6 for $\tau = 1/2$). If a gauge transformation $g : E \to E$ satisfies $g(A + a) = A + b$, then $a = b$ and $g = \pm 1$.*

*Proof.* For any $n \in \mathbb{Z}$, from Proposition 3.6 (9),

$$\|a - b\|_{L^2(nT<t<(n+1)T)} \leq \frac{1}{2}\|a - b\|_{n-K}^{n+K} \leq \frac{1}{2}\sup_{m\in\mathbb{Z}} \|a - b\|_{L^2(mT<t<(m+1)T)}.$$

Hence

$$\sup_{m\in\mathbb{Z}} \|a - b\|_{L^2(mT<t<(m+1)T)} \leq \frac{1}{2}\sup_{m\in\mathbb{Z}} \|a - b\|_{L^2(mT<t<(m+1)T)}.$$

This implies $a = b$. Then Proposition 3.6 (10) shows $g = \pm 1$. ∎

## 4. Parameter space of the deformation

For a connection $A$ on $E$, we set

$$D_A := d_A^* + d_A^+ : \Omega^1(\mathrm{ad}E) \to \Omega^0(\mathrm{ad}E) \oplus \Omega^+(\mathrm{ad}E).$$

Here $d_A^*$ is the formal adjoint of $d_A : \Omega^0(\mathrm{ad}E) \to \Omega^1(\mathrm{ad}E)$, and $d_A^+$ is the self-dual part of $d_A : \Omega^1(\mathrm{ad}E) \to \Omega^2(\mathrm{ad}E)$. We define a linear space $H_A^1$ by

$$H_A^1 := \{a \in \Omega^1(\mathrm{ad}E) | D_A a = 0,\ \|a\|_{L^\infty(X)} < \infty\}. \tag{18}$$

$(H_A^1, \|\cdot\|_{L^\infty(X)})$ is a (possibly infinite-dimensional) Banach space. This space will be the parameter space of the deformation theory developed in the next section. The main purpose of this section is to prove the following proposition:

Proposition 4.1: *Let $A$ be a non-degenerate ASD connection on $E$ satisfying $\|F_A\|_{\mathrm{op}} \leq d$. Then for any interval $(\alpha, \beta) \subset \mathbb{R}$ of length $> 2$ there exists a finite-dimensional linear subspace $V \subset H_A^1$ satisfying the following two conditions.*

(i)
$$\dim V \geq \frac{1}{\pi^2}\int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - C_3(A).$$
*Here $C_3(A)$ is a positive constant depending only on $A$. The important point is that it is independent of the interval $(\alpha, \beta)$.*

(ii) *All $a \in V$ satisfy $\|a\|_{L^\infty(X)} \leq 2\|a\|_{L^\infty(\alpha<t<\beta)}$.*

The following is a preliminary version of Proposition 4.1:

PROPOSITION 4.2: *Let $A$ be an ASD connection on $E$ satisfying $\|F_A\|_{\mathrm{op}} \leq d$. For any $\varepsilon > 0$ and any interval $(\alpha, \beta) \subset \mathbb{R}$ of length $> 2$, there exists a finite-dimensional linear subspace $W \subset \Omega^1(\mathrm{ad}E)$ such that*

(i)
$$\dim W \geq \frac{1}{\pi^2} \int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - C(\varepsilon).$$

(ii) *All $a \in W$ satisfy* $\mathrm{supp}(a) \subset (\alpha, \beta) \times S^3$.

(iii) *All $a \in W$ satisfy* $\mathrm{supp}(D_A a) \subset (\alpha, \alpha+1) \times S^3 \cup (\beta-1, \beta) \times S^3$ *and* $\|D_A a\|_{L^\infty(X)} \leq \varepsilon \|a\|_{L^\infty(X)}$.

*Proof.* From the compactness of $\mathcal{M}_d$, there is a bundle trivialization $g$ of $E$ over $U := \{\alpha - 1 < t < \alpha + 1\} \cup \{\beta - 1 < t < \beta + 1\} \subset X$ such that the connection matrix $g(A)$ satisfies

$$\|g(A)\|_{C^k(U)} \lesssim C(k) \quad (\forall k \geq 0).$$

Let $\psi : \mathbb{R} \to [0, 1]$ be a cut-off function such that $\psi = 1$ over a small neighborhood of $[\alpha + 1, \beta - 1]$, $\mathrm{supp}(\psi) \subset (\alpha + 1/2, \beta - 1/2)$ and $|d\psi| \leq 4$. Define a connection $A'$ over $(\alpha-1, \beta+1) \times S^3$ by $A' := \psi A$. (The precise definition is as follows: $A'$ is equal to $A$ on a small neighborhood of $[\alpha+1, \beta-1] \times S^3$, and it is equal to $g^{-1}(\psi g(A))$ over $U$.) We have $F(A') = \psi F(A) + d\psi \wedge A + (\psi^2 - \psi)A^2$;

$$|F(A')| \leq \sqrt{3}d + 4|A| + |A^2| \lesssim 1.$$

Set $X' := (\mathbb{R}/(\beta - \alpha)\mathbb{Z}) \times S^3$, and let $\pi : X \to X'$ be the natural projection. We define a principal $SU(2)$ bundle $E'$ on $X'$ as follows: We identify the region $\{\alpha < t < \beta\} \subset X$ with its projection $\pi\{\alpha < t < \beta\}$ and set

$$E' := E|_{\alpha<t<\beta} \cup (\pi(U) \times SU(2)),$$

where we glue the two terms of the right-hand side by using the trivialization $g$. We can naturally identify the connection $A'$ with a connection on $E'$ (also denoted by $A'$);

$$c_2(E') = \frac{1}{8\pi^2} \int_{X'} \mathrm{tr}(F_{A'}^2) \geq \frac{1}{8\pi^2} \int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - \mathrm{const}.$$

Let $H^1_{A'}$ be the linear space of $a \in \Omega^1_{X'}(\mathrm{ad}E')$ satisfying $D_{A'}a = d^*_{A'}a + d^+_{A'}a = 0$. From the Atiyah–Singer index theorem,

$$\dim H^1_{A'} \geq 8c_2(E') \geq \frac{1}{\pi^2} \int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - \mathrm{const}.$$

LEMMA 4.3: *All* $a \in H^1_{A'}$ *satisfy*

$$\|\nabla_{A'} a\|_{L^\infty(X')} \lesssim \|a\|_{L^\infty(X')}.$$

*Proof.* Take any $\gamma \in \mathbb{R}$. From the construction, we can choose a connection matrix of $A'$ over $\pi\{\gamma < t < \gamma + 1\}$ so that

$$\|A'\|_{C^k(\pi\{\gamma<t<\gamma+1\})} \lesssim C(k) \quad (\forall k \geq 0).$$

Then the standard elliptic regularity theory (Gilbarg–Trudinger [10, Theorem 9.11]) shows that

$$\|\nabla_{A'} a\|_{L^\infty(\pi\{\gamma+1/4<t<\gamma+3/4\})} \lesssim \|a\|_{L^\infty(\pi\{\gamma<t<\gamma+1\})}.$$

A similar argument will be also used in the proof of Lemma 6.2. ∎

Set $\Omega := \pi(U) \subset X'$. Let $\tau = \tau(\varepsilon) > 0$ be a small number which will be fixed later. Take points $x_1, x_2, \ldots, x_N$ ($N \lesssim 1/\tau^4$) in $\Omega$ such that for any $x \in \Omega$ there is some $x_i$ satisfying $d(x, x_i) \leq \tau$. Let $V$ be the kernel of the following linear map:

$$H^1_{A'} \to \bigoplus_{i=1}^{N} (\Lambda^1(\mathrm{ad}E'))_{x_i}, \quad a \mapsto (a(x_i))_{i=1}^N.$$

We have

$$\dim V \geq \dim H^1_{A'} - 12N \geq \frac{1}{\pi^2} \int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - \mathrm{const} - 12N.$$

Take any $a \in V$ and $x \in \Omega$. Choose $x_i$ satisfying $d(x, x_i) \leq \tau$. From Lemma 4.3 and $a(x_i) = 0$,

$$|a(x)| \leq \tau \|\nabla_{A'} a\|_{L^\infty(X')} \lesssim \tau \|a\|_{L^\infty(X')}.$$

