FILTERED INSTANTON HOMOLOGY AND COSMETIC SURGERY

ALIAKBAR DAEMI, TYE LIDMAN, AND MIKE MILLER EISMEIER

Abstract. The cosmetic surgery conjecture predicts that for a non-trivial knot in the three-sphere, performing two different Dehn surgeries results in distinct oriented three-manifolds. Hanselman reduced the problem to ± 2 or $\pm 1/n$-surgeries being the only possible cosmetic surgeries. We remove the case of $\pm 1/n$-surgeries using the Chern–Simons filtration on Floer's original irreducible-only instanton homology, reducing the conjecture to the case of ± 2-surgery on genus 2 knots with trivial Alexander polynomial. Along the way, we establish a new surgery relationship for Floer's instanton homology and prove some similar results for surgeries on knots in $S^2 \times S^1$.

1. Introduction

1.1. **Cosmetic surgeries.** For a knot K in S^3, let the result of p/q-surgery be denoted $S^3_{p/q}(K)$. If U is the unknot, then $S^3_{p/q}(U) = L(p, -q)$ and hence there are infinitely many different surgeries on U that can produce orientation-preserving diffeomorphic three-manifolds. The *cosmetic surgery conjecture* [Gor91, Conjecture 6.1] (see also [Kir97, Problem 1.81A]) predicts that the unknot is very special in this regard:

Conjecture 1.1. *Let K be a non-trivial knot in S^3. If $p/q \neq p'/q'$, then $S^3_{p/q}(K)$ and $S^3_{p'/q'}(K)$ are not orientation-preserving diffeomorphic.*

For notation, we say that p/q and p'/q' are a cosmetic pair. Note that the orientations here are key. For example, $S^3_{p/q}(4_1)$ is orientation-reversing diffeomorphic to $S^3_{-p/q}(4_1)$ for any p/q since the figure-eight knot is amphichiral. As a less tautological example, the trefoil admits the chirally cosmetic surgery $S^3_9(3_1) \cong -S^3_{9/2}(3_1)$ [DM91].

There has been quite a lot of progress on this conjecture. For example, ∞ can never be part of a cosmetic pair [GL89, Theorem 2]. A sequence of results using Heegaard Floer homology [Wan06, OS11, Wu11, NW15, Han23] gave increasingly stronger constraints on the potential surgery slopes. In particular, [NW15, Theorem 1.2] and [Han23, Theorem 2(i)] combine to give:

Theorem 1.2. *If p/q, p'/q' are a cosmetic pair for a non-trivial knot $K \subset S^3$, then $p/q = -p'/q'$. Furthermore, $p/q = \pm 2$ or $\pm 1/n$ for some non-zero integer n.*

We remark that these arguments using Heegaard Floer homology also give further constraints on the cosmetic surgery slopes given more information about the knot (e.g. its knot Floer homology). Hanselman verified the cosmetic surgery conjecture for all knots up to 16 crossings [Han23, Theorem 6]. There have also been many important advances using other aspects of low-dimensional topology. Detcherry found various obstructions using quantum invariants, and checked the conjecture for knots up to 17 crossings [Det21, Corollary 1.10]. Additionally, work of Futer–Purcell–Schleimer [FPS22, Theorem 7.29] rules out cosmetic pairs for hyperbolic knots roughly whenever the slopes have large enough length. They were then able to use a variety of bounds, including Chern–Simons invariants, to verify the cosmetic surgery conjecture for all knots up to 19 crossings [FPS24, Theorem 2.10]. For some other examples of the variety of progress on the cosmetic surgery conjecture, see [Ito23, SS21, Tao22].

TL was partially supported by NSF grant DMS-2105469.

There is some upper limit on what the Heegaard Floer homology techniques can give: Ozsváth–Szabó observed that the Heegaard Floer homology of the 1 and -1 surgeries on the knot 9_{44} are graded isomorphic [OS11]. In the current article, we incorporate the values of the Chern–Simons functional on instanton Floer homology to obtain more information.

Theorem 1.3. *If K is a non-trivial knot in S^3, then $S^3_{1/n}(K)$ and $S^3_{-1/n}(K)$ are not orientation-preserving diffeomorphic for any integer $n \neq 0$. Furthermore, they cannot be related by a ribbon homology cobordism.*

Recall that a ribbon homology cobordism is a homology cobordism that admits a handle decomposition without any 3-handles [DLVVW22].

Combining this theorem with the previous literature and a quick Casson invariant computation (see Theorem 6.1 below) yields:

Corollary 1.4. *If a non-trivial knot admits a cosmetic surgery, then the pair of slopes is ± 2, the knot has genus 2, the Alexander polynomial is trivial, and the value of the Jones polynomial at $e^{2\pi i/5}$ is 1. In particular, the cosmetic surgery conjecture holds for fibered knots, knots which are not topologically slice, and HFK-thin knots.*

Proof. That the only possible pair of cosmetic slopes is ± 2 is a combination of Theorems 1.2 and 1.3. Hanselman shows that if ± 2 is a cosmetic pair, then the knot has genus 2. The claim about the Jones polynomial of K follows from [Det21, Theorem 1.4]. Finally, in Theorem 6.1 below, we show that in this special case the Alexander polynomial must be trivial. Since non-trivial knots with Alexander polynomial one are non-fibered, are topologically slice [Fre84, FQ90], and have thick knot Floer homology [OS04b, OS04a], we have the desired result. □

Remark 1.5. Since any alternating knot is HFK-thin, Corollary 1.4 implies that the cosmetic surgery conjecture holds for alternating knots. For a large family of alternating knots, this was already verified in the recent work [IJ24].

Remark 1.6. Dave Futer has reported to the authors that there are in fact no non-trivial knots with at most 17 crossings which have both $J_K(e^{2\pi i/5}) = 1$ and $\Delta_K = 1$. We also remark that the results of [Det21] can be used to give further constraints for the remaining case of surgery slopes ± 2 in terms of colored Jones polynomial of K.

Remark 1.7. Combined with work of Ravelomanana [Rav18], this proves that there are no exceptional cosmetic surgeries on hyperbolic knots in S^3.

Using similar arguments, we are also able to prove a restricted cosmetic surgery statement in $S^2 \times S^1$.

Theorem 1.8. *Let L be a knot in $S^2 \times S^1$ which is homologous but not isotopic to $\{*\} \times S^1$. No two integral surgeries on L can produce the same oriented three-manifold. Furthermore, they cannot be related by a ribbon homology cobordism.*

Note that the three-manifolds described in Theorem 1.8 are all homology spheres. Furthermore, they are boundaries of *Mazur manifolds* and hence are homology cobordant to S^3. (A Mazur manifold is a four-manifold obtained by attaching a 2-handle to $S^1 \times D^3$ along a curve which generates H_1.) In particular, the homology spheres from Theorem 1.8 are homology cobordant to each other, but the ones provided by surgeries on the same knot are not related by homology cobordisms without any 3-handles.

A Mazur manifold is B^4 if and only if the attaching curve is isotopic to $\{*\} \times S^1$ [Gab87]. Theorem 1.8 has the following related four-dimensional consequence. We are not aware of any other three- or four-dimensional proof of this fact in general.

Corollary 1.9. *Let M be a Mazur manifold other than B^4. Changing the framing of the 2-handle in the Mazur description results in a four-manifold not oriented homeomorphic to M.*

As another application of Theorem 1.8, we are able to reprove and extend part of a theorem of Josh Wang [Wan22, Theorem 1.1].

Corollary 1.10. *Let K_0 be obtained as a non-trivial band surgery on a two-component unlink. Let K_n be the result of putting n full twists in the band. For any non-zero integer p, if $S^3_{1/p}(K_n) = S^3_{1/p}(K_m)$ then $m = n$. Consequently, K_n and K_m are distinct knots.*

Proof. Fix p. Let γ denote the meridian of the band, thought of as a knot in $S^3_{1/p}(K_0)$. This surgery is not S^3 as long as $p \neq 0$ and K_0 is non-trivial, which is the case when the band is non-trivial [Sch85]. Then $S^2 \times S^1$ is 0-surgery on γ, while $S^3_{1/p}(K_n)$ is the result of $-1/n$-surgery on γ, or alternatively, integral surgery on the core of γ in $S^2 \times S^1$. The result follows from Theorem 1.8. $\square$

Note that the statement is in some ways slightly stronger than Wang's theorem for full twists, as his results do not immediately distinguish the surgered manifolds. However, his results apply to a broader class of links, namely possibly half-twisted bandings between arbitrary two-component split links. Wang has proposed a strategy for generalizing Corollary 1.10 to this more general case, but it seems like the requisite technical machinery for instanton Floer homology is not currently developed.

Remark 1.11. While we do not do this here, similar ideas to the ones in this paper can be applied to prove analogous results in other manifolds. For example, suppose K is a knot in $\mathbb{RP}^3$ which is neither null-homologous nor a core curve of the genus 1 Heegaard splitting. Then the two integer homology spheres which arise as integral surgeries on K are distinct as oriented three-manifolds.

1.2. **The strategy and surgery exact triangles in instanton Floer homology.** We now describe our overarching strategy. For the introduction, we give broad strokes for non-experts. Instanton Floer homology morally assigns to a homology sphere a chain complex over $\mathbb{Z}[t, t^{-1}]$ whose generators corresponds to conjugacy classes of non-trivial $SU(2)$ representations of π_1. Any monomial generator α gets a real value from the Chern–Simons functional denoted $\mathrm{CS}(\alpha)$ and a $\mathbb{Z}$-grading denoted by $\mathrm{gr}(\alpha)$. Multiplication by t raises $\mathrm{gr}(\alpha)$ by 8 and $\mathrm{CS}(\alpha)$ by 1.[1] Furthermore, the differential has degree -1 with respect to gr and is filtered by the Chern–Simons functional. Our strategy is to show that the Chern–Simons filtered instanton homologies of different surgeries on knots are different. The strategy is roughly to find an isomorphism which strictly lowers the filtration. These maps will come from functoriality of instanton Floer complexes with respect to cobordisms.

A cobordism $W : Y \to Y'$ between integer homology spheres induces a map of instanton Floer homologies by counting certain objects called ASD connections, given a choice of bundle on W. In the simplest case, if W is a negative-definite cobordism and A is an ASD connection over the trivial $SU(2)$-bundle on W which is asymptotic to monomial generators α at Y and α' at Y', then $\mathrm{CS}(\alpha) \geq \mathrm{CS}(\alpha')$ and $\mathrm{gr}(\alpha) = \mathrm{gr}(\alpha')$. This inequality is strict if $\pi_1(W) = 0$. Consequently, if W is simply-connected, negative-definite, and $W : Y \to Y'$ induces an isomorphism on instanton homology, then we get a filtration decreasing isomorphism of Floer homologies. If the instanton Floer homology of Y is non-trivial, this is only possible if the filtered instanton Floer homologies of Y and Y' are different, and so the three-manifolds are not orientation-preserving diffeomorphic to

[1]For the experts, we consider flat connections modulo degree zero gauge transformations, so Chern–Simons can be viewed as real-valued and the associated Floer homology is $\mathbb{Z}$-graded. Of course, we also need to introduce perturbations everywhere to obtain the desired transversality results. We assume in this introduction that all moduli spaces are cut out transversely and so no perturbations are needed.

each other. This strategy is sufficient to prove Theorem 1.8, which we next describe. The strategy for $1/n$-surgeries is more complicated, and is discussed after the warm-up case of knots in $S^2 \times S^1$.

1.2.1. *The strategy to prove Theorem 1.8.* We will use Floer's original exact triangle for a knot K in a homology sphere Y:

$$\ldots \to I_*(Y) \to I_*(Y_{-1}(K)) \to I_*^w(Y_0(K)) \to \ldots \tag{1}$$

where $I_*^w(Y_0(K))$ is the admissible version of instanton homology for three-manifolds with w determining the bundle data [Flo90, BD95]. If $L \subset S^2 \times S^1$ is given as in the statement of Theorem 1.8, we apply (1) to the case that Y is given by an integral surgery on L and $K \subset Y$ is the dual knot. Then $Y_0(K) = S^2 \times S^1$ and $Y_{-1}(K)$ is the integral surgery on K where we increase the surgery coefficient by one. Since $I_*^w(S^2 \times S^1) = 0$, the map $I_*(Y) \to I_*(Y_{-1}(K))$ in (1) is an isomorphism. This map is induced by the 2-handle cobordism $W : Y \to Y_{-1}(K)$, which is negative definite and simply-connected. Furthermore, $I_*(Y)$ is non-zero because L is not isotopic to $\{*\} \times S^1$ [LPCZ23, Theorem 1.3]. Thus the observation in the previous paragraph can be applied to W to show that Y and $Y_{-1}(K)$ are not orientation-preserving diffeomorphic to each other. By stacking such cobordisms together, we can more generally show that any two integral surgeries on K are not orientation-preserving diffeomorphic to each other. The proof of the second part of Theorem 1.8 uses a result from [DLVVW22] about the behavior of instanton Floer homology with respect to ribbon homology cobordisms to obtain a similar contradiction.

1.2.2. *The strategy to prove Theorem 1.3 for $n = \pm 1$.* Unfortunately, it is generally impossible to construct a simply-connected, negative definite cobordism from $S^3_{1/n}(K)$ to $S^3_{-1/n}(K)$ whose usual cobordism map is an isomorphism. Such a map would necessarily have degree zero, but for example $S^3_1(3_1)$ and $S^3_{-1}(3_1)$ have their instanton Floer homology supported in different gradings; we need a variation on the constructions above to make this work. We begin with ± 1-surgery.

In order to relate $S^3_1(K)$ and $S^3_{-1}(K)$, we need to establish another surgery exact triangle — which we call the *distance-two surgery triangle* — and extract information from it in a novel way. The exact triangle we prove is predicted in [CDX20], who prove an analogous version for Floer homology with admissible bundles:

Theorem 1.12. *Let K be a knot in a homology sphere Y. Then there is an exact triangle of the following form:*

$$\ldots I_*(Y_1(K)) \to I_*(Y) \oplus I_*(Y) \to I_*(Y_{-1}(K)) \to \ldots \tag{2}$$

Remark 1.13. A similar exact triangle for the Heegaard Floer homology groups $\widehat{HF}$ is established as [OS08, Theorem 3.1]. While the groups $\widehat{HF}$ should be understood as analogous to the instanton homology groups $I^{\#}(Y)$, there is no known analogue of Floer's irreducible instanton homology groups $I_*(Y)$ in Heegaard Floer theory.

Taking $Y = S^3$, we obtain an isomorphism from $I_*(S^3_{-1}(K))$ to $I_*(S^3_1(K))$ because $I_*(S^3) = 0$. Unfortunately, the cobordism $W : S^3_{-1}(K) \to S^3_1(K)$ inducing the isomorphism has $b^+ > 0$, and it does not behave in the desired way with respect to the Chern–Simons filtration; the existence of this map does not lead to a contradiction.