We can choose $\tau > 0$ so that the maximum of $|a|$ is attained at a point in $X' \setminus \Omega$. For $a \in V$, we define $\tilde{a} \in \Omega^1(\mathrm{ad}E)$ over $X$ by $\tilde{a} := \psi a$. (The precise definition is as follows: We identify the region $\{\alpha < t < \beta\}$ with its projection in $X'$; $\tilde{a}$ is equal to $\psi a$ over $\alpha < t < \beta$, and it is equal to 0 outside of $\mathrm{supp}(\psi)$.) Set $W := \{\tilde{a} | \, a \in V\} \subset \Omega^1(\mathrm{ad}E)$; $W$ satisfies the condition (ii) in the statement. We have $\|\tilde{a}\|_{L^\infty(X)} = \|a\|_{L^\infty(X')}$ because the maximum of $|a|$ is attained at a point in $X' \setminus \Omega$. Hence

$$\dim W = \dim V \geq \frac{1}{\pi^2} \int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - \mathrm{const} - 12N.$$

We have

$$D_A\tilde{a} = (A - A') * \tilde{a} + D_{A'}(\psi a) = (A - A') * \tilde{a} + (d\psi) * a.$$

Here $*$ are algebraic multiplications; $D_A\tilde{a}$ is supported in

$$\{\alpha < t < \alpha + 1\} \cup \{\beta - 1 < t < \beta\}$$

and

$$\|D_A\tilde{a}\|_{L^\infty(X)} \lesssim \|a\|_{L^\infty((\alpha,\alpha+1)\times S^3 \cup (\beta-1,\beta)\times S^3)} \lesssim \tau\|\tilde{a}\|_{L^\infty(X)}.$$

We can choose $\tau = \tau(\varepsilon) > 0$ so that $\|D_A\tilde{a}\|_{L^\infty(X)} \leq \varepsilon\|\tilde{a}\|_{L^\infty(X)}$. Then $W$ satisfies the conditions (i), (ii), (iii) in the statement. ∎

LEMMA 4.4: *Let $\alpha < \beta$. Let $A$ be an ASD connection on $E$ satisfying $\|F_A\|_{\mathrm{op}} \leq d$.*

(i) *If $A$ is non-degenerate, then there is a linear map*

$$\{u \in \Omega^0(\mathrm{ad}E) | \mathrm{supp}(u) \subset (\alpha, \alpha+1) \times S^3 \cup (\beta-1, \beta) \times S^3\} \to \Omega^0(\mathrm{ad}E), \quad u \mapsto v,$$

*satisfying*

$$d_A^* d_A v = u, \quad \|v\|_{L^\infty(X)} + \|d_A v\|_{L^\infty(X)} \lesssim_A \|u\|_{L^\infty(X)}.$$

(ii) *There is a linear map*

$$\{\xi \in \Omega^+(\mathrm{ad}E) | \mathrm{supp}(\xi) \subset (\alpha, \alpha+1) \times S^3 \cup (\beta-1, \beta) \times S^3\} \to \Omega^+(\mathrm{ad}E), \quad \xi \mapsto \eta,$$

*satisfying*

$$d_A^+ d_A^* \eta = \xi, \quad \|\eta\|_{L^\infty(X)} + \|\nabla_A \eta\|_{L^\infty(X)} \lesssim \|\xi\|_{L^\infty(X)}.$$

*The statement* (ii) *does not require the non-degeneracy of $A$.*

*Proof.* (i) Set $L^2_{1,A}(\mathrm{ad}E) := \{w \in L^2(\mathrm{ad}E) | \, d_A w \in L^2(X)\}$ with the inner product $(w_1, w_2)' := (d_A w_1, d_A w_2)_{L^2(X)}$. From Lemma 3.2 (i), every compactly supported $w \in \Omega^0(\mathrm{ad}E)$ satisfies $\|w\|_{L^2(X)} \leq \sqrt{C_1}\|d_A w\|_{L^2(X)} = \sqrt{C_1}\|w\|'$. Hence the norm $\|\cdot\|'$ is equivalent to $\|\cdot\|_{L^2_{1,A}(X)}$. In particular, $(L^2_{1,A}(\mathrm{ad}E), (\cdot,\cdot)')$ becomes a Hilbert space.

The rest of the argument is the standard $L^2$-method: Take $u \in \Omega^0(\mathrm{ad}E)$ with $\mathrm{supp}(u) \subset (\alpha, \alpha+1) \times S^3 \cup (\beta-1, \beta) \times S^3$. We apply the Riesz representation theorem to the following bounded linear functional:

$$(\cdot, u)_{L^2(X)} : L^2_{1,A}(\mathrm{ad}E) \to \mathbb{R}, \quad w \mapsto (w, u)_{L^2(X)}.$$

(From Lemma 3.2 (i), $|(w,u)_{L^2(X)}| \leq \sqrt{C_1}\|w\|'\|u\|_{L^2(X)}$.) Then there uniquely exists $v \in L^2_{1,A}(\mathrm{ad}E)$ satisfying $(d_A w, d_A v) = (w,v)' = (w,u)_{L^2(X)}$. This means that $d_A^* d_A v = u$ as a distribution. Moreover,

$$\|d_A v\|_{L^2(X)} = \|v\|' \leq \sqrt{C_1}\|u\|_{L^2(X)} \lesssim_A \|u\|_{L^\infty(X)}.$$

From Lemma 3.2 (i), $\|v\|_{L^2(X)} \lesssim_A \|u\|_{L^\infty(X)}$. As in the proof of Lemma 4.3, the elliptic regularity theory gives

$$\|v\|_{L^\infty(X)} + \|d_A v\|_{L^\infty(X)} \lesssim \|v\|_{L^2(X)} + \|d_A^* d_A v\|_{L^\infty(X)} \lesssim_A \|u\|_{L^\infty(X)}.$$

(ii) We have the Weitzenböck formula [9, Chapter 6]:

$$d_A^+ d_A^* \eta = \frac{1}{2}\left(\nabla_A^* \nabla_A + S/3\right)\eta \quad \text{for } \eta \in \Omega^+(\mathrm{ad}E).$$

Here $S$ is the scalar curvature of $X$, and it is a positive constant. Then the $L^2$-method shows the above statement. (Indeed a stronger result will be given in Lemma 6.1 in Section 6.2.) ∎

*Proof of Proposition 4.1.* Let $\varepsilon = \varepsilon(A) > 0$ be a small number which will be fixed later. For this $\varepsilon$ and the interval $(\alpha, \beta) \subset \mathbb{R}$ there is a finite-dimensional subspace $W \subset \Omega^1(\mathrm{ad}E)$ satisfying the conditions (i), (ii), (iii) in Proposition 4.2.

From Lemma 4.4, there is a linear map $W \to \Omega^0(\mathrm{ad}E) \oplus \Omega^+(\mathrm{ad}E)$, $a \mapsto (v, \eta)$, satisfying $d_A^* d_A v = d_A^* a$, $d_A^+ d_A^* \eta = d_A^+ a$ and

$$\begin{aligned}\|d_A v\|_{L^\infty(X)} + \|d_A^* \eta\|_{L^\infty(X)} \leq & C\|D_A a\|_{L^\infty(X)} \\ \leq & \varepsilon C \|a\|_{L^\infty(X)} (= \varepsilon C \|a\|_{L^\infty(\alpha<t<\beta)}),\end{aligned}$$

where $C = C(A)$ is a positive constant depending only on $A$. Here we have used the conditions (ii) and (iii) in Proposition 4.2. Set $a' := a - d_A v - d_A^* \eta$. This satisfies $D_A a' = 0$. Set $V := \{a' \mid a \in W\} \subset H^1_A$. We have

$$\|a'\|_{L^\infty(X)} \geq (1 - \varepsilon C)\|a\|_{L^\infty(X)} \quad \text{for } a \in W.$$

We choose $\varepsilon > 0$ sufficiently small so that $(1 - \varepsilon C) > 0$. Then $\dim V = \dim W$. From the condition (i) of Proposition 4.2,

$$\dim V \geq \frac{1}{\pi^2}\int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - \mathrm{const}_\varepsilon.$$

We have $\|a'\|_{L^\infty(X)} \leq (1+\varepsilon C)\|a\|_{L^\infty(X)}$ for $a \in W$. On the other hand, from the conditions (ii) and (iii) of Proposition 4.2,

$$\|a'\|_{L^\infty(\alpha<t<\beta)} \geq \|a\|_{L^\infty(\alpha<t<\beta)} - \varepsilon C\|a\|_{L^\infty(X)} = (1-\varepsilon C)\|a\|_{L^\infty(X)}.$$

Hence

$$\|a'\|_{L^\infty(X)} \leq \frac{1+\varepsilon C}{1-\varepsilon C}\|a'\|_{L^\infty(\alpha<t<\beta)}.$$

We choose $\varepsilon > 0$ so that $(1+\varepsilon C)/(1-\varepsilon C) \leq 2$. Then $\|a'\|_{L^\infty(X)} \leq 2\|a'\|_{L^\infty(\alpha<t<\beta)}$ for all $a' \in V$. ∎

## 5. Deformation theory and the proof of Theorem 1.10

In this section we develop a deformation theory of non-degenerate ASD connections and prove Theorem 1.10. (The paper [18] studied a deformation theory of periodic ASD connections.) Let $A$ be a non-degenerate ASD connection on $E$ satisfying $\|F_A\|_{\mathrm{op}} < d$. Note that this is a strict inequality. We fix this $A$ throughout this section.

5.1. Deformation theory. Let $H^1_A \subset \Omega^1(\mathrm{ad}E)$ be the Banach space defined by (18). Let $k \geq 0$ and $0 \leq i \leq 4$ be integers. For $\xi \in L^2_{k,loc}(\Lambda^i(\mathrm{ad}E))$ (a locally $L^2_k$-section of $\Lambda^i(\mathrm{ad}E)$), we set

$$\|\xi\|_{\ell^\infty L^2_k} := \sum_{j=0}^{k} \sup_{n\in\mathbb{Z}} \|\nabla_A^j \xi\|_{L^2(n<t<n+1)}.$$

From the elliptic regularity, we have $\|a\|_{L^\infty(X)} \lesssim \|a\|_{\ell^\infty L^2_k} \lesssim \mathrm{const}_k \|a\|_{L^\infty(X)}$ for $a \in H^1_A$ (cf. the proof of Lemma 4.3).

Let $\ell^\infty L^2_k(\Lambda^+(\mathrm{ad}E))$ be the Banach space of $\xi \in L^2_{k,loc}(\Lambda^+(\mathrm{ad}E))$ satisfying $\|\xi\|_{\ell^\infty L^2_k} < \infty$. From the Sobolev embedding theorem, $\|\xi\|_{L^\infty(X)} \lesssim \|\xi\|_{\ell^\infty L^2_3}$ for $\xi \in \ell^\infty L^2_3(\Lambda^+(\mathrm{ad}E))$. Consider

$$\begin{aligned}\Phi : H^1_A \times \ell^\infty L^2_5(\Lambda^+(\mathrm{ad}E)) \to & \ell^\infty L^2_3(\Lambda^+(\mathrm{ad}E)),\\ (a,\phi) \mapsto & F^+(A + a + d_A^*\phi)\\ = & (a\wedge a)^+ + d_A^+ d_A^* \phi + [a \wedge d_A^*\phi]^+ + (d_A^*\phi \wedge d_A^*\phi)^+.\end{aligned}$$

This is a smooth map between the Banach spaces with $\Phi(0,0) = 0$. We want to describe the fiber $\Phi^{-1}(0)$ around the origin by using the implicit function theorem. Let $(\partial_2\Phi)_0 : \ell^\infty L^2_5(\Lambda^+(\mathrm{ad}E)) \to \ell^\infty L^2_3(\Lambda^+(\mathrm{ad}E))$ be the derivative of $\Phi$ at the origin with respect to the second variable $\phi$. We have

$(\partial_2\Phi)_0(\phi) = d_A^+ d_A^* \phi = \frac{1}{2}(\nabla_A^*\nabla_A + S/3)\phi$ for $\phi \in \ell^\infty L_5^2(\Lambda^+(\mathrm{ad}E))$ by the Weitzenböck formula. (Here $S$ is the scalar curvature of $X$, and it is a positive constant.) The following $L^\infty$-estimate is proved in [18, Proposition A.5]:

Lemma 5.1: *Let $\xi$ be a $C^2$-section of $\Lambda^+(\mathrm{ad}E)$ over $X$. We set*

$$\eta := (\nabla_A^*\nabla_A + S/3)\xi,$$

*and suppose $\|\xi\|_{L^\infty(X)} < \infty$ and $\|\eta\|_{L^\infty(X)} < \infty$. Then*

$$\|\xi\|_{L^\infty(X)} \leq (24/S)\|\eta\|_{L^\infty(X)}.$$

Lemma 5.2: *The operator $(\partial_2\Phi)_0 : \ell^\infty L_5^2(\Lambda^+(\mathrm{ad}E)) \to \ell^\infty L_3^2(\Lambda^+(\mathrm{ad}E))$ is an isomorphism. This means that a local deformation of $A$ is "unobstructed".*

*Proof.* This can be proved by using Lemma 6.1 in Section 6.2. But here we give a direct proof. From the $L^\infty$-estimate in Lemma 5.1, the above operator is injective. Hence the problem is its surjectivity. Take $\eta \in \ell^\infty L_3^2(\Lambda^+(\mathrm{ad}E))$. Let $\varphi_n : \mathbb{R} \to [0,1]$ be a cut-off function such that $\varphi_n = 1$ over $[-n,n]$ and $\mathrm{supp}(\varphi_n) \subset (-n-1, n+1)$. Set $\eta_n := \varphi_n\eta$. By the $L^2$-method (see the proof of Lemma 4.4), there exists $\xi_n \in L_{1,A}^2(\Lambda^+(\mathrm{ad}E))$ satisfying $(\nabla_A^*\nabla_A + S/3)\xi_n = \eta_n$ as a distribution and $\|\xi_n\|_{L^2(X)} \lesssim \|\eta_n\|_{L^2(X)} < \infty$. From the elliptic regularity, $\xi_n$ is in $L_{5,loc}^2$ and hence of class $C^2$. Moreover,

$$\|\xi_n\|_{L^\infty(X)} \lesssim \|\xi_n\|_{L^2(X)} + \|\eta_n\|_{L^\infty} < \infty.$$

Hence by the $L^\infty$-estimate (Lemma 5.1)

$$\|\xi_n\|_{L^\infty(X)} \leq (24/S)\|\eta_n\|_{L^\infty(X)} \leq (24/S)\|\eta\|_{L^\infty(X)} \lesssim \|\eta\|_{\ell^\infty L_3^2}.$$

For any integer $m$,

$$\|\xi_n\|_{L_5^2(m<t<m+1)} \lesssim \|\xi_n\|_{L^\infty(X)} + \|\eta_n\|_{\ell^\infty L_3^2} \lesssim \|\eta\|_{\ell^\infty L_3^2}.$$

By choosing a subsequence $\{\xi_{n_k}\}_{k\geq 1}$, there exists $\xi \in L_{5,loc}^2(\Lambda^+(\mathrm{ad}E))$ such that $\xi_{n_k}$ converges to $\xi$ weakly in $L_5^2((m,m+1)\times S^3)$ for every $m \in \mathbb{Z}$. Then $(\nabla_A^*\nabla_A + S/3)\xi = \eta$ and $\|\xi\|_{\ell^\infty L_5^2} \lesssim \|\eta\|_{\ell^\infty L_3^2} < \infty$. ∎