However, this is not the only cobordism map available to us. In the proof of the exact triangle, one needs to find a nullhomotopy of the chain map that represents the composition

$$I_*(Y_1(K)) \to I_*(Y) \oplus I_*(Y) \to I_*(Y_{-1}(K)).$$

This nullhomotopy comes from a count of ASD connections over a 1-parameter family of metrics on a different cobordism W'. In the case of $Y = S^3$, this count turns out to be a chain map g_1, and even better, a quasi-isomorphism which is a homotopy inverse of the cobordism map for

W. (It is important to emphasize that the 1-parameter family map is different from the usual cobordism map for W', which is identically 0.) The cobordism W' is negative-definite and if $\alpha_\pm$ on $S^3_{\pm 1}(K)$ are related by an ASD connection over W', then $8\mathrm{CS}(\alpha_+) - \mathrm{gr}(\alpha_+) > 8\mathrm{CS}(\alpha_-) - \mathrm{gr}(\alpha_-)$; that is, W' induces a strictly filtered map with respect to a degree-shifted Chern–Simons filtration. The instanton Floer complex is finitely generated over $\mathbb{Z}[t, t^{-1}]$ and thus, the function $8\mathrm{CS} - \mathrm{gr}$ is bounded on the collection of all monomial generators. Because g_1 is strictly filtered, this once again implies that the filtered Floer homologies of $S^3_1(K)$ and $S^3_{-1}(K)$ are not isomorphic.

This can be explained more succinctly by defining an invariant $\ell(Y)$, which is loosely the minimal value of $8\mathrm{CS} - \mathrm{gr}$ on a generator of instanton homology. The exact triangle (2) is used to show that $\ell(S^3_1(K)) > \ell(S^3_{-1}(K))$. By further applying (1) and (2), we obtain a sequence of inequalities

$$\ldots \ell(S^3_{1/2}(K)) > \ell(S^3_1(K)) > \ell(S^3_{-1}(K)) > \ell(S^3_{-1/2}(K)) \ldots \tag{3}$$

which completes the proof that $S^3_{1/n}(K)$ is not orientation-preserving diffeomorphic to $S^3_{-1/n}(K)$. (The additional inequalities do not require the use of the nullhomotopy maps.) The claim about ribbon homology cobordisms follows as in the proof of Theorem 1.8.

Remark 1.14. It is worth pointing out that this argument does not require knowing the theorem of Ni–Wu that $S^3_{p/q}(K) = S^3_{p/q'}(K)$ implies $q = \pm q'$.

Remark 1.15. Cobordism maps with respect to families of metrics are widely used in Floer theory; for instance, these are used to prove that the cobordism map is independent of metric, and they are used in the proof of Floer's exact triangle. They are also relevant in the study of diffeomorphism groups of 3- and 4-manifolds. To the best of the authors' knowledge, however, Theorem 1.3 is the first known example where such a map is used to define an isomorphism of Floer homology groups.

In [LLP23, Proposition 1.15], it was shown that the existence of a non-trivial knot K with $S^3_1(K) = S^3_{-1}(K)$ would produce an exotic $S^1 \times S^3 \# S^2 \times S^2$. This is now ruled out by Theorem 1.3. Similarly, if there exist non-trivial knots $K_1, \ldots, K_n$ such that $S^3_1(K_i) = S^3_{-1}(K_{i+1})$ for $i = 1, \ldots, n$ with indices computed mod n, then there exists an exotic $S^3 \times S^1 \#_n S^2 \times S^2$. By the inequalities in (3), we see this is also impossible.

Theorem 1.16. *Let $K_1, \ldots, K_n$ be a sequence of knots and $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ integers with $a_i > b_{i+1}$ with indices computed mod n. If $S^3_{1/a_i}(K_i) = S^3_{1/b_{i+1}}(K_{i+1})$ for all i, then all K_i are unknotted.*

Remark 1.17. It is natural to wonder whether our arguments can be applied to the case of ± 2 surgery. Even though $S^3_{\pm 2}(K)$ are not integer homology spheres, these still have a well-defined filtered irreducible instanton homology $I_*(S^3_{\pm 2}(K))$, and one can define invariants $\ell(Y_{\pm 2})$. We have $\ell(S^3_1(K)) > \ell(S^3_2(K))$ and $\ell(S^3_{-2}(K)) > \ell(S^3_{-1}(K))$, but our initial investigations suggest that it is more difficult to compare the values $\ell(S^3_{\pm 2}(K))$ even though these manifolds have isomorphic Floer homology. In addition, while there exist variations on Theorem 1.12 relating the instanton homologies of ± 2 surgery, the variations known to the authors at the time of writing are all inadequate in some way: for one variation, the complex replacing $C_*(S^3) \oplus C_*(S^3)$ is no longer trivial so that the map g_1 is no longer a chain map; for another variation, this complex is trivial and g_1 is a chain map, but it behaves unfavorably with respect to the Chern–Simons filtration. We plan to explore this circle of ideas elsewhere.

Organization

In Section 2, we review the basics of persistent homology, including a variation in the instanton setting. We then introduce the invariant ℓ used above and verify its basic properties. In Section 3, we recall the necessary background material on filtered instanton Floer homology, and state the new results relating to the distance-two exact triangle. In Section 4 we prove Theorems 1.8 and 1.3, taking for granted the distance-two exact triangle of Theorem 1.12; we establish the existence of this exact triangle in Section 5, which comprises about half of the paper. The final Section 6 gives a short proof that if K admits cosmetic surgeries then $\Delta_K(t) = 1$.

Acknowledgments

We gladly thank Josh Wang for some critical discussions from which this paper was developed and Peter Kronheimer for pointing out a crucial error in a primitive version of this strategy. We also thank Dave Futer and Jonathan Hanselman for helpful comments.

2. Filtered groups and numerical invariants

Real-valued lifts of the Chern–Simons functional have been used extensively to define filtered algebraic structures and extract numerical invariants of 3-manifolds, and these numerical invariants have had extensive topological applications [Fur90, HK11, Dae20, NST24].

Here we will use the language of persistent homology to define the relevant algebraic structure. A *persistence module* is a collection of R-modules F_rV indexed by $r \in \mathbb{R} \cup \{\infty\}$, together with connecting homomorphisms $i_r^{r'} : F_rV \to F_{r'}V$ when $r \le r'$ which are functorial in the sense that $i_{r'}^{r''} i_r^{r'} = i_r^{r''}$ [ELZ02, ZC05]. The basic example is the *interval module*, and over a field every persistence module decomposes into these basic pieces:

Example 2.1. Suppose $I \subset \mathbb{R}$ is an interval I. The *interval module* R_I has

$$F_rR_I = \begin{cases} R & r \in I \\ 0 & \text{otherwise,} \end{cases}$$

with connecting homomorphisms $i_r^{r'}$ equal to the identity if $r, r' \in I$ and equal to zero otherwise.

Theorem 2.2 ([CB15])**.** *Suppose V is a persistence module over a field k for which F_rV is finite-dimensional for all r. Then there exist a collection of intervals $I_1, I_2, \ldots$ and an isomorphism of persistence modules $V \cong \bigoplus_j k_{I_j}$.*

The collection of intervals in a direct sum decomposition $V = \bigoplus_j k_{I_j}$ is well-defined, and called the *barcode* of V. Notice that the statement above does not imply that there are finitely many intervals in the barcode for V, only that each $r \in \mathbb{R}$ lies in finitely many of these intervals.

In the instanton theory, the filtration has a natural 8-periodicity, so we introduce the closely related notion of *instanton persistence modules* or IP-modules.

Definition 2.3. An IP**-module** A over the ring R consists of the following data.

(i) For each integer d and each $r \in \mathbb{R} \cup \{\infty\}$, an R-module F_rA_d, and an isomorphism $\varphi_{r,d} : F_rA_d \to F_{r+1}A_{d+8}$.
(ii) For each integer d and each $r \le r'$, a homomorphism $i_r^{r'} : F_rA_d \to F_{r'}A_d$.

We demand that these satisfy the following:

(a) The map i_r^r is the identity, $i_{r'}^{r''} i_r^{r'} = i_r^{r''}$, and $\varphi_{r',d} i_r^{r'} = i_{r+1}^{r'+1} \varphi_{r,d}$.
(b) For each d, we have $F_rA_d = 0$ for sufficiently small $r \in \mathbb{R}$, while the map $i_r^\infty : F_rA_d \to F_\infty A_d$ is an isomorphism for sufficiently large $r \in \mathbb{R}$.

When there is no risk of confusion, we will often write A_d in place of $F_\infty A_d$.

In the case of an IP-module V over a field k, with each F_rV_d finite-dimensional, so that $V_d \cong \bigoplus_j k_{I_{j,d}}$, the periodicity property is essentially equivalent to the statement that the barcode in degree $d + 8$ is a shift of the barcode in degree d. More precisely, the number of intervals $I_{j,d}$ and

$I_{j,d+8}$ is the same, and each interval $I_{j,d+8}$ of V_{d+8} is the shift $I_{j,d}+1$ of a corresponding interval for V_d.

We will only use fairly crude numerical invariants, which ignore much of the structure of an IP-module.

Definition 2.4. If A is an IP-module, define $\kappa_A : \mathbb{Z} \to [-\infty, \infty]$ and $\ell(A) \in [-\infty, \infty]$ as

$$\begin{aligned} \kappa_A(d) &= \inf\{r \in \mathbb{R} \mid i_r^\infty : F_r A_d \to A_d \text{ is nonzero}\}, \\ \ell(A) &= \inf\{\kappa_A(d) - d/8 \mid d \in \mathbb{Z}\}. \end{aligned}$$

These are transparently invariant under IP-module isomorphisms (isomorphisms $F_r A_d \cong F_r B_d$ which commute with the structure maps). If we work over a field k and A is a direct sum of interval modules, then $\kappa_A(d)$ is the least r such that $[r, \infty]$ appears in the barcode of A_d, while $\ell(A)$ is the least r such that $[r + d/8, \infty]$ appears in the barcode of A_d for some d.

Lemma 2.5. *The κ and ℓ invariants satisfy the following properties.*

(a) We have $\kappa_A(d) > -\infty$ for all d, and $\kappa_A(d) = \infty$ if and only if $A_d = 0$.
(b) We have $\kappa_A(d+8) = \kappa_A(d) + 1$.
(c) We have $\ell(A) > -\infty$, and $\ell(A) = \infty$ if and only if $A_d = 0$ for all d.

Proof. The first claim follows immediately from axiom (b) of an IP-module. For the second claim, we may assume $\kappa_A(d) = r < \infty$. Suppose $x \in F_{r+\epsilon} A_d$ has $i^\infty_{r+\epsilon} x = y \neq 0$. Then by axiom (a), we have

$$i^\infty_{r+1+\epsilon} \varphi_{r+\epsilon,d}(x) = \varphi_{\infty,d} i^\infty_{r+\epsilon}(x) = \varphi_{\infty,d}(y) \neq 0.$$

The last claim holds because $y \neq 0$ and $\varphi_{\infty,d}$ is an isomorphism. Because $y \in F_{r+1} A_{d+8}$, it follows that $\kappa_A(d+8) \leq \kappa_A(d) + 1 + \epsilon$ for all $\epsilon > 0$, and hence that $\kappa_A(d+8) \leq \kappa_A(d) + 1$. The other inequality is similar. For the last claim, observe that $\kappa_A(d) - d/8$ is an 8-periodic function, and hence takes on only finitely many values. The minimum of these values is infinite if and only if $\kappa_A(d) = \infty$ for all d, so the result follows from (a). ☐

Next, we introduce the type of morphisms relevant to us.

Definition 2.6. Suppose A and B are IP-modules. An **IP-morphism $f : A \to B$ of degree D and level L** is a collection of homomorphisms $f_{r,d} : F_r A_d \to F_{r+L} B_{d+D}$ satisfying

$$f_{r',d} i_r^{r'} = i_{r+L}^{r'+L} f_{r,d}, \qquad f_{r+1,d+8} \varphi_{r,d} = \varphi_{r+L,d+D} f_{r,d}.$$

Lemma 2.7. *Suppose $f : A \to B$ is an IP-morphism of degree D and level L.*

(a) If $f_{\infty,d} : A_d \to B_{d+D}$ is injective, then $\kappa_B(d+D) \leq \kappa_A(d) + L$.
(b) If $f_{\infty,d}$ is injective for all d, then $\ell(B) \leq \ell(A) + (L - D/8)$.

Proof. If $\kappa_A(d) = \infty$ the first claim is vacuous, so suppose $\kappa_A(d) = r < \infty$, and that $x \in F_{r+\epsilon} A_d$ has $i^\infty_{r+\epsilon}(x) \neq 0$. Setting $y = f_{r+\epsilon,d}(x) \in F_{r+\epsilon+L} B_{d+D}$, we have

$$i^\infty_{r+\epsilon+L}(y) = f_{\infty,d} i^\infty_{r+\epsilon}(x) \neq 0,$$

as $f_{\infty,d}$ is injective. It follows that $\kappa_B(d+D) \leq \kappa_A(d) + L + \epsilon$ for all $\epsilon > 0$, proving the first claim.

As for the second claim, we have

$$\begin{aligned} \ell(B) &= \inf\{\kappa_B(d+D) - d/8 - D/8 \mid d \in \mathbb{Z}\} \\ &\leq \inf\{\kappa_A(d) - d/8 + L - D/8 \mid d \in \mathbb{Z}\} = \ell(A) + (L - D/8). \end{aligned}$$ ☐

Remark 2.8. The above statement also has an interpretation in terms of barcodes when working over a field. If $f_{r',d}$ is injective, and $[r, r']$ is an interval in the barcode of A_d, then $[r+L, r'+L]$ is *contained* in an interval in the barcode of B_{d+D}. Lemma 2.7(a) above is the case $r' = \infty$.

We will later need a mild generalization of the above claim.

Lemma 2.9. *Suppose A, B are IP-modules, and $f^i : A \to B$ is an IP-morphism of degree D_i and level L_i for $i = 1, 2$. If the map $(f^1_\infty, f^2_\infty) : A_d \to B_{d+D_1} \oplus B_{d+D_2}$ is injective for all d, then we have*

$$\ell(B) \leq \ell(A) + \max(L_1 - D_1/8, L_2 - D_2/8).$$

Proof. If $\ell(A) = r$, for all $\epsilon > 0$ there exists some d and $x \in F_{r+\epsilon+d/8}A_d$ so that $i^\infty_{r+\epsilon+d/8}(x) \neq 0$. Because (f^1_∞, f^2_∞) is injective, for some $i \in \{1, 2\}$ we have

$$0 \neq f^i_\infty i^\infty_{r+\epsilon+d/8}(x) = i^\infty_{r+\epsilon+d/8+L_i}(f^i_{r+\epsilon+d/8}x).$$

As $f^i_{r+\epsilon+d/8}(x) \in F_{r+\epsilon+d/8+L_i}B_{d+D_i}$, this implies that

$$\begin{aligned}\ell(B) &\leq \kappa_B(d + D_i) - d/8 - D_i/8 \leq (r + \epsilon + d/8 + L_i) - d/8 - D_i/8 \\ &= \ell(A) + \epsilon + (L_i - D_i/8) \leq \ell(A) + \max(L_1 - D_1/8, L_2 - D_2/8) + \epsilon.\end{aligned}$$

Taking ϵ to zero completes the argument. ☐

3. Instanton Floer homology

In this section, we review the aspects of instanton Floer theory necessary to prove our main results. At first, we let R be an arbitrary commutative ring. Later we will specialize to $R = \mathbb{F}_2$.

Proposition 3.1. *Floer's instanton homology groups satisfy the following properties.*

(a) If Y is an integer homology sphere, there is an associated $\mathbb{Z}/8$-graded module $I_d(Y; R)$, invariant under orientation-preserving diffeomorphism and functorial under cobordisms $(W, c) : Y \to Y'$ with $b_1(W) = b^+(W) = 0$, where $c \subset W$ is a closed, oriented, embedded surface.

(b) If Y' is a homology $S^2 \times S^1$, there is an associated $\mathbb{Z}/4$-graded module $I^w_d(Y'; R)$, again invariant under orientation-preserving diffeomorphism.

(c) We have $I_d(S^3; R) = 0$ and $I^w_d(S^2 \times S^1; R) = 0$ for all d.

Proof. For integer homology spheres see [Flo88]. For homology $S^2 \times S^1$ see [Flo90]; this is the instanton Floer homology of Y' equipped with a $U(2)$-bundle with odd first Chern class. This bundle is a *non-trivial admissible bundle*, which supports no reducible connections. The vanishing result follows because neither S^3 nor $S^2 \times S^1$ (with the corresponding non-trivial $U(2)$ bundle) admits irreducible projectively flat $U(2)$-connections. ☐

More general functoriality statements hold, but we will not need them. We will also suppress the base ring R from notation.

In the case of integer homology spheres, these groups have a natural enrichment to IP-modules.

Proposition 3.2. *If Y is an integer homology sphere, the graded module $I_*(Y)$ may be enriched with the structure of an IP-module $I(Y)$, invariant under orientation-preserving diffeomorphism.*

Remark 3.3. Keep in mind that $I(Y)$ is the data of $F_r I_d(Y)$ for all $r \in \mathbb{R} \cup \{\infty\}$ and $d \in \mathbb{Z}$, together with connecting homomorphisms i and periodicity maps φ. On the other hand, $I_*(Y)$ is the data of $I_d(Y)$ for all $d \in \mathbb{Z}/8$. In particular, $I_*(Y)$ is the direct sum of $F_\infty I_d(Y)$ for all $d \in \mathbb{Z}/8$. (Note that our assumptions on IP-modules imply that $F_\infty I_d(Y)$ depends only mod 8 value of d.) In summary, $I_*(Y)$ is part of the information contained in $I(Y)$.

Proof. This claim is strictly weaker than the result of the main constructions of [Dae20, NST24]. (The language of IP-modules is closer to the language used in [NST24].) Both papers also pay close attention to the interaction with the reducible connection, which is irrelevant for our purposes. We briefly review the construction.

Write $\mathcal{B}^*(Y)$ for the space of irreducible $SU(2)$-connections on the trivial bundle over Y, modulo gauge. This space is equipped with the Chern–Simons functional $\mathrm{CS} : \mathcal{B}^*(Y) \to \mathbb{R}/\mathbb{Z}$. The universal cover of this space, the space of irreducible $SU(2)$-connections modulo degree-zero gauge transformations, carries a canonical $\mathbb{Z}$-periodic lift $\widetilde{\mathrm{CS}} : \widetilde{\mathcal{B}}^*(Y) \to \mathbb{R}$. We determine the lift by requiring that the natural extension to the space of *all* $SU(2)$-connections modulo degree-zero gauge transformations has $\widetilde{\mathrm{CS}}(\theta) = 0$, where θ is the trivial connection.

So long as r is not a critical value of the Chern–Simons function on Y, we define $F_r I_d(Y)$ to be the degree-d Morse homology of the sublevel set $\widetilde{\mathrm{CS}}^{-1}(-\infty, r]$, with respect to an appropriately perturbed Chern–Simons functional. The degree $i(\alpha)$ of an irreducible connection α is defined to be the index of the (appropriately perturbed) ASD operator D_A^+ associated to a connection on $\mathbb{R} \times Y$ which is equal to α at $-\infty$ and the trivial connection θ at $+\infty$.

For r not a critical value, that this Morse homology is well-defined is given as [NST24, Lemma 2.6] (when comparing, take $s = -\infty$). To simplify the definition of IP-module, we extend the definition of $F_r I_d(Y)$ to include the critical values by demanding this assignment be right-continuous: for r a critical value we set $F_r I_d(Y) = F_{r+\epsilon} I_d(Y)$ for sufficiently small $\epsilon > 0$. □

The functoriality of Proposition 3.1 extends to IP-morphisms.

Proposition 3.4. *Suppose $W : Y \to Y'$ is a cobordism between integer homology spheres with $b_1(W) = b^+(W) = 0$, and $c \subset W$ is an embedded oriented surface. Then there is a constant $\eta(W,c) \geq 0$ which is strictly positive if $\pi_1(W) = 0$ and an induced IP-morphism $(W,c)_* : I(Y) \to I(Y')$ of degree $D = -2c^2$ and level $L = -c^2/4 - \eta(W,c)$, which enriches the cobordism maps of Proposition 3.1(a).*

Remark 3.5. The data of $(W,c)_* : I(Y) \to I(Y')$ consists of induced maps $(W,c)_* : F_r I_d(Y) \to F_{r+L} I_{d+D}(Y')$ for all $r \in \mathbb{R} \cup \{\infty\}$ and $d \in \mathbb{Z}$ compatible with the connecting homomorphisms i and periodicity maps φ. In particular, we use the same notation $(W,c)_*$ for the induced map $I_d(Y) \to I_{d+D}(Y')$ on Floer's instanton homology.

Proof. When $c = \varnothing$, the final claim is established for a more complicated chain complex in [Dae20, Proposition 2.15, Lemma 2.35] and similarly [NST24, Lemma 2.10-2.11]. These maps arise by counting irreducible connections of index zero on the trivial $SU(2)$-bundle over W satisfying an appropriately perturbed ASD equation. While the latter reference only defines these maps for r outside the critical set of the two Chern–Simons functions, again we may extend them to all r in a right-continuous fashion.

The cobordism map for nonempty c is given by the same construction, counting connections on the $U(2)$-bundle E_c over W with $c_1(E_c)$ Poincaré dual to c, for which the traceless part of the curvature $F_0(A)$ satisfies a perturbed version of the ASD equation $F_0(A)^+ = 0$. Here we count connections whose induced connection on the determinant bundle is fixed, modulo determinant-1 gauge transformations. The proof that this map is well-defined goes through with no change, but we will discuss the computation of its degree and level.

Given critical values α, α' of the perturbed Chern–Simons functional of Y, Y', we can form the moduli space of connections on E_c satisfying the perturbed ASD equation and asymptotic to α and α' on the ends. This moduli space has connected components with different expected dimensions. Furthermore, the expected dimension of the connected component containing a connection A is uniquely determined by the *topological energy* of A defined as

$$\mathcal{E}(A) = \frac{1}{8\pi^2} \int \mathrm{tr}(F_0(A)^2). \tag{4}$$

We use the following convention to fix a subspace $M(W, c; \alpha, \alpha')$ of the moduli space of instantons on E_c with a fixed expected dimension.

The bundle E_c has a splitting as $L_c \oplus \underline{\mathbb{C}}$ where L_c is a line bundle whose c_1 is Poincaré dual to c. If A_c is a connection on L_c which is asymptotic to the trivial connection on the ends, then it induces a reducible connection on E_c compatible with the splitting $L_c \oplus \underline{\mathbb{C}}$, which we still denote it by A_c. The quantity $\mathcal{E}(A_c)$ is equal to $-c^2/4$ by Chern–Weil theory. Then the moduli space $M(W, c; \alpha, \alpha')$ consists of gauge equivalence classes of instantons A on E_c that are asymptotic to α, α' on the ends, and their topological energy satisfies

$$\begin{aligned} \mathcal{E}(A) &= \mathcal{E}(A_c) + \widetilde{\mathrm{CS}}(\alpha) - \widetilde{\mathrm{CS}}(\alpha') \\ &= -c^2/4 + \widetilde{\mathrm{CS}}(\alpha) - \widetilde{\mathrm{CS}}(\alpha'). \end{aligned} \tag{5}$$

Now we turn to the definition of $(W, c)_*$. The coefficient of α' in $(W, c)_*(\alpha)$ is given by counting the number of points in $M(W, c; \alpha, \alpha')$ when this space has expected dimension zero. Observe that by additivity of the ASD index [Don02, Equation (3.2), Proposition 3.10], we have

$$\begin{aligned} i(W, c; \alpha, \alpha') &= i(\alpha) + 3 + i(W, c; \theta, \theta') + 3 + i(Y'; \theta', \alpha') \\ &= i(\alpha) + 3 + i(W, c; \theta, \theta') - i(\alpha'). \end{aligned} \tag{6}$$

Here $i(W, c; \alpha, \alpha')$ denotes the index of the ASD operator associated to connections on W that are asymptotic to α, α'. The other terms above are defined in a similar way. Using the index formula for the ASD operators, we have $i(W, c; \theta, \theta') = -2c^2 - 3$. Using this and (**??**), $i(W, c; \alpha, \alpha') = 0$ if and only if $i(\alpha') = i(\alpha) - 2c^2$, giving the degree.

To determine the level, we argue similarly. Following [Dae20, Definition 3.46], let $\eta(W, c)$ denote the least topological energy of any ASD connection on (W, E_c) with irreducible flat limits. Since the topological energy of any instanton is non-negative, the quantity $\eta(W, c)$ is also non-negative. Furthermore, if $\eta(W, c)$ is zero, then E_c admits some projectively flat $U(2)$-connection with irreducible limits, and therefore its adjoint bundle admits a flat connection with irreducible limits. It follows that there is a homomorphism $\pi_1(W) \to SO(3)$ which restricts to a non-trivial homomorphism on both ends. Conversely, if $\pi_1(W) = 0$, then $\eta(W, c) > 0$. In the definition of $\langle (W, c)_*\alpha, \alpha' \rangle$, we count instantons with irreducible flat limits α and α'. Therefore, (5) implies that

$$\widetilde{\mathrm{CS}}(\alpha') = \widetilde{\mathrm{CS}}(\alpha) - c^2/4 - \mathcal{E}(A) \leq \widetilde{\mathrm{CS}}(\alpha) - \frac{1}{4}c^2 - \eta(W, c).$$

This gives the claim about the level of $(W, c)_*$. □

Proposition 3.2 allows us to define numerical invariants of integer homology spheres, while Proposition 3.4 — together with Lemma 2.7 — allows us to analyze their behavior under certain cobordisms. We will now focus on $R = \mathbb{F}_2$.

Definition 3.6. Suppose Y is an integer homology sphere. We define $\kappa_Y : \mathbb{Z} \to \mathbb{R} \cup \{\infty\}$ and $\ell(Y) \in \mathbb{R} \cup \{\infty\}$ as

$$\kappa_Y(d) = \kappa_{I(Y;\mathbb{F}_2)}(d), \quad \ell(Y) = \ell(I(Y; \mathbb{F}_2)),$$

the numerical invariants of the IP-module associated with Y.

These quantities are invariant under orientation-preserving diffeomorphisms of integer homology spheres. Notice that Lemma 2.5(c) immediately implies that $\ell(Y) \in \mathbb{R}$ if and only if $I_d(Y; \mathbb{F}_2) \neq 0$ for some $d \in \mathbb{Z}/8$.

3.1. **Exact triangles.** We state the two exact triangles relevant to this paper. The first is now classical, and due to Floer; the second originates, in the context of admissible bundles, in [CDX20]. In this section, Y is an integer homology sphere, $K \subset Y$ is a knot, and $Y_r(K)$ is the manifold obtained by r-surgery on K.

The following is [Flo90, Theorem 2.4]. Our index convention compares to Floer's as $i(\alpha) = -3 - i_{\mathrm{Floer}}(\alpha)$. Here, we collapse the $\mathbb{Z}/8$-grading on $I_*(Y)$ to a $\mathbb{Z}/4$-grading, so for instance the degree 1 mod 4 part of $I_*(Y)$ is $I_1(Y) \oplus I_5(Y)$.

Proposition 3.7. *There is an exact triangle of $\mathbb{Z}/4$-graded modules*

$$\begin{tikzcd} I_*(Y) \arrow[rr, "W_*"] & & I_*(Y_{-1}(K)) \arrow[dl] \\ & I^w_*(Y_0(K)) \arrow[ul, "{[-3]}"] & \end{tikzcd}$$

The horizontal homomorphism is the cobordism map associated to the 2-handle cobordism, and the leftmost homomorphism has degree -3.

In particular, by Proposition 3.4, the homomorphism $W_* : I_*(Y) \to I_*(Y_{-1})$ naturally extends to an IP-morphism $I(Y) \to I(Y_{-1})$ of degree and level zero.

We will establish the following distance-two surgery exact triangle only over $R = \mathbb{F}_2$, to avoid checking tedious details with signs. The result is expected to hold with coefficients in any commutative ring.

Proposition 3.8. *There is an exact triangle of $\mathbb{Z}/8$-graded $\mathbb{F}_2$-vector spaces*

$$\begin{tikzcd} I_{*+2}(Y_{-1}(K);\mathbb{F}_2) \arrow[rr, "{[-1]}"] & & I_*(Y_1(K);\mathbb{F}_2) \arrow[dl, "W_*\oplus(W,c)_*"] \\ & I_*(Y;\mathbb{F}_2) \oplus I_{*+2}(Y;\mathbb{F}_2) \arrow[ul, "(W',c')_*\oplus W'_*"] & \end{tikzcd}$$

The diagonal homomorphisms are induced by direct sums of cobordism maps, where W and W' are the 2-handle cobordisms with c, c' the cocore and core respectively, capped off with a Seifert surface in Y to give a closed surface.

Here, because of the degree shift in the domain, the statement that the top homomorphism $I_{*+2}(Y_{-1}) \to I_*(Y_1)$ has degree -1 means that $I_d(Y_{-1})$ maps to $I_{d-3}(Y_1)$.

Now Proposition 3.4 states that the rightmost and leftmost homomorphisms have natural extensions to a *direct sum* of IP-morphisms. The morphisms W_* and W'_* have degree zero and nonpositive level, while $(W,c)_*$ and $(W',c')_*$ have degree 2 and level at most 1/4. It will be relevant later that $L - D/8 \le 0$ in both cases, with strict inequality if $\pi_1(Y \setminus K)$ is normally generated by the meridian μ_K. The proof of Proposition 3.8 will be given in Section 5.4.

When $Y = S^3$, the top map is an isomorphism. The proof of the exact triangle in this particular case gives an explicit description of its inverse, a map $g_1 : I_*(Y_1(K);\mathbb{F}_2) \to I_{*-1}(Y_{-1}(K);\mathbb{F}_2)$. This map is obtained by counting instantons on the composite cobordism with respect to a certain *one-parameter family* of metrics on the composite. As a result, it gives rise to an IP-morphism. We record its properties in the following proposition, whose proof is given in Section 5.5.

Proposition 3.9. *Let K be a knot in S^3. Then there is an IP-morphism $g_1 : I(S^3_1(K);\mathbb{F}_2) \to I(S^3_{-1}(K);\mathbb{F}_2)$ of degree 3 and level $\frac{1}{4} - \eta(K)$, where $\eta(K) \ge 0$ is strictly positive when K is not the unknot. The induced map on Floer's instanton homology $I_d(S^3_1(K);\mathbb{F}_2) \to I_{d-1}(S^3_{-1}(K);\mathbb{F}_2)$ is an isomorphism.*

4. Proof of the main theorems

Here we use the results stated in the previous section to prove our main results.