By the implicit function theorem, we can choose $R > 0$ and $R' > 0$ such that for any $a \in H_A^1$ with $\|a\|_{L^\infty(X)} \leq R$ there uniquely exists $\phi_a \in \ell^\infty L_5^2(\Lambda^+(\mathrm{ad}E))$ satisfying $F^+(A + a + d_A^*\phi_a) = 0$ and $\|\phi_a\|_{\ell^\infty L_5^2} \leq R'$. We have $\phi_0 = 0$. For $a \in B_R(H_A^1) := \{a \in H_A^1 | \|a\|_{L^\infty} \leq R\}$ we set

$$a' := a + d_A^*\phi_a.$$

This satisfies the ASD equation $F^+(A+a')=0$ and the Coulomb gauge condition $d_A^* a'=0$. Since $\|F_A\|_{\mathrm{op}}<d$, we can choose $R>0$ sufficiently small so that

$$\|F(A+a')\|_{\mathrm{op}} \le d \quad \text{for all } a \in B_R(H_A^1).$$

Thus we get a deformation map:

$$B_R(H_A^1) \to \mathcal{M}_d, \quad a \mapsto [A+a']. \tag{19}$$

The derivative $(\partial_1\Phi)_0 : H_A^1 \to \ell^\infty L_3^2(\Lambda^+(\mathrm{ad}E))$ of $\Phi$ at the origin with respect to the first variable is equal to zero. Hence the derivative of the following map at the origin is also zero:

$$B_R(H_A^1) \to \ell^\infty L_5^2(\Lambda^+(\mathrm{ad}E)), \quad a \mapsto \phi_a.$$

Then we get

$$\|\phi_a-\phi_b\|_{\ell^\infty L_5^2} \lesssim_A (\|a\|_{L^\infty(X)}+\|b\|_{L^\infty(X)})\|a-b\|_{L^\infty(X)} \tag{20}$$

for $a,b \in B_R(H_A^1)$. In particular, the map $(B_R(H_A^1), \|\cdot\|_{L^\infty(X)}) \to \mathcal{M}_d$ is continuous.

*Remark 5.3:* Note that the construction of the deformation map (19) does not use the non-degeneracy condition of $A$. It will be used for the further study of the deformation map. Indeed, since $A$ is non-degenerate, we can apply Corollary 3.7 to this situation. Then we can show that the above map (19) is injective if $R$ is sufficiently small. Moreover, if $B_R(H_A^1)$ is endowed with the topology of uniform convergence over compact subsets (this is not equal to the norm topology), then $B_R(H_A^1)$ is compact and the map (19) becomes a topological embedding. We don't need these facts for the proof of Theorem 1.10. So we omit the details. But it is not difficult.

*Remark 5.4:* In the above argument we have solved the equation

$$F^+(A+a+d_A^*\phi)=0$$

by using the implicit function theorem. But indeed we can solve it more directly by using the method of Section 6.2. So there exists a little redundancy in our way of the explanation. We can prepare a unified method for both Sections 5.1 and 6.2. But we don't take this way here because this redundancy is not so heavy and the above implicit function theorem argument seems conceptually easier (at least for the authors) to understand.

5.2. Proof of Theorem 1.10. We need a distance on $\mathcal{M}_d$. Any choice will do. One choice is: For $[A_1], [A_2] \in \mathcal{M}_d$, we define the distance $\mathrm{dist}([A_1], [A_2])$ as the infimum of

$$\sum_{n=1}^{\infty} 2^{-n} \frac{\|g(A_1) - A_2\|_{L^\infty(-n<t<n)}}{1 + \|g(A_1) - A_2\|_{L^\infty(-n<t<n)}}$$

over all gauge transformations $g : E \to E$. We don't need this explicit formula. But probably it will be helpful for understanding.

Recall the following notation: For $\Omega \subset \mathbb{R}$ we define $\mathrm{dist}_\Omega([A_1], [A_2])$ as the supremum of $\mathrm{dist}([s^*(A_1)], [s^*(A_2)])$ over $s \in \Omega$; $s^*(\cdot)$ is the pull-back by $s : E \to E$. In particular, for $s \in \mathbb{R}$, the distance $\mathrm{dist}_{\{s\}}([A_1], [A_2])$ is the infimum of

$$\sum_{n=1}^{\infty} 2^{-n} \frac{\|g(A_1) - A_2\|_{L^\infty(s-n<t<s+n)}}{1 + \|g(A_1) - A_2\|_{L^\infty(s-n<t<s+n)}}$$

over all gauge transformations $g : E \to E$. We will abbreviate $\mathrm{dist}_{\{s\}}([A_1], [A_2])$ to $\mathrm{dist}_s([A_1], [A_2])$.

For the proof of Theorem 1.10, we need to compare the distances $\mathrm{dist}_{(\alpha,\beta)}$ on $\mathcal{M}_d$ and $\|\cdot\|_{L^\infty(\alpha<t<\beta)}$ on $B_R(H^1_A)$ for intervals $(\alpha, \beta) \subset \mathbb{R}$. The next lemma gives us a solution. This is a consequence of Proposition 3.6.

Lemma 5.5: *We can choose $0 < R_1 < R$ so that the following statement holds. For any $\delta > 0$ there exists $\varepsilon > 0$ such that if $a, b \in H^1_A$ with $\|a\|_{L^\infty(X)}, \|b\|_{L^\infty(X)} < R_1$ satisfy*

$$\mathrm{dist}_s([A + a'], [A + b']) \le \varepsilon$$

*for some $s \in \mathbb{R}$, then*

$$\|a - b\|_{L^\infty(s<t<s+1)} \le \frac{1}{4}\|a - b\|_{L^\infty(X)} + \delta.$$

*Proof.* Let $T = T(A) > 3$ be the positive constant introduced in Section 3. See the discussion around (8). We choose $n \in \mathbb{Z}$ so that

$$nT \le s - 1 < s + 2 \le (n + 2)T.$$

Then from the elliptic regularity

$$\|a - b\|_{L^\infty(s<t<s+1)} \lesssim \|a - b\|_{L^2(s-1<t<s+2)} \le \|a - b\|_{L^2(nT<t<(n+2)T)}.$$

Let $0 < \tau < 1$ be a small number which will be fixed later. Let $\varepsilon_2 = \varepsilon_2(A, \tau) > 0$ and $K = K(A, \tau) \in \mathbb{Z}_{>0}$ be the positive constants introduced in Proposition

3.6. From (20), if $R_1 \ll 1$,

$$\|a-b\|_{L^2(nT<t<(n+2)T)} \le \|a'-b'\|_{L^2(nT<t<(n+2)T)} + \tau\|a-b\|_{L^\infty(X)}.$$

Hence

$$\|a-b\|_{L^\infty(s<t<s+1)} \lesssim \|a'-b'\|_{L^2(nT<t<(n+2)T)} + \tau\|a-b\|_{L^\infty(X)}. \tag{21}$$

We estimate the term $\|a'-b'\|_{L^2(nT<t<(n+2)T)}$ by using Proposition 3.6.

We can assume $\delta^2 < \varepsilon_2$. From the Uhlenbeck compactness we can choose $\varepsilon > 0$ so that if two connections $[A_1],[A_2] \in \mathcal{M}_d$ satisfy $\mathrm{dist}([A_1],[A_2]) \le \varepsilon$, then there exists a gauge transformation $g : E \to E$ satisfying

$$\|g(A_1)-A_2\|_{L^\infty(-KT-2T<t<KT+2T)} + \|d^*_{A_2}(g(A_1)-A_2)\|_{L^\infty(-KT-2T<t<KT+2T)} < \tau^2\delta^2.$$

Then the assumption $\mathrm{dist}_s([A+a'],[A+b']) \le \varepsilon$ implies that there exists a gauge transformation $g : E \to E$ satisfying (set $\alpha := g(A+a') - (A+b')$)

$$\|\alpha\|_{L^\infty((n-K)T<t<(n+K+2)T)} + \|d^*_{A+b'}\alpha\|_{L^\infty((n-K)T<t<(n+K+2)T)} < \tau^2\delta^2 < \varepsilon_2.$$

In particular, $\|\alpha\|_{L^\infty((n-K)T<t<(n+K+2)T)} < \varepsilon_2$. Hence if $R_1 \ll \varepsilon_2$, then we can apply Proposition 3.6 to the present situation:

$$\begin{aligned}\|a'-b'\|_{L^2(nT<t<(n+2)T)} &\lesssim_A \tau\|a'-b'\|_{\ell^\infty L^2} + \tau\delta \\ &\lesssim_A \tau\|a-b\|_{L^\infty(X)} + \tau\delta \quad \text{(by (20))}.\end{aligned}$$

By applying this estimate to the above (21), we get

$$\|a-b\|_{L^\infty(s<t<s+1)} \lesssim_A \tau\|a-b\|_{L^\infty(X)} + \tau\delta.$$

We choose $\tau > 0$ sufficiently small. Then

$$\|a-b\|_{L^\infty(s<t<s+1)} \le \frac{1}{4}\|a-b\|_{L^\infty(X)} + \delta. \qquad \blacksquare$$

Recall that $B_r([A])_{\mathbb{R}} \subset \mathcal{M}_d$ is the closed ball of radius $r$ centered at $[A]$ with respect to the distance $\mathrm{dist}_{\mathbb{R}}$.