4.1. **Proof of Theorem 1.8.** Theorem 1.8 follows from the following more precise claim. To set notation, let L be a knot in $S^2 \times S^1$ which generates H_1 and fix a choice of framing curve λ. Let $S_n(L)$ be the three-manifold obtained by surgery on L with framing $n\mu + \lambda$. Note that $S_n(L)$ is a homology sphere for every n. We write $K_n \subset S_n(L)$ for the dual knot.

Theorem 4.1. *In the situation above, if L is not isotopic to $\{*\} \times S^1$, we have*

$$\infty > \cdots > \ell(S_{n-1}(L)) > \ell(S_n(L)) > \ell(S_{n+1}(L)) > \cdots$$

Proof. Our assumption on L implies that $S_n(L) \setminus N(K_n) \cong S^2 \times S^1 \setminus N(L)$ is an irreducible 3-manifold whose boundary is incompressible. By [LPCZ23, Theorem 1.3], each $S_n(L)$ has non-trivial instanton homology, hence $\ell(S_n(L)) < \infty$ for all n.

The 2-handle cobordism $W(n) : S_n(L) \to S_{n+1}(L)$ has $b^+ = 0$ and trivial π_1. To see the latter claim, note that the Seifert–Van Kampen theorem implies that $\pi_1(W(n))$ is isomorphic to the quotient of $\pi_1(S^2 \times S^1 \setminus N(L))$ by the normalizer of $\pi_1(\partial N(L))$. This is equivalent to the quotient of $\pi_1(S^2 \times S^1)$ by the normal subgroup generated by the class of L, and hence it is trivial. Proposition 3.4 gives an IP-morphism $W(n)_* : I(S^3_n(L)) \to I(S^3_{n+1}(L))$ with degree 0 and level $-\eta\big(W(n)\big) < 0$. Applying Proposition 3.7 to the pair $(S_n(L), K_n)$ gives an exact triangle

$$\begin{tikzcd}
I_*(S_n(L)) \arrow[rr, "W(n)_*"] & & I_*(S_{n+1}(L)) \arrow[dl] \\
& I^w_*(S^2 \times S^1) \arrow[ul, "{[-3]}"] &
\end{tikzcd}$$

By Proposition 3.1(c), $I^w_*(S^2 \times S^1) = 0$, so the induced map $W(n)_* : I_*(S_n(L)) \to I_*(S_{n+1}(L))$ is an isomorphism. By Lemma 2.7(b), we see that

$$\ell(S_{n+1}(L)) \leq \ell(S_n(L)) - \eta(W(n)).$$

Because $\ell(S_n(L))$ is finite, we in fact have a strict inequality $\ell(S_n(L)) > \ell(S_{n+1}(L))$ for all n. □

Proof of Theorem 1.8. Theorem 4.1 implies that for any pair of distinct integers n and m, the 3-manifolds $S^3_n(L)$ and $S^3_m(L)$ are not orientation-preserving diffeomorphic. Next, let $X : S^3_n(L) \to S^3_m(L)$ be a ribbon homology cobordism, and let $\overline{X} : S^3_m(L) \to S^3_n(L)$ be the reverse of X (the orientation-reversal $-X$ considered as a cobordism in the other direction). Proposition 3.4 implies that

$$X_* : I(S^3_n(L)) \to I(S^3_m(L)), \qquad \overline{X}_* : I(S^3_m(L)) \to I(S^3_n(L))$$

are IP-morphisms of degree 0 and level 0. Furthermore, [DLVVW22, Theorem 4.1] asserts that $X_* : I_*(S_n(L)) \to I_*(S_m(L))$ and $\overline{X}_* : I_*(S_m(L)) \to I_*(S_n(L))$ are respectively injective and surjective homomorphisms of vector spaces for any d. Since it is shown in the proof of Theorem 4.1 that $I_*(S_n(L))$ and $I_*(S_m(L))$ are isomorphic, the maps X_* and $\overline{X}_*$ are also isomorphisms on Floer homology, which extend to IP-morphisms of degree 0 and level 0. By another application of Lemma 2.7(b), we conclude that $\ell(S_n(L)) = \ell(S_m(L))$, which is a contradiction. □

4.2. **Proof of Theorem 1.3.** The proof of Theorem 1.3 is similar, but more intricate, and requires some more initial input. Henceforth, K denotes a knot in S^3. We begin with a computation of instanton Floer groups for surgeries on a knot.

The following statement is restricted to the case $R = \mathbb{F}_2$ because Proposition 3.8 is. An interested reader willing to lift the proof of Proposition 3.8 to the integers would be able to generalize the following statement to allow coefficients in an arbitrary commutative ring.

Lemma 4.2. *For any integer $n > 0$ we have an isomorphism of $\mathbb{Z}/8$-graded $\mathbb{F}_2$-vector spaces*

$$I_*\big(S^3_{-1/n}(K);\mathbb{F}_2\big) \cong \bigoplus_{i=0}^{n-1} I_{*-2i}\big(S^3_{-1}(K);\mathbb{F}_2\big), \qquad I_*\big(S^3_{1/n}(K);\mathbb{F}_2\big) \cong \bigoplus_{i=0}^{n-1} I_{*+2i}\big(S^3_{1}(K);\mathbb{F}_2\big).$$

Furthermore, we have isomorphisms of $\mathbb{Z}/4$-graded $\mathbb{F}_2$-vector spaces

$$I_*(S^3_{-1}(K);\mathbb{F}_2) \cong I^w_*(S^3_0(K);\mathbb{F}_2) \cong I_{*-3}(S^3_1(K);\mathbb{F}_2).$$

Proof. The coefficient ring $R = \mathbb{F}_2$ will be suppressed from notation for the rest of the argument. We will prove the first statement for $I_*(S^3_{1/n}(K))$ for $n > 0$; the argument for $I_*(S^3_{-1/n}(K))$ can be proved similarly. To simplify notation, for this proof we write $S_{1/n} = S^3_{1/n}(K)$.

We will prove the following stronger claim by induction on $n \geq 1$:

- For each n there is an isomorphism
$$\varphi_n : I_*(S_{1/n}) \to \bigoplus_{i=0}^{n-1} I_{*+2i}(S_1).$$
- These isomorphisms can be chosen so that for all $n \geq 2$, the 2-handle cobordism map $I_*(S_{1/n}) \to I_*(S_{1/(n-1)})$ is identified with the projection to the first $n-1$ coordinates (and in particular is surjective).[2]

The base case $n = 1$ is tautological; one may take φ_1 to be the identity.

Suppose the claim is proved for $n \geq 2$. Apply Proposition 3.8 to the pair $(Y, K) = (S_{1/n}, \tilde{K})$. By inductive hypothesis the 2-handle cobordism map $W_* : I_*(S_{1/n}) \to I_*(S_{1/(n-1)})$ is known to be surjective, so our exact triangle is in fact a short exact sequence

$$0 \to I_*(S_{1/(n+1)}) \xrightarrow{W_* \oplus (W,c)_*} I_*(S_{1/n}) \oplus I_{*+2}(S_{1/n}) \xrightarrow{(W',c')_* \oplus W'_*} I_{*+2}(S_{1/(n-1)}) \to 0.$$

Applying the inductive hypothesis, this sequence is isomorphic to the short exact sequence

$$0 \to I_*(S_{1/(n+1)}) \xrightarrow{\varphi_n W_* \oplus \varphi_n (W,c)_*} \bigoplus_{i=0}^{n-1} I_{*+2i}(S_1) \bigoplus_{j=1}^{n} I_{*+2j}(S_1) \xrightarrow{f \oplus \pi} \bigoplus_{j=1}^{n-1} I_{*+2j}(S_1) \to 0,$$

where π is projection onto the first $n-1$ coordinates.

Thus by exactness the map $\varphi_n W_* \oplus \varphi_n(W,c)_*$ induces an isomorphism

$$\begin{aligned} I_*(S_{1/(n+1)}) &\cong \{(x,y,z) \in \bigoplus_{i=0}^{n-1} I_{*+2i}(S_1) \bigoplus_{j=1}^{n-1} I_{*+2j}(S_1) \oplus I_{*+2n}(S_1) \mid f(x) = y\} \\ &\cong \bigoplus_{i=0}^{n-1} I_{*+2i}(S_1) \oplus I_{*+2n}(S_1), \end{aligned}$$

where the final map sends (x, y, z) to (x, z). The isomorphism φ_{n+1} is the composite of these two identifications. Finally, with respect to these isomorphisms the map $I_*(S_{1/(n+1)}) \to I_*(S_{1/n})$ is

[2] The corresponding statement for $-1/n$ is that the 2-handle cobordism map $W_* : I_*(S_{-1/(n-1)}) \to I_*(S_{-1/n})$ is identified with the inclusion of the first $n-1$ coordinates, and in particular is injective.

given by sending (x, z) to x, completing the induction.

The second claim is Proposition 3.7 applied to the pairs $(Y, K) = (S^3, K)$ and $(S^3_1(K), \tilde{K})$. $\square$

Corollary 4.3. *If K is any non-trivial knot in S^3, the vector space $I_*(S_{1/n}(K); \mathbb{F}_2)$ is non-trivial for all integers $n \neq 0$.*

Proof. By the universal coefficient theorem, it suffices to show $I_*(S_{1/n}(K); \mathbb{C}) \neq 0$. By the preceding lemma, it is equivalent to show $I^w_*(S_0(K); \mathbb{C}) \neq 0$. Finally, because K is non-trivial, [KM10, Proposition 7.16] gives the nonvanishing of a certain summand $KHI(K, g) \subset I^w_*(S_0(K); \mathbb{C})$. $\square$

Finally, Theorem 1.3 follows immediately from the following more precise claim.

Theorem 4.4. *If K is a non-trivial knot in S^3, we have*

$$\infty > \cdots > \ell\big(S^3_{1/2}(K)\big) > \ell\big(S^3_1(K)\big) > \ell\big(S^3_{-1}(K)\big) > \ell\big(S^3_{-1/2}(K)\big) > \cdots$$

In fact, we have $\ell\big(S^3_{1/n}(K)\big) - \ell\big(S^3_{-1/n}(K)\big) > 1/8$ for all $n > 0$.

Proof. That these ℓ-invariants are all finite follows from Corollary 4.3 and Lemma 2.5(c).

It was established in the proof of Lemma 4.2 that for $n > 0$, the cobordism $W_n : S_{-1/n}(K) \to S_{-1/(n+1)}(K)$ induces an injection on Floer homology. Thus, Floer's exact triangle collapses to a short exact sequence

$$0 \to I_*(S_{-1/n}(K)) \to I_*(S_{-1/(n+1)}(K)) \to I^w_*(S_0(K)) \to 0,$$

where the first map $(W_n)_*$ is the cobordism map induced by the simply-connected negative-definite cobordism $W_n : S_{-1/n}(K) \to S_{-1/(n+1)}(K)$ given by attaching a (-1)-framed handle along $\tilde{K}$. Because $(W_n)_*$ is injective, and by Proposition 3.4 extends to an IP-morphism

$$I\big(S_{-1/n}(K)\big) \to I\big(S_{-1/(n+1)}(K)\big)$$

of degree 0 and level $-\eta(W_n) < 0$, it follows from Lemma 2.7(b) that $\ell(S_{-1/n}(K)) > \ell(S_{-1/(n+1)}(K))$ for all $n > 0$.

Next, Proposition 3.9 and Lemma 2.7(b) immediately combine to give

$$\ell\big(S^3_1(K)\big) \geq \ell\big(S^3_{-1}(K)\big) - \big(-\frac{1}{4} - \eta(K) + 1/8\big) = \ell\big(S^3_{-1}(K)\big) + \frac{1}{8} + \eta(K) > \ell\big(S^3_{-1}(K)\big) + \frac{1}{8}.$$

Finally, we use that the triangle of Proposition 3.8 also collapses into a short exact sequence

$$0 \to I_*(S^3_{1/(n+1)}(K)) \to I_*(S^3_{1/n}(K)) \oplus I_{*+2}(S^3_{1/n}(K)) \to I_{*+2}(S^3_{1/(n-1)}(K)) \to 0.$$

In particular, the first map is injective. This map is the direct sum $W_* \oplus (W, c)_*$ of cobordism maps, where W is simply-connected and negative-definite. Proposition 3.4 implies that this enriches to a direct sum of morphisms of I-modules, the first of which has $L_1 - D_1/8 = -\eta(W) < 0$, the latter of which has $L_2 - D_2/8 = -\eta(W, c) < 0$. In particular, the larger of the two is still negative. Lemma 2.9 immediately gives

$$\ell\big(S^3_{1/(n+1)}(K)\big) > \ell\big(S^3_{1/n}(K)\big),$$

completing the proof. $\square$

Proof of Theorem 1.3. By Lemma 4.2, we see that $I_*(S^3_{1/n}(K); \mathbb{F}_2)$ and $I_*(S^3_{-1/n}(K); \mathbb{F}_2)$ have the same rank. As in the proof of Theorem 1.8, we see that if $X : S^3_{1/n}(K) \to S^3_{-1/n}(K)$ is a ribbon homology cobordism and $\overline{X}$ is its reverse, then the induced maps $X_*, \overline{X}_*$ are both isomorphisms. From this it follows that $\ell(S^3_{1/n}(K)) = \ell(S^3_{-1/n}(K))$, a contradiction. $\square$

5. The distance-two surgery triangle

The goal of this section is to construct the exact triangle in Proposition 3.8. The existence of such an exact triangle is proposed in [CDX20], and our proof here follows closely the proof of [CDX20, Theorem 1.6], which concerns the analogue of Proposition 3.8 for instanton homology of admissible bundles with respect to the more general gauge group $SU(N)$. In particular, the proof of the $N = 2$ case of [CDX20, Theorem 1.6] provides the skeleton of the proof of Proposition 3.8, except that we also need to analyze the reducible ASD connections over various cobordisms to guarantee that they do not cause any issue in the construction of the distance-two surgery triangle. Our proof of Proposition 3.8 is also formally similar to the proof of Floer's surgery exact triangle as given in [Sca15], though the argument is complicated by the presence of the more complicated 'middle ends' discussed in Section 5.2.

In the first subsection below, we review a standard homological algebra lemma about exact triangles. In the next subsection, we review the definition of various cobordisms with families of Riemannian metrics relevant in the proof of Proposition 3.8. Then we study the reducible ASD connections with respect to these families of metrics. The proofs of Proposition 3.8 and Proposition 3.9 are given in the final two subsections.

To avoid dealing with the study of orientations of moduli spaces, we work with $\mathbb{F}_2$ coefficients in this section. In particular, all chain complexes in this section are defined over $\mathbb{F}_2$.

5.1. **Homological algebra of surgery exact triangles.** As with many exact triangles in Floer theory, the proof of Proposition 3.8 is given by the *triangle detection lemma* (see [Sei08, Lemma 3.7] and [OS05, Lemma 4.2]).

Proposition 5.1. *For each $i \in \mathbb{Z}$, let (C_i, d_i) be a chain complex. Suppose that for all $i \in \mathbb{Z}$ we are given maps*

$$f_i : C_i \to C_{i-1} \qquad g_i : C_i \to C_{i-2} \qquad h_i : C_i \to C_{i-3},$$

which satisfy the following properties:

$$\begin{aligned} d_i^2 &= 0 \\ d_{i-1}f_i + f_i d_i &= 0 \\ d_{i-2}g_i + f_{i-1}f_i + g_i d_i &= 0 \\ d_{i-3}h_i + f_{i-2}g_i + g_{i-1}f_i + h_i d_i &= q_i \end{aligned}$$

where $q_i : C_i \to C_{i-3}$ is an isomorphism. Then the map

$$C_{i+1} \xrightarrow{(f_i, g_i)} \mathrm{Cone}(f_{i-1}) \overset{\text{def}}{=} \left(C_{i-1} \oplus C_{i-2}[1], \begin{bmatrix} d_{i-1} & 0 \\ f_{i-1} & d_{i-2} \end{bmatrix} \right)$$

is a chain homotopy equivalence. In particular, if H_i denotes the homology of (C_i, d_i), then we have the following exact triangle:

$$\begin{tikzcd} H_{-1} \arrow[rr, "(f_{-1})_*"] & & H_1 \arrow[dl, "(f_1)_*"] \\ & H_0 \arrow[ul, "(f_0)_*"] & \end{tikzcd}$$

To prove Proposition 3.8, we apply this lemma to the case that

$$C_{-1} := CI_*(Y_{-1}(K)), \qquad C_0 := CI_*(Y) \oplus CI_{*-2}(Y), \qquad C_1 := CI_*(Y_1(K)).$$

More generally, C_k is defined by requiring that $C_k = C_{k+3}$ for any k. The homomorphism d_i in each case is given by the corresponding Floer differential and the maps f_i, g_i and h_i are given by cobordism maps in instanton Floer theory. We recall the definition of these cobordisms in the next subsection.

5.2. **Cobordisms and families of metrics.** First we fix some notations for the discussion of the cobordisms involved in the proof of Proposition 3.8. A cobordim from a 3-manifold Z to another 3-manifold Z' with a middle end L is a 4-manifold W with

$$\partial W = -Z \sqcup Z' \sqcup L.$$

We write $W : Z \xrightarrow{L} Z'$ for any such cobordism, and we drop L from the notation when the choice of the middle end is clear from the context. Given two such cobordisms $W_0 : Z \xrightarrow{L} Z'$ and $W_1 : Z' \xrightarrow{L'} Z''$, we may compose them to obtain $W_0 \circ W_1 : Z \xrightarrow{L \sqcup L'} Z''$. In the following, fix a knot K in an integer homology sphere Y, and let $E(K)$ denote the exterior of K. We assume that an identification of $\partial E(K)$ with $S^1 \times S^1$ is fixed. For $-1 \le i \le 1$, let Z_i denote the result of $1/i$ surgery on the knot K in Y, and extend the definition of Z_i to any integer i by requiring that $Z_{k+3} = Z_k$ for any k.

Form a 4-manifold by gluing $[-3,-1] \times S^1 \times D^2$ and $[1,3] \times S^1 \times D^2$ to $[-3,3] \times E(K)$ respectively along $[-3,-1] \times \partial E(K)$ and $[1,3] \times \partial E(K)$ using the identifications $\mathbf{1}_{[-3,-1]} \times f_{-1}$ and $\mathbf{1}_{[1,3]} \times f_1$ where $f_{\pm 1} : \partial(S^1 \times D^2) \to \partial E(K)$ corresponds to ± 1-surgery on K. The resulting 4-manifold X has three boundary components (after smoothing the corners) which may be identified with $-Z_{-1}$, Z_1 and $\mathbb{RP}^3$. In particular, X can be regarded as a cobordism from Z_{-1} to Z_1 with the middle end $\mathbb{RP}^3$. (This is a special case of the construction described in [CDX20, Section 3.1].) A schematic diagram of this construction is presented as Figure 1 below. Using a similar construction, we construct cobordisms

$$X' : Z_1 \xrightarrow{S^3} Z_0, \qquad X'' : Z_0 \xrightarrow{S^3} Z_{-1}.$$

Filling the S^3 boundary components of the latter two cobordisms with 4-balls give rise to the standard 2-handle cobordisms. For any integer i, let $W^i_{i-1} : Z_i \xrightarrow{L_i} Z_{i-1}$ be given by either of X, X' or X'' where L_i is either $\mathbb{RP}^3$ or S^3. More generally, we define $W^i_j : Z_i \xrightarrow{L} Z_j$ for any $j \le i$ as the iterated composite

$$W^i_j := W^{j+1}_j \circ \cdots \circ W^i_{i-1}$$

with $L = L_i \sqcup \cdots \sqcup L_{j+1}$.

An explicit understanding of the cohomology groups of these cobordisms will be useful later when studying their reducible instantons. The 4-manifolds X, X' and X'' all have trivial first cohomology, and their second cohomology groups are isomorphic to $\mathbb{Z}$. In the description of X above, the cylinders

$$\mathfrak{c}_- := [-3,-1] \times S^1 \times \{0\} \subset [-3,-1] \times S^1 \times D^2, \quad \mathfrak{c}_+ := [1,3] \times S^1 \times \{0\} \subset [1,3] \times S^1 \times D^2$$

determine relative homology classes for $(X, \partial X)$, and the Poincaré dual of each of these determines a generator for the second cohomology group of X. The intersection of $\mathfrak{c}_\pm$ with $Z_{\pm 1}$ is the dual knot of the Dehn surgery, and we can glue a Seifert surface of this knot to $\mathfrak{c}_\pm$ to obtain a properly embedded surface $c_\pm$ representing the same cohomology class as $\mathfrak{c}_\pm$ in X. Note that the only boundary component of $c_\pm$ is in $\mathbb{RP}^3$ and the restriction of the cohomology class of $c_\pm$ to this middle end is the generator of $H^2(\mathbb{RP}^3)$. We may similarly form embedded surfaces $c'_\pm \subset X'$ and $c''_\pm \subset X''$ which represent generators of the second cohomologies of X' and X''. The intersection form of X is positive-definite and the intersection forms of X' and X'' are negative definite. Because each of X, X', X'' has boundary a union of rational homology spheres, a second cohomology class has a well-defined square $c^2 \in \mathbb{Q}$, defined using the isomorphism $H^2(X, \partial X; \mathbb{Q}) \to H^2(X; \mathbb{Q})$. For the cohomology classes above, we have $c_\pm^2 = 1/2$, while $(c'_\pm)^2 = (c''_\pm)^2 = -1$.

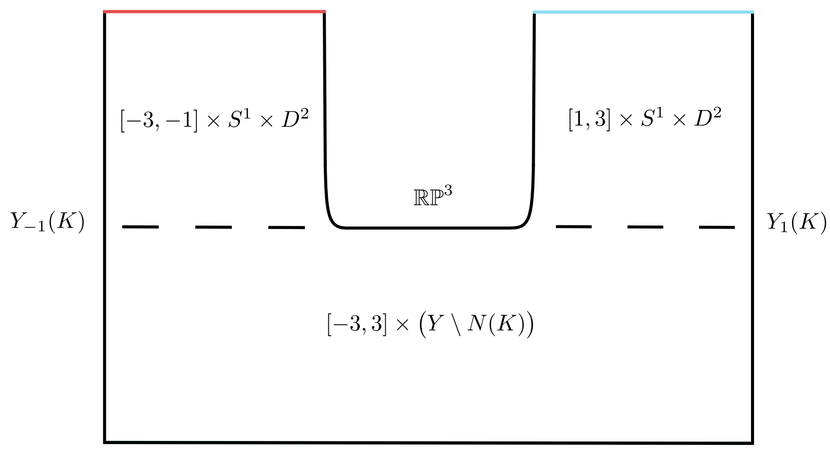

FIGURE 1. A representation of the cobordism $X = W^{-1}_{-2}$. The three boundary components are labeled by their oriented diffeomorphism types; the embedded surface $\mathfrak{c}_- = [-3,-1] \times S^1 \times \{0\}$ is represented as the darker red curve, while $\mathfrak{c}_+$ is represented as the lighter blue curve. We visualize c_+ and c_- by the same picture, with the understanding that the intersections with integer homology spheres are capped off by Seifert surfaces.

We define a pair of embedded surfaces $\hat{c}^i_j, \check{c}^i_j \subset W^i_j$ by the formulas

$$\begin{aligned} &\hat{c}^1_0 = \varnothing \subset W^1_0 && \hat{c}^0_{-1} = c_+ \subset W^0_{-1}, && \hat{c}^{-1}_{-2} = c_- \subset W^{-1}_{-2}, \\ &\check{c}^1_0 = c_+ \subset W^1_0, && \check{c}^0_{-1} = c_- \sqcup c_+ \subset W^0_{-1}, && \check{c}^{-1}_{-2} = c_- \subset W^{-1}_{-2} \end{aligned}$$

in the case that $j = i-1$. For general $i > j$, we define $\hat{c}^i_j = \hat{c}^i_{i-1} \circ \cdots \circ \hat{c}^{j+1}_j$, and similarly for $\check{c}^i_j$. (This is a special case of the more general construction of [CDX20, Section 3.2]).

The cobordism W^i_j admits a family of metrics G^i_j parametrized by the *associahedron* of dimension $i-j-1$. We need these families of metrics with $i-j \leq 3$ to define the homomorphisms used in the proof of Proposition 3.8. So, we only focus on recalling the definition of these families of metrics within this range. First fix Riemannain metrics on the 3-manifolds Z_1, Z_0 and Z_{-1}. (This choice is already assumed in the definition of the complexes C_i in the previous subsection.) Fix Riemannian metrics G^i_{i-1} on the cobordisms $W^1_0, W^0_{-1}, W^{-1}_{-2}$ with cylindrical ends modeled on the chosen metrics on Z_1, Z_0 and Z_{-1} and the round metrics on S^3 and $\mathbb{RP}^3$. This construction may be extended to arbitrary i by requiring that $G^{i+3}_{i+2} = G^i_{i-1}$.

The family of metrics G^i_{i-2} on W^i_{i-2} is parametrized by the interval $[-1,1]$. Each end of the interval corresponds to a Riemannian metric which is broken along a separating codimension-1 submanifold; for a detailed reference discussing such families including broken metrics, see [KMOS07, Section 5]. Since W^i_{i-2} is the composition of the cobordisms W^i_{i-1} and W^{i-1}_{i-2}, the chosen metrics on these cobordisms determine a metric on W^i_{i-2}, which is broken along an embedded copy of Z_{i-1}. This metric is the element of the family of metrics G^i_{i-2} associated to the endpoint 1 of the interval $[-1,1]$. For $1/3 \leq t < 1$, let $T = 1/(1-t)$. Then the metric corresponding to t in the family of metrics is given by removing the subspaces $[T,\infty) \times Z_{i-1}$ from W^i_{i-1} and $(-\infty,-T] \times Z_{i-1}$ from W^{i-1}_{i-2} and then gluing the boundary components Z_{i-1} together.

The metric corresponding to the other endpoint is given by a broken metric fully stretched along an embedded copy of S^3 or $\mathbb{RP}^3$ in W^i_{i-2}. First we consider the cobordism $W^1_{-1} = X' \circ X''$. This cobordism is given by gluing

$$[-5,-3] \times S^1 \times D^2, \qquad [-1,1] \times S^1 \times D^2, \qquad [3,5] \times S^1 \times D^2$$

to $[-5,5] \times E(K)$ respectively using the gluing maps $\mathbf{1}_{[-5,-3]} \times f_1$, $\mathbf{1}_{[-1,1]} \times f_0$ and $\mathbf{1}_{[3,5]} \times f_{-1}$. (Analogous to $f_{\pm 1}$, the map f_0 is a diffeomorphism $\partial(S^1 \times D^2) \to \partial E(K)$, which in this case corresponds to ∞-surgery.) Then the union of following subspaces of W^1_{-1}

$$\{-4\} \times S^1 \times D^2 \subset [-5,-3] \times S^1 \times D^2, \qquad \{4\} \times S^1 \times D^2 \subset [3,5] \times S^1 \times D^2, \tag{7}$$

$$\gamma \times S^1 \times S^1 \subset [-5,5] \times N(\partial E(K)) \tag{8}$$

gives a submanifold M^1_{-1} of W^1_{-1} diffeomorphic to $\mathbb{RP}^3$. Here $N(\partial E(K))$ is a regular neighborhood of the boundary $\partial E(K)$ in $E(K)$, and in particular, it can be identified with $(-1,0] \times S^1 \times S^1$. This gives an identification of $[-5,5] \times N(\partial E(K))$ with $[-5,5] \times (-1,0] \times S^1 \times S^1$. In (7), γ is a properly embedded path in $[-5,5] \times (-1,0]$ whose endpoints are $(-4,0)$ and $(4,0)$.

The submanifold M^1_{-1} is separating and the two connected components of the complement of a tubular neighborhood of M^1_{-1} can be described as follows. One of the components can be regarded as a cobordism from Z_1 to Z_{-1} with the middle end $\mathbb{RP}^3$, and is given by reversing the cobordism W^{-1}_{-2}. The other connected component is a 4-manifold N with three boundary components, one of which is $\mathbb{RP}^3$ and the other two are S^3. Let D be a twice punctured 3-ball with three boundary components S_{-1}, S_0 and S_1 which are diffeomorphic to the 2-sphere. Then N is diffeomorphic to the total space of the S^1-bundle over D whose Euler class evaluates to -1 on the boundary components $S_{\pm 1}$ of D and to 2 on the boundary component S_0. In particular, the boundary component $\mathbb{RP}^3$ of N corresponds to the circle bundle over S_0. The manifold N contains the embedded surface $c = \hat{c}^1_{-1} \cap N$, and the Poincaré dual cohomology class has $c^2 = -1/2$. (This discussion is a special case of the discussion of 'spherical cuts' and their complements in [CDX20, Section 3.1].)

The metric corresponding to -1 in the family of metrics G^i_{i-2} is given by a broken metric on W^1_{-1}, which is broken along M^1_{-1}. To be more precise, we fix metrics with cylindrical ends on the complement of M^1_{-1}, which has two cylindrical ends corresponding to the round metrics on $\mathbb{RP}^3$, two cylindrical ends corresponding to the round metric on S^3 and one cylindrical end for each of $Z_{\pm 1}$. Similar to the previous case, we define our family of metrics for the interval by $(-1,-1/3]$ by removing half cylinders from the $\mathbb{RP}^3$ ends, and then gluing the two connected components along their $\mathbb{RP}^3$ boundaries. Finally we extend our family of metrics on W^1_{-1} to $(-1/3, 1/3)$ in an arbitrary way so that all metrics g_t coincide on the four ends of W^1_{-1} with the chosen cylindrical end metrics.

The definitions of the families of metrics on W^2_0 and W^0_{-2} follow a similar scheme. In the same way as in the previous case, we can form separating submanifolds $M^2_0 \subset W^2_0$ and $M^0_{-2} \subset W^0_{-2}$, which are in this case diffeomorphic to S^3. Then the family of metrics G^2_0 (resp. G^0_{-2}) for $t = 1$ is given by a broken metric that is fully stretched along Z_1 (resp. Z_{-1}), for $t = -1$ is given by a broken metric that is fully stretched along M^2_0 (resp. M^0_{-2}) and is extended by (non-broken metrics) for $t \in (-1,1)$. Finally we extend this construction to any W^i_{i-2} by requiring that $G^{k+3}_{k+1} = G^k_{k-2}$ for any k. We also remark that the complement of M^i_{i-2} in W^i_{i-2} has a similar description as before; one of the components is still diffeomorphic to N and the other component is given by reversing the cobordism W^{i+1}_i.