Proposition 5.6: *For any $r > 0$ there exists $\varepsilon(r) > 0$ such that for any $0 < \varepsilon \le \varepsilon(r)$ and any interval $(\alpha,\beta) \subset \mathbb{R}$ of length $> 2$ we have*

$$\mathrm{Widim}_\varepsilon(B_r([A])_{\mathbb{R}}, \mathrm{dist}_{(\alpha,\beta)}) \ge \frac{1}{\pi^2}\int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - C_3.$$

*Here $C_3 = C_3(A)$ is the positive constant introduced in Proposition 4.1.*

*Proof.* We can choose $0 < r' < R_1$ such that every $a \in B_{r'}(H^1_A)$ satisfies $[A + a'] \in B_r([A])_{\mathbb{R}}$. ($R_1$ is the constant introduced in the previous lemma.) From Lemma 5.5 we can choose $\varepsilon(r) > 0$ so that if $a, b \in B_{r'}(H^1_A)$ satisfy

$$\mathrm{dist}_{(\alpha,\beta)}([A + a'], [A + b']) \leq \varepsilon(r),$$

then

$$\|a - b\|_{L^\infty(\alpha<t<\beta)} \leq \frac{1}{4}\|a - b\|_{L^\infty(X)} + \frac{r'}{8}. \tag{22}$$

By Proposition 4.1, there exists a linear subspace $V \subset H^1_A$ such that

$$\dim V \geq \frac{1}{\pi^2}\int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - C_3$$

and that all $a \in V$ satisfy $\|a\|_{L^\infty(X)} \leq 2\|a\|_{L^\infty(\alpha<t<\beta)}$. We investigate the restriction of the deformation map (19) to $B_{r'}(V) := \{a \in V | \|a\|_{L^\infty(X)} \leq r'\}$.

By applying the above (22) to $B_{r'}(V)$, we get the following: If $a, b \in B_{r'}(V)$ satisfy $\mathrm{dist}_{(\alpha,\beta)}([A + a'], [A + b']) \leq \varepsilon(r)$, then

$$\|a - b\|_{L^\infty(X)} \leq 2\|a - b\|_{L^\infty(\alpha<t<\beta)} \leq \frac{1}{2}\|a - b\|_{L^\infty(X)} + \frac{r'}{4}$$

and hence $\|a - b\|_{L^\infty(X)} \leq r'/2$. Therefore we get: For $0 < \varepsilon \leq \varepsilon(r)$,

$$\begin{aligned}
&\mathrm{Widim}_\varepsilon(B_r([A])_{\mathbb{R}}, \mathrm{dist}_{(\alpha,\beta)}) \\
&\qquad \geq \mathrm{Widim}_{r'/2}(B_{r'}(V), \|\cdot\|_{L^\infty(X)}) = \dim V \quad \text{(by Example 2.1)} \\
&\qquad \geq \frac{1}{\pi^2}\int_{\alpha<t<\beta} |F_A|^2 d\mathrm{vol} - C_3. \qquad \blacksquare
\end{aligned}$$

*Proof of Theorem 1.10.* The upper bound $\dim_{[A]}(\mathcal{M}_d : \mathbb{R}) \leq 8\rho(A)$ is given by Theorem 1.9. So the problem is the lower bound.

Now $\dim_{[A]}(\mathcal{M}_d : \mathbb{R}) = \lim_{r\to 0} \dim(B_r([A])_{\mathbb{R}} : \mathbb{R})$, and $\dim(B_r([A])_{\mathbb{R}} : \mathbb{R})$ is given by

$$\lim_{\varepsilon\to 0}\Big(\lim_{n\to+\infty} \frac{\sup_{x\in\mathbb{R}} \mathrm{Widim}_\varepsilon(B_r([A])_{\mathbb{R}}, \mathrm{dist}_{(x,x+n)})}{n}\Big).$$

By Proposition 5.6, for $0 < \varepsilon \leq \varepsilon(r)$ and $n > 2$,

$$\sup_{x\in\mathbb{R}} \mathrm{Widim}_\varepsilon(B_r([A])_{\mathbb{R}}, \mathrm{dist}_{(x,x+n)}) \geq \frac{1}{\pi^2}\sup_{x\in\mathbb{R}}\int_{x<t<x+n} |F_A|^2 d\mathrm{vol} - C_3.$$

Since

$$\rho(A) = \lim_{n\to\infty} \frac{1}{8\pi^2 n}\sup_{x\in\mathbb{R}}\int_{x<t<x+n} |F_A|^2 d\mathrm{vol},$$

we have

$$\dim(B_r([A])_{\mathbb{R}} : \mathbb{R}) \geq 8\rho(A).$$

Thus $\dim_{[A]}(\mathcal{M}_d : \mathbb{R}) \geq 8\rho(A)$. $\blacksquare$

## 6. Gluing infinitely many instantons

In this section we prove Theorem 1.11: Suppose $d > 1$. Let $\varepsilon > 0$, and let $A$ be an ASD connection on $E$ with $\|F_A\|_{\mathrm{op}} < d$. We want to find a non-degenerate ASD connection $\tilde{A}$ on $E$ satisfying

$$\|F(\tilde{A})\|_{\mathrm{op}} < d, \quad \rho(\tilde{A}) > \rho(A) - \varepsilon. \tag{23}$$

If $A$ itself is non-degenerate, then $\tilde{A} := A$ satisfies the condition. So we assume that $A$ is degenerate.

As we described in Section 1.3, the idea of the proof is gluing instantons. We glue infinitely many copies of the instanton $I$ (given in Example 1.1) to $A$ over the regions where the curvature $F_A$ has very small norm. Then we get a non-degenerate ASD connection $\tilde{A}$. The technique of gluing infinitely many instantons in the context of Yang–Mills theory was first developed in [21]. It was further expanded in [23]. Infinite gluing techniques (in other words, shadowing lemmas) for other equations can be found in Angenent [1], Eremenko [8], Macrì–Nolasco–Ricciardi [17] and Gournay [11, 12].

Throughout this section, we fix a positive number $\tau$ such that

$$\|F_A\|_{\mathrm{op}} < d - \tau, \quad d - \tau > 1.$$

Let $\delta = \delta(\varepsilon, \tau) > 0$ be a sufficiently small number, and $T = T(\varepsilon, \tau, \delta) > 0$ be a sufficiently large number. We choose $\delta$ and $T$ so that the following argument works well.

The variable $t$ means the natural projection $t : X \to \mathbb{R}$.

6.1. Cut and paste. Let $I$ be an ASD connection on $E$ defined in Example 1.1. For $s \in \mathbb{R}$ let $I_s := (-s)^*(I)$ be the pull-back of $I$ by $(-s) : E \to E$; $I_s$ is an ASD connection on $E$ with

$$|F(I_s)|_{\mathrm{op}} = \frac{4}{(e^{t-s} + e^{-t+s})^2}, \quad \|F(I_s)\|_{\mathrm{op}} = 1.$$

Most of its energy is contained in a neighborhood of $t = s$.

We define $J \subset \mathbb{Z}$ as the set of $n \in \mathbb{Z}$ satisfying $\|F_A\|_{L^\infty(nT<t<(n+1)T)} < \delta$. Since $A$ is degenerate, $J$ is an infinite set. In this subsection we describe a "cut and paste" procedure: We cut and paste the instanton $I_{nT+\frac{T}{2}}$ to $A$ over $[nT, (n+1)T] \times S^3$ for each $n \in J$. The resulting new connection will be denoted by $B$. ($B$ is not ASD in general.)