Next, we turn into the description of the family of metrics G^i_{i-3} on W^i_{i-3} parametrized by a pentagon P, which is the 2-dimensional associahedron. First we consider the case of G^2_{-1}. We start by describing five codimension-1 submanifolds of the cobordism $W^2_{-1} = W^0_{-1} \circ W^1_0 \circ W^2_1$. The definition of this cobordism as a composite implies that there are natural embeddings of Z_0 and Z_1 in W^2_{-1}. Since $W^2_{-1} = W^2_0 \circ W^0_{-1} = W^2_1 \circ W^1_{-1}$, there are also natural embeddings of M^2_0 and M^1_{-1} into W^2_{-1}. To describe the fifth submanifold M^2_{-1}, note that W^2_{-1} is given by gluing

$$[-7,-5] \times S^1 \times D^2, \qquad [-3,-1] \times S^1 \times D^2, \qquad [1,3] \times S^1 \times D^2, \qquad [5,7] \times S^1 \times D^2$$

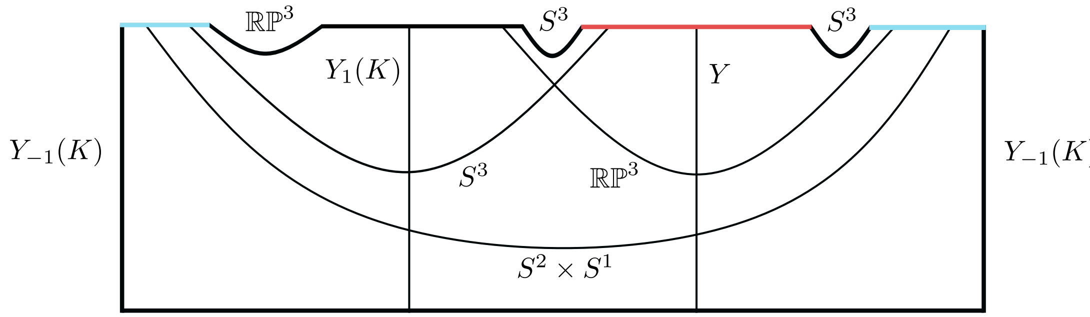

FIGURE 2. A representation of the cobordism W^2_{-1}, with its five boundary components labeled. The more thin arcs represent the five submanifolds used in the construction of the family of metrics G^2_{-1}, and each arc is labeled by the diffeomorphism type of the corrresponding submanifold. The surface $\hat{c}^2_{-1}$ is represented as the lighter blue curve; the surface $\check{c}^2_{-1}$ is the union of the darker red and lighter blue curves. The middle end $\mathbb{RP}^3$ is drawn differently than the S^3 middle ends to make the asymmetry more apparent. The pictures for other W^i_{i-3} are similar; for instance, the bundle data can on W^i_{i-3} obtained by cyclically permuting the three depicted pieces.

to $[-7,7] \times E(K)$ respectively using the gluing maps $\mathbf{1}_{[-7,-5]} \times f_{-1}$, $\mathbf{1}_{[-3,-1]} \times f_1$, $\mathbf{1}_{[1,3]} \times f_0$ and $\mathbf{1}_{[5,7]} \times f_{-1}$. Then M^2_{-1} is a submanifold of W^2_{-1} given as the union of the following three parts:

$$\{-6\} \times S^1 \times D^2 \subset [-7,-5] \times S^1 \times D^2, \qquad \{6\} \times S^1 \times D^2 \subset [5,7] \times S^1 \times D^2,$$
$$\gamma \times S^1 \times S^1 \subset [-7,7] \times N(\partial E(K))$$

where similar to (7), $\gamma \subset [-7,7]\times(-1,0]$ is a properly embedded path with the endpoints $(-6,0)$ and $(6,0)$. The manifold M^2_{-1} is diffeomorphic to $S^1 \times S^2$. The complement of a tubular neighborhood of this submanifold of W^2_{-1} has two connected components; one of the connected components, denoted V_{-1}, is given by removing from the product cobordism $[-7,7] \times Z_{-1}$ a tubular neighborhood of $\{0\} \times K_{-1}$. The other connected component, denoted by N', is given by removing from N a tubular neighborhood of one of the fibers of the S^1 fibration of N.

Label the edges of the pentagon P cyclically by e_j with $j \in \mathbb{Z}/5$, and label the common vertex of e_j and e_{j+1} with $v_{j,j+1}$. Then the edge e_j of P corresponds to the submanifold Q_j of W^2_{-1}, where

$$Q_0 = Z_0, \qquad Q_1 = Z_1, \qquad Q_2 = M^1_{-1}, \qquad Q_3 = M^2_{-1}, \qquad Q_4 = M^2_0.$$

In particular, for any j the submanifolds Q_j and Q_{j+1} are disjoint from each other. The metrics corresponding to the edge e_j of P are fully stretched along the submanifold Q_j. Furthermore, the endpoint $v_{j-1,j}$ (respectively $v_{j,j+1}$) of this edge corresponds to the metric that is also fully stretched respectively along Q_{j-1} (respectively, Q_{j+1}). As we move from $v_{j-1,j}$ to $v_{j,j+1}$ the corresponding metrics vary from having large necks along Q_{j-1} to small necks along Q_{j-1} and then from small necks along Q_{j+1} to large necks along Q_{j+1}, while all the metrics in parametrized by this edge are broken along Q_j. In particular, we require that the metric parametrized by the edge e_0 is given by the metric G^2_1 on W^2_1 and the family of metrics G^1_{-1} on W^1_{-2}. A similar requirement applies to the edge e_1. We extend the family of metrics from the boundary of P first to a tubular neighborhood of the vertices $v_{j,j+1}$ in P by replacing infinite length necks along Q_j and Q_{j+1} with finite length necks. Then we extend this family to a tubular neighborhood of e_j by replacing the infinite length necks along Q_j with finite length necks along this submanifold. Finally we extend this family to the rest of P in a smooth way by non-broken metrics on W^2_{-1}, which are assumed to all coincide with

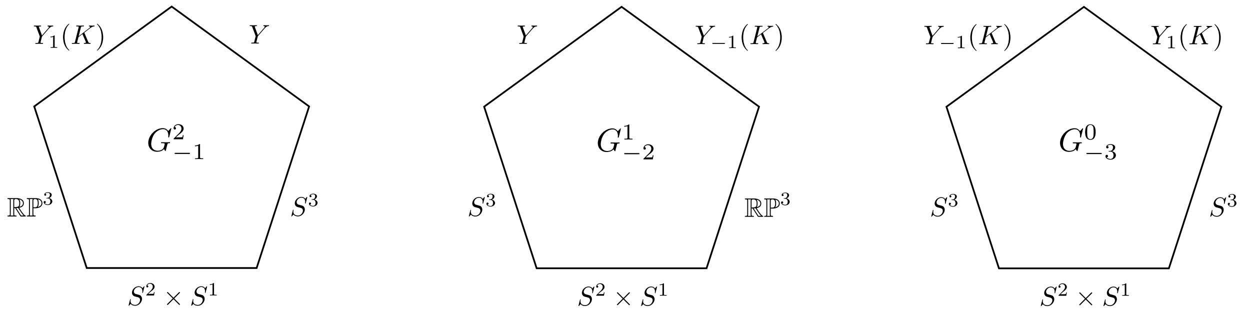

FIGURE 3. The three pentagons of metrics. Each boundary edge corresponds to an interval of metrics broken along a given submanifold, and each edge in the diagram is labeled by the diffeomorphism type of the breaking submanifold. The vertex $v_{0,1}$ is the topmost vertex, and the edges e_j proceed counter-clockwise.

the chosen cylindrical metrics on the five ends of W^2_{-1}. (This is a special case of the construction of an associahedron of metrics in [CDX20, Section 4.3].)

The definition of the families of metrics G^1_{-2} and G^0_{-3} are similar, and we extend the construction of this family of metrics on W^i_{i-3} to any i by requiring that $G^{k+3}_k = G^k_{k-3}$ for any k. In particular, as a part of the construction of the family of metrics, we from a submanifold M^i_{i-3} of W^i_{i-3} which is diffeomorphic to $S^2 \times S^1$, and the complement of its tubular neighborhood has two connected components. One of the connected components is again diffeomorphic to N' and the other connected component, denoted by V_i, is the complement of a tubular neighborhood of $\{0\} \times K_i$ in $[-7, 7] \times Z_i$.

5.3. **Reducibles.** The maps f_i, g_i, h_i will be defined by counting isolated irreducible instantons on W^i_j with respect to the family of metrics G^i_j. The relations for these maps will be computed by studying the ends of the 1-dimensional moduli spaces of irreducible instantons.

To ensure that these counts are defined, we need the 0-dimensional moduli spaces of irreducible instantons to be compact. Similarly, to get the expected boundary relations, we need some control over the reducible instantons that appear in 1-dimensional moduli spaces. There was no such issue in [CDX20], where the 3-manifolds were equipped with *admissible bundles*, which support no reducible flat connections. In this subsection, we will analyze the reducible instantons with respect to the families of metrics G^i_j. Using the study of such reducibles, we will show in the following subsection that the argument of [CDX20] goes through with minimal change, by showing that the only reducibles which appear in the compactification of the moduli spaces of interest are precisely those used in the proof of [CDX20, Theorem 1.6].

First, we review the definition of the relevant moduli spaces of instantons. Previously, we introduced properly embedded oriented surfaces $\hat{c}^i_j$ and $\check{c}^i_j$ in W^i_j, which are abbreviated to c for now. For flat connections α on Z_i and β on Z_j, there is a moduli space $M(W^i_j, c, \mathrm{int}(G^i_j); \alpha, \beta)$ of pairs (g, A), where $g \in \mathrm{int}(G^i_j)$ and A is a g-ASD connection on (W^i_j, c) which is asymptotic to α on Z_i and β on Z_j; these pairs are considered up to gauge equivalence.

To first extend this to a moduli space $M(W^i_j, c, G^i_j; \alpha, \beta)$, we include instantons with respect to the broken metrics. Suppose $g \in G^i_j$ is a broken metric, broken along a single connected submanifold Y (so $W^i_j = W \cup_Y W'$, with Y being the end intermediate to these). Then an instanton with respect to this broken metric is a pair (A, A') of an instanton on W and an instanton on W' with respect to the relevant cylindrical-end metrics, so that the A and A' are asymptotic to the same flat connection along Y, now considered up to simultaneous gauge equivalence.[3] The definition is similar for metrics

[3]A better definition in general includes *gluing factors* associated to the stabilizers Γ of the intermediate flat connections. However, these will not play any role in the examples that appear below. This is because when we glue

broken along more than one connected submanifold. We define $M(W_j^i, c, G_j^i; \alpha, \beta)$ to be the moduli space of pairs (g, A), where $g \in G_j^i$ and A is a g-instanton on (W_j^i, c) asymptotic to α on Z_i and β on Z_j, considered up to gauge equivalence; when g is broken, this is understood in the generalized sense described above. We topologize this with respect to convergence on compact subsets in the complement of the broken submanifolds.

This moduli space remains noncompact. Just as we pass from $M(W, c, \mathrm{int}(G))$ to $M(W, c, G)$ by introducing instantons with respect to these broken metrics, there exists a compactification $M^+(W, c, G; \alpha, \beta)$ which introduces *broken solutions* to the ASD equations. In the statement below, if A is an instanton on a cobordism W, the *positive limit* of A is the flat connection A is asymptotic to on its outgoing boundary component; similarly with *negative limit* and *middle limit*.

Definition 5.2. Let $(W, c) : Y \xrightarrow{L} Y'$ be a cobordism with a middle end. Suppose W is equipped with an unbroken Riemannian metric g. A **broken instanton** on (W, c) is a sequence

$$A = (B_1, \cdots, B_n, A_W, C_1, \cdots, C_m, B'_1, \cdots, B'_\ell)$$

of connections satisfying the following conditions:

(a) The connection B_i (resp. B'_i, C_i) is a nonconstant instanton on $\mathbb{R} \times Y$ (resp. $\mathbb{R} \times Y'$, $\mathbb{R} \times L$) considered modulo translation, while A_W is an instanton on (W, c).

(b) For $1 \le i < n$, the positive limit of B_i is equal to the negative limit of B_{i+1}, and similarly for C_i and B'_i.

(c) The negative limit of A_W is the positive limit of B_n, while the middle limit of A_W is the negative limit of C_1 and the positive limit of A_W is the negative limit of B'_1.

The **outer limits** of a broken instanton are given by the negative limit of B_1 and the positive limit of B'_ℓ. If every instanton in the sequence is irreducible and all limiting flat connections are irreducible, we say this broken instanton is **fully irreducible**.

A similar definition holds in the case that W is equipped with a broken metric g. We will say 'instanton' to refer to a possibly-broken instanton, and specify 'unbroken instanton' when we want to assume $n = m = \ell = 0$ (including in the case that g is a broken metric).

A broken instanton has a total of $n + m + \ell$ *intermediate limits*, the positive limits of B_i (resp. negative limits of B'_j, C_j). Each intermediate limit β gives rise to a gluing factor $\dim \Gamma_\beta$, the dimension of the stabilizer of β. When β is irreducible, this is zero; when β has abelian holonomy this dimension is one, and when β has central holonomy this dimension is three. We write $g(A)$ for the sum of all of these gluing factors.

Definition 5.3. Given a broken instanton A, its **index** is defined to be

$$i(A) = g(A) + i(A_W) + \sum_{j=1}^{n} i(B_j) + \sum_{j=1}^{m} i(L_j) + \sum_{j=1}^{\ell} i(B'_\ell),$$

where i is the index of the ASD operator.

This coincides with the index of the connection obtained by gluing the constintuent instantons of A into a connection on W. These definitions are relevant as follows. We define the space $M^+(W, c, G; \alpha, \beta)_d$ to be the space of pairs (g, A), where $g \in G$ is a (possibly broken) metric and A is (possibly broken) instanton on (W, c) with index $d - \dim G$. This is given the topology of 'chain-convergence', as in [Don02, Section 5.1]. This moduli space has expected dimension d. It is established as [Don02, Proposition 5.5] that if $M^+(W, c, G; \alpha, \beta)_{d-8n}$ is empty for all $n > 0$, the moduli space $M^+(W, c, G; \alpha, \beta)_d$ is compact.

instantons along some reducible flat connection, one will be irreducible and the other will be reducible with the same stabilizer as the intermediate flat connection.

By a generic choice of perturbation, irreducible instantons can be assumed to be cut out transversely, and thus that any fully irreducible instanton on (W, c, G) has index $i(A) \geq -\dim G$. We will define our maps by counting the number of elements in two moduli spaces $M(W^i_j, c, G^i_j; \alpha, \beta)_0$ for $c = \hat{c}^i_j$ or $c = \check{c}^i_j$. To ensure these counts are well-defined, we will need to ensure that this space coincides with its compactification; that is, that there are no broken instantons of index $i(A) \leq -\dim G$.

When verifying the boundary relations for these maps, we will investigate the moduli space $M(W, c, G; \alpha, \beta)_1$. Our boundary relations mostly involve counts of fully irreducible broken instantons, and we therefore need to rule out the possibility of small index instantons which are not fully irreducible. Because reducible flat connections on a cylinder are constant, there are two ways a connection could fail to be fully irreducible:

- It could be possible that all components of A are irreducible, but some interior limit is reducible;
- The connection A_W could be reducible.

The first type is easy to rule out.

Lemma 5.4. *Suppose A is a broken instanton on $(W^i_j, \hat{c}^i_j)$ or $(W^i_j, \check{c}^i_j)$ with respect to a metric $g \in \mathrm{int}(G^i_j)$. If the constituent instantons are all irreducible but A is not fully irreducible, then $i(A) \geq 2 - (i - j - 1)$.*

Proof. As discussed above, all irreducible instantons on (W^i_j, c, G^i_j) have $i(A_W) \geq -\dim G^i_j = -(i - j - 1)$, and all irreducible instantons on a cylinder have $i(B) \geq 1$. Any broken instanton has at least one component which is an instanton on a cylinder, and because A is not fully reducible at least one interior gluing factor is positive, so $i(A) \geq 1 + 1 - (i - j - 1)$. ☐

As for the second type, observe that if $W : Y \xrightarrow{L} Y'$ is a cobordism between integer homology spheres (with L a disjoint union of S^3s and $\mathbb{RP}^3$s) and A_W is reducible, then the positive and negative limits of A_W are also reducible, hence central because Y, Y' are integer homology spheres. Therefore, any broken instanton with irreducible outer limits for which A_W is reducible has at least two components which are instantons on cylinders, as well as two central gluing factors, so

$$i(A) \geq 1 + 3 + i(A_W) + 3 + 1 = 8 + i(A_W).$$

It suffices to give a suitable lower bound on the minimal index of such A_W.

Lemma 5.5. *If A_W is a reducible unbroken instanton on $(W^i_j, \hat{c}^i_j)$ or $(W^i_j, \check{c}^i_j)$ with respect to a metric $g \in G^i_j$, then $i(A_W) \geq -4$.*

Proof. In our case, the index formula of [MMR94, Proposition 8.4.2] simplifies to the following equation, where we write $L = L_1 \sqcup \cdots \sqcup L_n$ with $A_W|_{L_i} = \alpha_i$:

$$i(A_W) = 8\mathcal{E}(A_W) - 3(1 + b^+(W)) + \frac{1}{2}\sum_{i=1}^{n}(3 - h_{\alpha_i} - \rho(\alpha_i)). \tag{9}$$

The quantity h_{α_i} is 3 when α_i is central and 1 when α_i is abelian. For us, (L_i, w_i) are all either $(S^3, \varnothing)$ or $(\mathbb{RP}^3, w)$ for $[w] \neq 0 \in H_1(\mathbb{RP}^3; \mathbb{Z}/2)$, and the quantity $\rho(\alpha_i)$ is zero for any flat connection on S^3 or $\mathbb{RP}^3$ by explicit computation, e.g. [APS75, Proposition 2.12].

The manifold S^3 supports only the central flat connection, and $(\mathbb{RP}^3, w)$ with $[w] \neq 0 \in H_1(\mathbb{RP}^3; \mathbb{Z}/2)$ supports exactly one abelian flat connection. In particular, the sum over i gives twice the number of $\mathbb{RP}^3$ middle ends.

If the reducible A_W gives rise to a splitting $E_c \cong L_x \oplus L_y$, where $c = x + y$, then Chern–Weil theory gives that the energy term $8\mathcal{E}(A_W)$ is equal to $-2(x - y)^2 = -2(2x - c)^2$. The energy is non-negative and strictly positive unless A_W is projectively flat, which is the case precisely when $x - y = 2x - c$ is torsion. In particular, the energy is strictly positive when there is some $\mathbb{RP}^3$ end, as in this case $2x - c$ is never torsion.

For the manifolds W^i_j with $i - j \leq 3$, either $b^+(W) = 0$ or we have that $b^+(W) = 1$ and L is a disjoint union of $\mathbb{RP}^3$ and some number of copies of S^3 with $c|_{\mathbb{RP}^3}$ is non-trivial. In the former case, the index computation gives $i(A_W) \geq -3$. In the latter case, the index computation gives $i(A_W) > -6 + 1 = -5$ because $\mathcal{E}(A_W)$ is positive. □

These two lemmas will be the main way we *rule out* broken instantons which are not fully irreducible. However, there are some places in the argument of [CDX20] where reducible connections do enter in an essential way. In particular, we will need to understand the reducibles on the manifold (N, c) discussed in Section 5.2 for $c = N \cap \hat{c}^1_{-1}$ or $c = N \cap \check{c}^1_{-1}$.

Lemma 5.6. *For either generator $c \in H^2(N;\mathbb{Z})$, there is exactly one minimal-index reducible A_N on (N, c), for which $i(A_N) = -1$.*

Proof. By Hodge theory and the fact that $b^+(N) = 0$, reducible solutions to the ASD equation are in bijection with pairs $\{x, y\} \subset H^2(N;\mathbb{Z})$ for which $x + y = PD(c)$. Because $H^2(N;\mathbb{Z}) \cong \mathbb{Z}$ and c is a generator, there is one reducible solution $\{nc, (1-n)c\}$ for each integer $n \geq 1$.

As discussed above, the quantity $\mathcal{E}(A_W)$ in (8) for such a reducible is precisely $-2(x-y)^2$. The surface c represents a generator of $H^2(N;\mathbb{Z})$, for which $c^2 = -1/2$. Any pair $\{nc, (1-n)c\}$ gives $-2(x-y)^2 = (2n-1)^2$, which is minimized for $n = 1$. There is thus one minimal-index reducible.

Because (N, c) has boundary $(S^3, \varnothing) \sqcup (S^3, \varnothing) \sqcup (\mathbb{RP}^3, w)$, we see that A_N restricts trivially to two boundary components, and to an abelian connection on the third. Therefore, (8) gives

$$i(A_N) = 1 - 3 + 1 = -1. \quad \square$$

5.4. **Proof of Proposition 3.8.** Here we establish the existence of the maps f_i, g_i, h_i, as well as their boundary relations; by the discussion of Section 5.1, this gives us the desired exact triangle.

The general principle is as follows. If $(W, c) : Y \to Y'$ is a cobordism equipped with a family of metrics G, if the solutions to the G-ASD equation are cut out transversely and there are no reducible instantons with index less than $-\dim G$, then counting instantons over (W, c, G) with index equal to $-\dim G$ gives rise to a well-defined map $C_*(Y) \to C_*(Y')$ satisfying the relation

$$d_{Y'} \circ (W, c, G)_* + (W, c, G)_* \circ d_Y = (W, c, \partial G)_*.$$

The relation is derived by inspecting the ends of the moduli space of instantons $M^+(W, c, G; \alpha, \beta)_1$ with α and β irreducible.

Our task is to verify that instantons supported by the interior of each face of G are cut out transversely and to compute the map $(W, c, \partial G)_*$.

5.4.1. *The maps f_i.* Each manifold W^i_{i-1} is equipped with an unbroken metric. The manifold W^2_1 has $b^+(W) > 0$ and a geometric representative $c \subset W^2_1$ for a non-trivial admissible $U(2)$-bundle. A combination of Lemmas 5.4 and 5.5 imply that the moduli spaces of irreducible instantons on W^i_{i-1} of index 0, with irreducible outer limits, are compact, and thus consists of finitely many points.

We can therefore define

$$\begin{aligned}
f_2 &= (W^2_1, G^2_1, \hat{c}^2_1)_* : C_*(Z_{-1}) \to C_*(Z_1), \\
f_1 &= (W^1_0, G^1_0, \hat{c}^1_0)_* \oplus (W^1_0, G^1_0, \check{c}^1_0)_* : C_*(Z_1) \to C_*(Z_0) \oplus C_*(Z_0) \\
f_0 &= (W^0_{-1}, G^0_{-1}, \hat{c}^0_{-1})_* \oplus -(W^0_{-1}, G^0_{-1}, \check{c}^0_{-1}) : C_*(Z_0) \oplus C_*(Z_0) \to C_*(Z_{-1})
\end{aligned}$$

to be the maps which count index zero irreducible instantons with irreducible outer limits on the relevant cobordisms.

The moduli spaces of index 1 consists of fully irreducible solutions by Lemmas 5.4 and 5.5, where standard gluing techniques apply. We therefore obtain the following result.

Proposition 5.7. *The maps f_i are well-defined chain maps.*

5.4.2. *The maps g_i.* Now we consider the 1-parameter family of metrics G^i_{i-2} on W^i_{i-2}. We count instantons with irreducible limits and of index -1 with respect to this family to define the maps g_i, and use the moduli space of index 0 instantons with irreducible limits to prove the desired chain homotopy relation. The combination of Lemmas 5.4 and 5.5 imply that for any *unbroken* metric $g \in \mathrm{int}(G^i_{i-2})$ these moduli spaces consist of fully irreducible connections, but the broken metrics require somewhat more care. In fact, it is important to the argument that one boundary point of G^1_{-1} *does* support a broken instanton with a reducible component.

Lemma 5.8. *Suppose A is a broken instanton with irreducible outer limits on $(W^i_{i-2}, \hat{c}^i_{i-2})$ or $(W^i_{i-2}, \check{c}^i_{i-2})$ with respect to the metric broken along Z_{i-1}. If A is not fully irreducible, then $i(A) \geq 1$.*

Proof. Suppose A^i_{i-1} and A^{i-1}_{i-2} are the connections on W^i_{i-1} and W^{i-1}_{i-2} induced by A. If A has irreducible components but some reducible intermediate flat connection, then each constituent instanton has non-negative index but some positive gluing factor, so $i(A) \geq 1$. If A^i_{i-1} is reducible but A^{i-1}_{i-2} is not, then we have

$$i(A) \geq 1 + 3 + i(A^i_{i-1}) + 3 + i(A^{i-1}_{i-2}),$$

as A^i_{i-1} must be glued to some trajectory on the cylinder to have irreducible outer limits. Applying Lemma 5.5 and the fact that $i(A^{i-1}_{i-2}) \geq 0$, we see that $i(A) \geq 3$. The same argument applies to the case that A^{i-1}_{i-2} is reducible but A^i_{i-1} is not. Finally, if both the connections A^i_{i-1} and A^{i-1}_{i-2} are reducible, then Lemma 5.5 implies that

$$i(A) \geq 1 + 3 + (-4) + 3 + (-4) + 3 + 1 = 3. \qquad \square$$

The other endpoint of G^i_{i-2} corresponds to a decomposition $W^i_{i-2} = (-W^{i+1}_i) \cup_{-L_{i+1}} N$. The index i determines whether $-W^{i+1}_i$ has middle end diffeomorphic to S^3 or $\mathbb{RP}^3$, with the latter corresponding to the case $i \equiv 1 \mod 3$.

Lemma 5.9. *Suppose A is a broken instanton with irreducible outer limits on $(W^i_{i-2}, \hat{c}^i_{i-2})$ or $(W^i_{i-2}, \check{c}^i_{i-2})$ with respect to the metric broken along M^i_{i-2}. Suppose A is not fully itreducible. If $i \not\equiv 1 \mod 3$, then $i(A) \geq 2$. If $i \equiv 1 \mod 3$, then $i(A) \geq 0$ with equality if and only if $A = (A_W, A_N)$ for A_W an unbroken irreducible instanton of index zero on $W = -W^{i+1}_i$ and A_N the unique reducible of minimal index on (N, c).*

Proof. Suppose first that the component of A on $-W^{i+1}_i$ is reducible. As discussed before Lemma 5.5, this implies $i(A) \geq 8 + i(A_W) + i(A_N) + g$, where g is the dimension of the gluing factor associated to M^i_{i-2}. By Lemma 5.5 and Lemma 5.6, we see that in this case $i(A) \geq 4$. Thus we may assume that A_W is irreducible. We may also assume that A contains no instantons on a cylinder, as these increase the index by at least one. Thus, we may suppose $A = (A_W, A_N)$, for which the index is $i(A) = i(A_W) + i(A_N) + g$.

Because A_W is irreducible, $i(A_W) \geq 0$. By Lemma 5.6, we have $i(A_N) \geq -1$. Finally, $g = 1$ if $i \equiv 1 \mod 3$ and $g = 3$ otherwise. Thus if $i \equiv 1 \mod 3$ we have $i(A) \geq 0$ with equality if and only if $i(A_W) = 0$ and $i(A_N) = -1$, whereas if $i \not\equiv 1 \mod 3$ we have $i(A) \geq 2$. $\square$

Now we define maps

$$
\begin{aligned}
g_2 &= (W^2_0, G^2_0, \hat{c}^2_0)_* \oplus (W^2_0, G^2_0, \check{c}^2_0)_* : C_*(Z_{-1}) \to C_*(Z_0) \oplus C_*(Z_0)\\
g_1 &= (W^1_{-1}, G^1_{-1}, \hat{c}^1_{-1})_* \; - \; (W^1_{-1}, G^1_{-1}, \check{c}^1_{-1}) : C_*(Z_1) \to C_*(Z_{-1})\\
g_0 &= (W^0_{-2}, G^0_{-2}, \hat{c}^0_{-2})_* \oplus (W^0_{-2}, G^0_{-2}, \check{c}^0_{-2})_* : C_*(Z_0) \oplus C_*(Z_0) \to C_*(Z_1)
\end{aligned}
$$

by counting instantons with irreducible limits and index -1 in the respective family of metrics and perturbations, with the approptiate bundle. By a combination of Lemmas 5.4, 5.5 and the Lemmas 5.8-5.9 above, we see that this moduli space is compact, so the maps g_i are well defined. The boundary relations are slightly more subtle.

Proposition 5.10. *The maps g_i defined above satisfy the relations*

$$d_{i-2}g_i + f_{i-1}f_i + g_i d_i = 0.$$

Proof. We focus first on g_2; the same argument will apply for g_0 with essentially no change. The desired relation decomposes as a direct sum of two relations involving the two components of g_2 and the two components of f_2, f_0, and the argument is now standard.

More precisely, if A is a broken instanton for $g \in G^i_{i-2}$ with $i(A) \leq 0$, then A is fully irreducible, so one finds by the usual gluing techniques that

$$d_0(W^2_0, G^2_0, \hat{c}^2_0)_* + (W^2_0, G^2_0, \hat{c}^2_0)d_2 = (W^2_0, \partial G^2_0, \hat{c}^2_0)_*,$$

and similarly for $\check{c}^2_0$. The map induced by the incoming boundary component of $(\partial G^2_0, \hat{c}^2_0)$ is the composite map $(W^1_0, G^1_0, \hat{c}^1_0) \circ (W^2_1, G^2_1, \hat{c}^2_1)$ of the first component of $f_1 f_2$. By Lemma 5.9, the other boundary component induces the zero map. Running the same argument for $\check{c}^2_0$ (which gives the second component of $f_1 f_2$) and taking a direct sum, we obtain the relation $d_0 g_2 + f_1 f_2 + g_2 d_2 = 0$.