For simplicity of the notation, we suppose $0 \in J$, and we explain the cut and paste procedure over the region $[0,T] \times S^3$. Let $\varphi : X \to [0,1]$ be a cut-off function such that

$$\varphi = 0 \text{ on } \{t \leq T/3\} \cup \{t \geq 2T/3\}, \quad \varphi = 1 \text{ on } \{T/3+1 \leq t \leq 2T/3-1\}.$$

Set $U := \{T/3-1 < t < T/3+2\} \cup \{2T/3-2 < t < 2T/3+1\} \subset X$. Since $T \gg 1$ and $\|F_A\|_{L^\infty(0<t<T)} < \delta \ll 1$, we can choose connection matrices of $A$ and $I_{T/2}$ over $U$ such that

$$\|A\|_{C^k(U)} < C(k)\delta, \quad \|I_{T/2}\|_{C^k(U)} < C(k)\delta \quad (\forall k \geq 0).$$

Then we define a connection $B$ on $[0,T] \times S^3$ by

$$B := \begin{cases} A & \text{on } \{0 \leq t \leq T/3\} \cup \{2T/3 \leq t \leq T\}, \\ (1-\varphi)A + \varphi I_{T/2} & \text{on } U, \\ I_{T/2} & \text{on } \{T/3+1 \leq t \leq 2T/3-1\}. \end{cases}$$

In the same way we construct a connection $B$ by cutting and pasting the instanton $I_{nT+\frac{T}{2}}$ to $A$ over $[nT,(n+1)T] \times S^3$ for every $n \in J$.

Since $\delta \ll 1$ and $\|F_I\|_{\mathrm{op}} = 1 < d - \tau$, the connection $B$ satisfies

$$\|F_B\|_{\mathrm{op}} < d - \tau. \tag{24}$$

For $n \notin J$, we have $B = A$ over $nT \leq t \leq (n+1)T$. For $n \in J$, we have

$$\frac{1}{8\pi^2 T}\int_{nT<t<(n+1)T} |F_A|^2 d\mathrm{vol} < \frac{\delta^2 \mathrm{vol}(S^3)}{8\pi^2} < \frac{\varepsilon}{2} \quad (\delta \ll \varepsilon).$$

From this estimate we get

$$\rho(B) > \rho(A) - \frac{\varepsilon}{2}. \tag{25}$$

Moreover, $B$ satisfies the following non-degeneracy condition (cf. Lemma 1.7):

$$\begin{aligned} &\|F_B\|_{L^\infty(nT<t<(n+1)T)} \geq \delta \quad \text{for } n \notin J, \\ &\|F_B\|_{L^\infty(nT<t<(n+1)T)} \geq \|F_I\|_{L^\infty(-1<t<1)} \geq 1 \quad \text{for } n \in J. \end{aligned} \tag{26}$$

Therefore $B$ satisfies almost all the desired conditions. The only one problem is that $B$ is not ASD. But $B$ is an approximately ASD connection: $F_B^+$ is supported in

$$\bigcup_{n\in J}\left(\left\{nT + \frac{T}{3} \leq t \leq nT + \frac{T}{3} + 1\right\} \cup \left\{nT + \frac{2T}{3} - 1 \leq t \leq nT + \frac{2T}{3}\right\}\right).$$

Since $\delta \ll 1$,

$$\|F_B^+\|_{L^\infty(X)} \lesssim \delta, \quad \|\nabla_B F_B^+\|_{L^\infty(X)} \lesssim \delta. \tag{27}$$

6.2. Perturbation. In this subsection we construct an ASD connection $\tilde{A}$ by slightly perturbing the connection $B$ constructed in the previous subsection. We want to solve the equation $F^+(B + d_B^*\phi) = 0$ for $\phi \in \Omega^+(\mathrm{ad}E)$. By using the Weitzenböck formula [9, Chapter 6],

$$\begin{aligned} F^+(B + d_B^*\phi) =& F_B^+ + d_B^+ d_B^*\phi + (d_B^*\phi \wedge d_B^*\phi)^+ \\ =& F_B^+ + \frac{1}{2}\Big(\nabla_B^*\nabla_B + \frac{S}{3}\Big)\phi + F_B^+ \cdot \phi + (d_B^*\phi \wedge d_B^*\phi)^+ \end{aligned}$$

where $S$ is the scalar curvature of $X = \mathbb{R} \times S^3$; $S$ is a positive constant. The following fact on the operator $(\nabla_B^*\nabla_B + S/3)$ is proved in [18, Appendix, Proposition A.7, Lemmas A.1, A.2].

Lemma 6.1: *For any smooth $\xi \in \Omega^+(\mathrm{ad}E)$ with $\|\xi\|_{L^\infty} < \infty$, there uniquely exists a smooth $\phi \in \Omega^+(\mathrm{ad}E)$ satisfying*

$$\|\phi\|_{L^\infty} < \infty, \quad \Big(\nabla_B^*\nabla_B + \frac{S}{3}\Big)\phi = \xi.$$

*We will denote this $\phi$ by $(\nabla_B^*\nabla_B + S/3)^{-1}\xi$. It satisfies*

$$|\phi(x)| \le \int_X g(x,y)|\xi(y)|\mathrm{dvol}(y), \quad \|\phi\|_{L^\infty} \lesssim \|\xi\|_{L^\infty}.$$

*Here $g(x,y) > 0$ is the Green kernel of the operator $\nabla^*\nabla + S/3$ (this is the operator acting on functions). It is positive and uniformly integrable:*

$$\int_X g(x,y)\mathrm{dvol}(y) \lesssim 1 \quad (\textit{independent of } x).$$

*Moreover, it decays exponentially: For $d(x,y) > 1$,*

$$0 < g(x,y) \lesssim e^{-\sqrt{S/3}\, d(x,y)} \quad (d(x,y)\textit{: distance between } x \textit{ and } y).$$

Lemma 6.2: *Suppose $\xi \in \Omega^+(\mathrm{ad}E)$ is smooth and $\|\xi\|_{L^\infty} < \infty$. Then*

$$\phi := (\nabla_B^*\nabla_B + S/3)^{-1}\xi$$

*satisfies*

$$\|\phi\|_{L^\infty} + \|\nabla_B\phi\|_{L^\infty} \lesssim \|\xi\|_{L^\infty}.$$

*Proof.* $\|\phi\|_{L^\infty} \lesssim \|\xi\|_{L^\infty}$ was already given in Lemma 6.1. So we want to prove $\|\nabla_B \phi\|_{L^\infty} \lesssim \|\xi\|_{L^\infty}$. From the compactness of $\mathcal{M}_d$ (or the Uhlenbeck compactness) and the construction of $B$, for any $s \in \mathbb{R}$ we can choose a connection matrix of $B$ over $(s, s+1) \times S^3$ satisfying

$$\|B\|_{C^k(s<t<s+1)} \lesssim C(k) \quad (\forall k \geq 0).$$

Then from the $L^p$-estimate (Gilbarg–Trudinger [10, Theorem 9.11]) and $\|\phi\|_{L^\infty} \lesssim \|\xi\|_{L^\infty}$, for $1 < p < \infty$,

$$\|\phi\|_{L^p_{2,B}(s+1/4<t<s+3/4)} \lesssim C(p)\|\xi\|_{L^\infty(X)}. \tag{28}$$

Then the desired estimate $\|\nabla_B \phi\|_{L^\infty} \lesssim \|\xi\|_{L^\infty}$ follows from the Sobolev embedding $L^p_1 \hookrightarrow C^0$ $(p > 4)$. ∎

Set $\phi := 2(\nabla_B^* \nabla_B + S/3)^{-1}\xi$ where $\xi \in \Omega^+(\mathrm{ad}E)$ is smooth and $\|\xi\|_{L^\infty} < \infty$. We want to solve the equation $F^+(B + d_B^*\phi) = 0$, i.e.,

$$\xi = -F_B^+ - F_B^+ \cdot \phi - (d_B^*\phi \wedge d_B^*\phi)^+.$$

Set $Q(\xi) := -F_B^+ - F_B^+ \cdot \phi - (d_B^*\phi \wedge d_B^*\phi)^+$. From $\|F_B^+\|_{L^\infty} \lesssim \delta$ and Lemma 6.2,

$$\|Q(\xi) - Q(\eta)\|_{L^\infty} \lesssim (\delta + \|\xi\|_{L^\infty} + \|\eta\|_{L^\infty})\|\xi - \eta\|_{L^\infty}.$$

Then we can easily check that (when $\delta \ll 1$) the sequence $\{\xi_n\} \subset \Omega^+(\mathrm{ad}E)$ defined by

$$\xi_0 := 0, \quad \xi_{n+1} := Q(\xi_n)$$

satisfies $\|\xi_n\|_{L^\infty} \lesssim \delta$ (the implicit constant is independent of $n$) and becomes a Cauchy sequence in $L^\infty(X)$. Let $\xi_n \to \xi_\infty$ in $L^\infty(X)$. We have $\|\xi_\infty\|_{L^\infty} \lesssim \delta$. We will show that $\xi_\infty$ is smooth and satisfies $Q(\xi_\infty) = \xi_\infty$.

Set $\phi_n := 2(\nabla_B^* \nabla_B + S/3)^{-1}\xi_n$. Then

$$\xi_{n+1} = Q(\xi_n) = -F_B^+ - F_B^+ \cdot \phi_n - (d_B^*\phi_n \wedge d_B^*\phi_n)^+. \tag{29}$$

From the above (28) and $\|\xi_n\|_{L^\infty} \lesssim \delta$, the sequence $\{\phi_n\}$ is bounded in $L^p_{2,B}(K)$ for every $1 < p < \infty$ and compact subset $K \subset X$. Then from the equation (29) the sequence $\{\xi_n\}$ is bounded in $L^p_{1,B}(K)$. In the same way (the standard bootstrapping argument) we can show that the sequence $\{\xi_n\}$ is bounded in $L^p_{k,B}(K)$ for every $k \geq 0$, $1 < p < \infty$ and compact subset $K \subset X$. Therefore $\xi_\infty$ is smooth, and $\xi_n$ converges to $\xi_\infty$ in $C^\infty$ over every compact subset. Then

$$\xi_\infty = -F_B^+ - F_B^+ \cdot \phi_\infty - (d_B^*\phi_\infty \wedge d_B^*\phi_\infty)^+ \quad (\phi_\infty := 2(\nabla_B^* \nabla_B + S/3)^{-1}\xi_\infty). \tag{30}$$

Set $\tilde{A} := B + d_B^* \phi_\infty$. The connection $\tilde{A}$ is ASD. The rest of the work is to show that $\tilde{A}$ is non-degenerate and satisfies the condition (23).

From Lemma 6.2, $\|\phi_\infty\|_{L^\infty} + \|\nabla_B \phi_\infty\|_{L^\infty} \lesssim \|\xi_\infty\|_{L^\infty} \lesssim \delta$. Moreover, the equation

$$d_B^+ d_B^* \phi_\infty + (d_B^* \phi_\infty \wedge d_B^* \phi_\infty)^+ = -F_B^+$$

and $\|F_B^+\|_{L^\infty} + \|\nabla_B F_B^+\|_{L^\infty} \lesssim \delta$ (see (27)) implies $\|\nabla_B \nabla_B \phi_\infty\|_{L^\infty} \lesssim \delta$. (See the proof of Lemma 6.2.) Hence the curvature

$$F(\tilde{A}) = F_B + d_B d_B^* \phi_\infty + d_B^* \phi_\infty \wedge d_B^* \phi_\infty$$

satisfies $\|F(\tilde{A}) - F_B\|_{L^\infty} \lesssim \delta$. Since $B$ satisfies $\|F_B\|_{\mathrm{op}} < d - \tau$ and $\rho(B) > \rho(A) - \varepsilon/2$ (see (24) and (25)), if $\delta = \delta(\varepsilon, \tau) \ll 1$, we get

$$\|F(\tilde{A})\|_{\mathrm{op}} < d, \quad \rho(\tilde{A}) > \rho(A) - \varepsilon.$$

Therefore $\tilde{A}$ satisfies the condition (23).

Finally, we show that $\tilde{A}$ is non-degenerate. It is enough to prove that for all $n \in \mathbb{Z}$ the connection $\tilde{A}$ satisfies (see Lemma 1.7)

$$\|F(\tilde{A})\|_{L^\infty(nT < t < (n+1)T)} > \delta/2. \tag{31}$$

When $n \in J$, we have

$$\|F_B\|_{L^\infty(nT < t < (n+1)T)} \geq 1 \quad \text{(see (26))}$$

and

$$\|F(\tilde{A}) - F_B\|_{L^\infty} \lesssim \delta \ll 1.$$

So the above (31) holds for $n \in J$.

Choose $n \notin J$. For simplicity, we suppose $n = 0$. From the Green kernel estimate in Lemma 6.1,

$$|\phi_\infty(x)| \leq 2 \int_X g(x, y) |\xi_\infty(y)| d\mathrm{vol}(y).$$

From (30) and $|F_B^+|, |\phi_\infty|, |\nabla_B \phi_\infty| \lesssim \delta$,

$$|\xi_\infty| \lesssim |F_B^+| + \delta^2.$$

Since $0 \notin J$, the distance between $(-1, T+1) \times S^3$ and $\mathrm{supp}(F_B^+)$ is $\gtrsim T$. The Green kernel $g(x, y)$ decays exponentially. So if we choose $T = T(\varepsilon, \tau, \delta)$ sufficiently large, then

$$\|\phi_\infty\|_{L^\infty(-1 < t < T+1)} \lesssim \delta^2.$$

Now $\phi_\infty$ satisfies the following equation over $(-1, T+1) \times S^3$:

$$d_B^+ d_B^* \phi_\infty = -(d_B^* \phi_\infty \wedge d_B^* \phi_\infty)^+.$$

Since $\|d_B^* \phi_\infty \wedge d_B^* \phi_\infty\|_{L^\infty} \lesssim \|\nabla_B \phi_\infty\|_{L^\infty}^2 \lesssim \delta^2$, the bootstrapping argument shows that

$$\|\nabla_B \nabla_B \phi_\infty\|_{L^\infty(0<t<T)} \lesssim \delta^2.$$

Therefore, $|F(\tilde{A}) - F_B| \lesssim \delta^2$ over $(0, T) \times S^3$. Since $\|F_B\|_{L^\infty(0<t<T)} \geq \delta$ (see (26)) and $\delta \ll 1$, we get (31) for $n = 0$. We have finished the proof of Theorem 1.11.

*Remark 6.3:* If we start with the trivial flat connection $A$ in this gluing argument, then we can make the argument invariant under the action of the subgroup $T\mathbb{Z} \subset \mathbb{R}$. Then the resulting non-degenerate ASD connection $\tilde{A}$ becomes periodic (Example 1.6). So we can conclude that the space $\mathcal{M}_d$ $(d > 1)$ always contains a non-flat periodic ASD connection.

## Appendix A. Another ASD moduli space

Here we briefly discuss another possibility of the definition of the ASD moduli space. Let $X = \mathbb{R} \times S^3$ and $E = X \times SU(2)$ as in the main body of the paper. For $d \geq 0$ we define $\mathcal{N}_d$ as the space of the gauge equivalence classes of ASD connections $A$ on $E$ satisfying

$$\|F_A\|_{L^\infty(X)} \leq d.$$

Note that here we use the $L^\infty$-norm, which is different from the operator norm used in the definition of $\mathcal{M}_d$. The space $\mathcal{N}_d$ is endowed with the topology of $C^\infty$-convergence over compact subsets; $\mathcal{N}_d$ is compact and metrizable, and it admits a natural $\mathbb{R}$-action. The paper [18] studies the mean dimension and local mean dimension of this $\mathcal{N}_d$. In particular, [18, Theorem 1.2] shows the following upper bound on the local mean dimension:

Theorem A.1: *For any $[A] \in \mathcal{N}_d$,*

$$\dim_{[A]}(\mathcal{N}_d : \mathbb{R}) \leq 8\rho(A).$$

If $A$ is an ASD connection on $E$, then the operator norm $|F_A|_{\mathrm{op}}$ and the Euclidean norm $|F_A|$ bound each other by

$$\frac{1}{\sqrt{3}}|F_A| \leq |F_A|_{\mathrm{op}} \leq |F_A|.$$

(This uses the ASD condition.) Hence

$$\mathcal{N}_d \subset \mathcal{M}_d \subset \mathcal{N}_{\sqrt{3}\,d}.$$

Then for any $[A] \in \mathcal{M}_d$,

$$\dim_{[A]}(\mathcal{M}_d : \mathbb{R}) \leq \dim_{[A]}(\mathcal{N}_{\sqrt{3}\,d} : \mathbb{R}) \leq 8\rho(A).$$

This is Theorem 1.9 in Section 1.2. From knowledge on $\mathcal{M}_d$ we can prove the results on $\mathcal{N}_d$:

THEOREM A.2: *Let $A$ be a non-degenerate ASD connection on $E$ with $\|F_A\|_{L^\infty} < d$. Then*

$$\dim_{[A]}(\mathcal{N}_d : \mathbb{R}) = 8\rho(A).$$

*Proof.* We assume that $\mathcal{M}_d$ is endowed with a distance and that $\mathcal{N}_d$ is endowed with its restriction. Then we have $B_r([A]; \mathcal{N}_d)_{\mathbb{R}} = B_r([A]; \mathcal{M}_d)_{\mathbb{R}}$ for sufficiently small $r > 0$. Hence by Theorem 1.10,

$$\dim_{[A]}(\mathcal{N}_d : \mathbb{R}) = \dim_{[A]}(\mathcal{M}_d : \mathbb{R}) = 8\rho(A). \qquad \blacksquare$$

THEOREM A.3: *Suppose $d > \sqrt{3}$, and let $A$ be an ASD connection on $E$ with $\|F_A\|_{L^\infty} < d$. For any $\varepsilon > 0$ there exists a non-degenerate ASD connection $\tilde{A}$ on $E$ satisfying*

$$\|F(\tilde{A})\|_{L^\infty} < d, \quad \rho(\tilde{A}) > \rho(A) - \varepsilon.$$

*Proof.* The point is that the instanton $I$ defined in Example 1.1 satisfies

$$|F_I(t,\theta)| = \frac{4\sqrt{3}}{(e^t + e^{-t})^2}, \quad \|F_I\|_{L^\infty} = \sqrt{3}.$$

Then the gluing construction in Section 6 gives the result. $\blacksquare$

Let $\rho_{\mathcal{N}}(d)$ be the supremum of $\rho(A)$ over $[A] \in \mathcal{N}_d$. Let $\mathcal{D}_{\mathcal{N}} \subset [0, +\infty)$ be the set of left-discontinuous points of $\rho_{\mathcal{N}}(d)$. This is at most countable. From the above theorems, we can prove the following theorem. (The proof is the same as the proof of Theorem 1.2.)

THEOREM A.4: *For any $d \in (\sqrt{3}, +\infty) \setminus \mathcal{D}_{\mathcal{N}}$,*

$$\dim_{loc}(\mathcal{N}_d : \mathbb{R}) = 8\rho_{\mathcal{N}}(d).$$

So if $d > \sqrt{3}$ we have a good understanding of the local mean dimension of $\mathcal{N}_d$. For $d < 1$, $\mathcal{N}_d = \mathcal{M}_d = \{[\text{flat connection}]\}$ is the one-point space

(Example 1.1). The remaining problem is the case of $1 \leq d \leq \sqrt{3}$. We don't have any good information about this range.

The main good property of the operator norm $\|F_A\|_{\mathrm{op}}$ is our knowledge of the sharp threshold value described in Example 1.1.

## References


[1] S. Angenent, *The shadowing lemma for elliptic PDE*, in *Dynamics of Infinite-dimensional Systems (Lisbon, 1986)*, Nato Advanced Science Institutes Series F: Computer and Systems Sciences, Vol. 37, Springer, Berlin, 1987, pp. 7-22.

[2] M. F. Atiyah, *Elliptic operators, discrete groups, and von Neumann algebras*, in *Colloque Analyse et Topologie en l'Honneur de Henri Cartan (Orsay, 1974)*, Astérisque, Vol. 32–33, Société Mathématique de France, Paris, 1976, pp. 43–72.

[3] M. F. Atiyah, N. J. Hitchin and I. M. Singer, *Self-duality in four-dimensional Riemannian geometry*, Proceedings of the Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences **362** (1978), 425–461.

[4] A. A. Belavin, A. M. Polyakov, A. S. Schwartz and Y. S. Tyupkin, *Pseudo-particle solutions of the Yang–Mills equations*, Physics Letters **59** (1975), 85–87.

[5] R. Brody, *Compact manifolds and hyperbolicity*, Transactions of the American Mathematical Society **235** (1978), 213–219.

[6] S. K. Donaldson, *Floer Homology Groups in Yang–Mills Theory*, Cambridge Tracts in Mathematics, Vol. 147, Cambridge University Press, Cambridge, 2002.

[7] S. K. Donaldson and P. B. Kronheimer, *The Geometry of Four-manifolds*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1990.

[8] A. Eremenko, *Normal holomorphic curves from parabolic regions to projective spaces*, preprint, Purdue University (1998), arXiv: 0710.1281.

[9] D. S. Freed and K. K. Uhlenbeck, *Instantons and Four-manifolds*, Mathematical Sciences Research Institute Publications, Vol. 1, Springer-Verlag, New York, 1991.

[10] D. Gilbarg and N. S. Trudinger, *Elliptic Partial dDifferential Equations of Second Order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001.

[11] A. Gournay, *Dimension moyenne et espaces d'applications pseudo-holomorphes*, thesis, Département de Mathématiques d'Orsay, 2008.

[12] A. Gournay, *Complex surfaces and interpolation on pseudo-holomorphic cylinder*, arXiv: 1006.1775.

[13] A. Gournay, *Widths of $\ell^p$ balls*, Houston Journal of Mathematics **37** (2011), 1227–1248.

[14] M. Gromov, *Topological invariants of dynamical systems and spaces of holomorphic maps: I*, Mathematical Physics, Analysis and Geomery **2** (1999), 323–415.

[15] E. Lindenstrauss, *Mean dimension, small entropy factors and an embedding theorem*, Institut des Hautes Études Scientifiques. Publications Mathématiques **89** (1999), 227–262.

[16] E. Lindenstrauss and B. Weiss, *Mean topological dimension*, Israel Journal of Mathematics **115** (2000), 1–24.

[17] M. Macrì, M. Nolasco and T. Ricciardi, *Asymptotics for selfdual vortices on the torus and on the plane: a gluing technique*, SIAM Journal on Mathematical Analysis **37** (2005), 1–16.

[18] S. Matsuo and M. Tsukamoto, *Instanton approximation, periodic ASD connections, and mean dimension*, Journal of Functional Analysis **260** (2011), 1369–1427.

[19] S. Matsuo and M. Tsukamoto, *Brody curves and mean dimension*, Journal of the American Mathematical Society **28** (2015), 159–182.

[20] C. H. Taubes, *Self-dual Yang–Mills connections on non-self-dual 4-manifolds*, Journal of Differential Geometry **17** (1982), 139–170.

[21] M. Tsukamoto, *Gluing an infinite number of instantons*, Nagoya Mathematical Journal **188** (2007), 107–131.

[22] M. Tsukamoto, *Moduli space of Brody curves, energy and mean dimension*, Nagoya Mathematical Journal **192** (2008), 27–58.

[23] M. Tsukamoto, *Gauge theory on infinite connected sum and mean dimension*, Mathematical Physics, Analysis and Geometry **12** (2009), 325–380.

[24] M. Tsukamoto, *Deformation of Brody curves and mean dimension*, Ergodic Theory & Dynamical Systems **29** (2009), 1641–1657.

[25] M. Tsukamoto, *Sharp lower bound on the curvatures of ASD connections over the cylinder*, Journal of the Mathematical Society of Japan **66** (2014), 951–956.

[26] K. K. Uhlenbeck, *Connections with $L^p$ bounds on curvature*, Communicaations in Mathematical Physics **83** (1982), 31–42.

[27] K. Wehrheim, *Uhlenbeck compactness*, EMS Series of Lectures in Mathematics, European Mathematical Society, Zürich, 2004.

[28] K. Yosida, *On a class of meromorphic functions*, Proceedings of the Physico-Mathematical Society of Japan **16** (1934), 227–235.