As for g_1, now the strategy instead must follow [CDX20, Proposition 5.14(i)]. The argument is slightly more subtle because the map induced by $(W^1_{-1}, \hat{c}^1_{-1})$, corresponding to the metric broken along $\mathbb{RP}^3$, is not automatically zero. Rather, this map is the same for $\hat{c}^1_{-1}$ and $\check{c}^1_{-1}$, so when we take the *difference* we obtain zero. This corresponds to the definition

$$f_0 f_1 = (W^0_{-1}, G^0_{-1}, \hat{c}^0_{-1})_*(W^1_0, G^1_0, \hat{c}^1_0) - (W^0_{-1}, G^0_{-1}, \check{c}^0_{-1})_*(W^1_0, G^1_0, \check{c}^1_0)_*.$$

To see that the maps induced by the bundles $\hat{c}^1_{-1}$ and $\check{c}^1_{-1}$ corresponding to the metrics broken along $\mathbb{RP}^3$ are the same, observe that

$$\hat{c}^1_{-1} \cap (-W^{-1}_{-2}) = -\hat{c}^{-1}_{-2} = -\check{c}^{-1}_{-2} = \check{c}^1_{-1} \cap (-W^{-1}_{-2}),$$

while the intersection with N gives the two generators of $H^2(N; \mathbb{Z})$. Now by Lemma 5.9, with respect to the metric broken along $\mathbb{RP}^3$, the only broken instantons of index zero have A_W an irreducible instanton of index zero, and A_N the unique minimal index reducible. Thus, for $\hat{c}$ this map coincides on the chain level with the map $(-W^{-1}_{-2}, -\hat{c}^{-1}_{-2})_*$, and similarly for $\check{c}$; but as discussed above $\hat{c}^{-1}_{-2} = \check{c}^{-1}_{-2}$, so indeed the two maps are the same. ☐

5.4.3. *The maps h_i.* We move on to the maps h_i and their boundary relations. These are defined by counting index -2 irreducible instantons on the cobordisms W^i_{i-3} with respect to the family of metrics G^i_{i-3}. For this family,

- we need to show there are no instantons of index less than -2 supported by metrics in the interior of G^i_{i-3};
- we need to show there are no instantons of index less than -1 supported by ∂G^i_{i-3};
- we need to determine the induced map of ∂G^i_{i-3}.

Because there are three families G^i_{i-3} and each family contains five boundary strata, this involves a great deal of case analysis. The following four lemmas in our case analysis follow exactly as in the previous section, and we omit their proofs.

The first lemma asserts that there are no small-index reducibles on the interior of the family. This lemma follow from Lemma 5.4.

Lemma 5.11. *Suppose A is a broken instanton with irreducible outer limits on $(W^i_{i-3}, \hat{c}^i_{i-3})$ or $(W^i_{i-3}, \check{c}^i_{i-3})$ with respect to one of the metrics $g \in \operatorname{int}(G^i_{i-3})$. Then $i(A) \geq -2$ with equality if and only if A is unbroken and hence fully irreducible.*

The second lemma asserts that there are no small-index reducibles on the boundary faces broken along S^3 (and, in fact, no small-index instantons whatsoever).

Lemma 5.12. *Suppose A is a broken instanton with irreducible outer limits on $(W^i_{i-3}, \hat{c}^i_{i-3})$ or $(W^i_{i-3}, \check{c}^i_{i-3})$ with respect to one of the metrics $g \in G^i_{i-3}$ broken along S^3. Then $i(A) \geq 1$.*

The third lemma asserts that there are no small-index reducibles on the boundary faces broken along the various Z_j, and that counting instantons on these boundary faces gives precisely the composite of the relevant instanton-counting maps.

Lemma 5.13. *Suppose A is a broken instanton with irreducible outer limits on $(W^i_{i-3}, \hat{c}^i_{i-3})$ or $(W^i_{i-3}, \check{c}^i_{i-3})$ with respect to one of the metrics $g \in G^i_{i-3}$ broken along Z_{i-1} or Z_{i-2}. Then $i(A) \geq -1$ with equality if and only if A is a broken instanton with exactly two irreducible pieces.*

Stating the fourth lemma requires some preparation. For $i = 1, -1$, the family G^i_{i-3} contains an interval of metrics $I^i_{\mathbb{RP}^3}$ broken along $\mathbb{RP}^3$; it restricts to a single metric on N, but an interval of metrics I^i_W on $W^i_{i-3} \setminus N = \bar{W}^i_{i-3}$. The following lemma asserts that the chain-level map induced by $(W^i_{i-3}, I^i_{\mathbb{RP}^3})$ is equal to the chain-level map induced by $(\bar{W}^i_{i-3}, I^i_W)$.

Lemma 5.14. *Suppose $i \not\equiv 0 \mod 3$, and that A is a broken instanton with irreducible outer limits on $(W^i_{i-3}, \hat{c}^i_{i-3})$ or $(W^i_{i-3}, \check{c}^i_{i-3})$ with respect to one of the metrics $g \in G^i_{i-3}$ broken along $\mathbb{RP}^3$. Then $i(A) \geq -1$, with equality if and only if A is a broken instanton with exactly two pieces, the piece on $\bar{W}^i_{i-3}$ being an unbroken irreducible of index -1 and the piece on N being the unique index -1 abelian connection.*

For the final face, consisting of metrics broken along $S^2 \times S^1$, the argument is somewhat different. We follow the argument of [CDX20, Section 6.3].

Recall that $M^i_{i-3} \cong S^2 \times S^1$ divides W^i_{i-3} into two pieces, one diffeomorphic to V_i given by the complement of a regular neighborhood of $\{0\} \times K_i$ in $[-1,1] \times Z_i$ and the other diffeomorphic to the complement N' of a neighborhood of an S^1 fiber in N. The corresponding one-parameter family of metrics is constant on V_i, and restricts to a family of metrics on N' denoted by $I^i_{S^2 \times S^1}$).

The character variety of flat $SU(2)$-connections on $S^2 \times S^1$ may be identified as

$$\mathfrak{X}(S^2 \times S^1) \cong \operatorname{Hom}\left(\pi_1(S^2 \times S^1), SU(2)\right)/\text{conj} \cong SU(2)/\text{conj} \cong [-1,1],$$

with the overall map given by sending $[A]$ to $\operatorname{tr}(\operatorname{Hol}_A(\{*\} \times S^1))/2$. The endpoint 1 corresponds to the trivial connection, the endpoint -1 corresponds to the non-trivial central connection whose holonomy along $\{*\} \times S^1$ is $-I$, while the interior points of the interval correspond to abelian connections.

Given a 4-manifold W whose boundary is not necessarily a union of rational homology spheres, we define the index of an ASD connection over W to be the index of the operator defined using weighted Sobolev spaces with a positive weight $\delta > 0$. For us, W will have $S^2 \times S^1$ as its only component which is not a rational homology sphere. Restriction gives a map $M(W) \to \mathfrak{X}(S^2 \times S^1)$, and the index of an ASD connection computes the expected dimension of the fiber of this map.

The following statement can be proved similarly to those above, and is discussed for general $SU(N)$ following [CDX20, Remark 6.33].

Lemma 5.15. *Suppose that A is a broken instanton with irreducible outer limits on $(W^i_{i-3}, \hat{c}^i_{i-3})$ or $(W^i_{i-3}, \check{c}^i_{i-3})$ with respect to one of the metrics $g \in G^i_{i-3}$ broken along $S^2 \times S^1$. Then $i(A) \geq -1$, with equality if and only if A is a broken instanton with exactly two pieces, the piece on V_i being an unbroken irreducible of index -1 and the piece on N' being a reducible instanton of index -2.*

That is, we may identify the moduli space of index -1 instantons $M(W^i_{i-3}, \hat{c}^i_{i-3}, I^i_{S^2\times S^1}; \alpha, \alpha')_0$ with the fiber product

$$M(V_i, \hat{c}^i_{i-3}; \alpha, \alpha')_0 \times_{\mathfrak{X}(S^2\times S^1)} M(N', \hat{c}^i_{i-3}, I^i_{S^2\times S^1})^{\mathrm{red}}_1,$$

and similarly with $\check{c}^i_{i-3}$. (The subscripts indicate the dimensions of the relevant spaces, not the indices of their consitutient instantons.) It follows that there is a well-defined map defined by counting instantons supported by the family of metrics $(W^i_{i-3}, \hat{c}^i_{i-3}, I^i_{S^2\times S^1})$ which have index equal to -1. A similar claim holds for $\check{c}^i_{i-3}$.

These put together, define the maps h_i by the following formulae. We have

$$h_i : C_*(Z_i) \to C_*(Z_i), \qquad h_i = \begin{cases} (W^i_{i-3}, G^i_{i-3}, \hat{c}^i_{i-3})_* + (W^i_{i-3}, G^i_{i-3}, \check{c}^i_{i-3})_* & i \in \{1, -1\} \\ (W^i_{i-3}, G^i_{i-3}, \hat{c}^i_{i-3})_* \oplus (W^i_{i-3}, G^i_{i-3}, \check{c}^i_{i-3})_* & i = 0. \end{cases}$$

It will be convenient to write $q_i : C_*(Z_i) \to C_*(Z_i)$ for the map defined by the same formula as above, but using the family of metrics $I^i_{S^2\times S^1}$ in place of G^i_{i-3}.

Proposition 5.16. *The maps h_i written above are well-defined, and satisfy the relation*

$$d_{i-3}h_i + f_{i-2}g_i + g_{i-1}f_i + h_i d_i = q_i.$$

Proof. The combination of Lemmas 5.11-5.15 imply that the relevant moduli spaces for the definition of h_i support only irreducible instantons of index -2 and no instantons of index ≤ -3, so the Uhlenbeck compactness theorem implies the relevant counts are finite. To investigate the relation, observe that Lemmas 5.11-5.15 imply that the interior of the moduli space of index -1 instantons is a smooth manifold of dimension 1, and that these lemmas identify the boundaries of these moduli spaces. The boundary face in G^i_{i-3} which breaks along Z_{i-1} contributes $g_{i-1}f_i$; the boundary face which breaks along Z_{i-2} contributes $f_{i-2}g_i$. The boundary faces which break along S^3 contribute the zero map. When $i = 0$, this leaves us only the map q_0 coming from the $S^2 \times S^1$ breakings. When $i \in \{1, -1\}$, both families $(W^i_{i-3}, G^i_{i-3}, \hat{c}^i_{i-3})$ and $(W^i_{i-3}, G^i_{i-3}, \check{c}^i_{i-3})$ also have a face corresponding to breaking along $\mathbb{RP}^3$. By Lemma 5.14, these two faces induce the same map, so these terms cancel out upon adding the maps for $\check{c}^i_{i-3}$ and $\hat{c}^i_{i-3}$. This leaves only the q_i terms. $\square$

It remains to us to compute the maps q_i.

Proposition 5.17. *The map q_i determined above is homotopic to the identity.*

Proof. Again we follow the $N = 2$ case of [CDX20, Section 6.3]. As depicted in Figure 3, the case $i = 0$ is simpler, as both boundary components of $I^i_{S^2\times S^1}$ correspond to breakings along spheres. In this case [CDX20, Proposition 5.54] identifies the moduli space $M(N', c, I^0_{S^2\times S^1})^{\mathrm{red}}_1 \cong [-1, 1]$. For $i \in \{1, -1\}$, one needs to take a little extra care in dealing with the $\mathbb{RP}^3$-broken metrics, and here it becomes essential that we work with both choices of geometric representative.

One can still identify

$$M(N', \hat{c}^i_{i-3}, I^i_{S^2\times S^1})^{\mathrm{red}}_1 \cong [-1, 0], \qquad M(N', \check{c}^i_{i-3}, I^i_{S^2\times S^1})^{\mathrm{red}}_1 \cong [0, 1]$$

with 0 corresponding to the metric broken along $\mathbb{RP}^3$ in both cases. These two metrics coincide. Furthermore, on the part of N' bounded by $\mathbb{RP}^3$ and $S^2 \times S^1$, the two choices of bundle data coincide. It follows that on this piece of N', the instantons associated to $0 \in [-1, 0]$ and $0 \in [0, 1]$ coincide, and in particular so do their restrictions to $S^2 \times S^1$. It follows that these two moduli spaces $[-1, 0]$ and $[0, 1]$ can be pasted together along the broken metric to obtain a single moduli space homeomorphic to $[-1, 1]$, equipped with a continuous map to $\mathfrak{X}(S^2 \times S^1) \cong [-1, 1]$.

Now the argument proceeds the same for all i. That this map sends $\partial[-1, 1]$ to $\partial\mathfrak{X}(S^2 \times S^1)$ identically — and thus the map from $[-1, 1]$ has degree 1 in an appropriate sense — is the $N = 2$ case of [CDX20, Proposition 5.53]. This is established by an explicit computation [CDX20, Proposition 5.40] for the broken metrics in $\partial[-1, 1]$, ultimately because the restriction to $\mathfrak{X}(S^2 \times S^1)$ can be determined by a curvature integral.

Thus, the map q_i coincides with the map induced by the broken metric on $V_i \sqcup D^3 \times S^1$, as the natural map $M(D^3 \times S^1)^{\text{red}}_1 \to \mathfrak{X}(S^2 \times S^1)$ also has degree 1. Choosing a homotopy from the broken metric on $V_i \sqcup D^3 \times S^1$ to the product metric on $I \times Z_i$ gives rise to a homotopy from the map q_i to the identity map. □

Proof of Proposition 3.8. Using the triangle detection lemma, we have constructed an exact triangle whose vertices are of the expected form. We need to discuss the degree of the relevant maps and verify the given description of f_1 and f_0 as cobordism maps. The degree computation follows from the index formula (8) and our understanding of the topology of the manifolds W^i_{i-1}. The map f_1 is by definition

$$(W^1_0, \hat{c}^1_0)_* \oplus (W^1_0, \check{c}^1_0)_* = (W, \varnothing)_* \oplus (W, c_+)_*,$$

where c_+ is the cocore, as stated in Proposition 3.8. The map f_0 is by definition

$$(W^0_{-1}, \check{c}^1_0)_* \oplus -(W^0_{-1}, \check{c}^0_{-1})_* = (W', c_-)_* \oplus -(W', c_- \sqcup c_+)_*,$$

with c_- the core and c_+ the cocore. Now observe that the induced map $(W', c)_*$ depends only on the relative homology class of c. Because W' has intersection form (-1) and $c_- \cap c_+$ is a single positively-oriented point, we see that $c_- = -c_+$ in $H^2(W'; \mathbb{Z})$. Thus, $(W', c_- \sqcup c_+)_*$ coincides with $(W', \varnothing)_*$. This completes the proof of Proposition 3.8. □

5.5. **Proof of Proposition 3.9.** In the case $Y = Z_0 = S^3$, the chain complex $CI_*(S^3)$ is identically zero, and the maps f_0 and f_1 are trivial. This implies that the map $g_1 : C(Z_1) \to C(Z_{-1})$ is a chain map because the composite $f_0 f_1$ factors through the trivial group, and we have

$$d_{-1} g_1 + g_1 d_1 = f_0 f_1 = 0.$$

Similarly, because $f_0 g_2$ and $g_0 f_1$ are trivial, the relations for h_1 and h_{-1} imply that f_{-1} and g_1 are chain homotopy inverse maps. In particular, g_1 induces an isomorphism on homology.

It suffices to explain why g_1 extends to an IP-morphism of degree 3 and level $\frac{1}{4} - \eta(K)$. The map g_1 is obtained by summing over two maps, both of which are obtained by counting instantons on W^1_{-1} with respect to an appropriate family of metrics G and $U(2)$-bundle. These $U(2)$-bundles correspond to geometric representatives $\hat{c}^1_{-1}, \check{c}^1_{-1} \subset W^1_{-1}$ for which $c^2 = -1$. Because W^1_{-1} has rational homology sphere ends and $b_1(W) = b^+(W) = 0$, the degree and level of such a map is given by

$$D = \dim G - 2c^2 = 1 + 2 = 3, \qquad L = -c^2/4 - \eta(K) = 1/4 - \eta(K)$$

by a slight generalization of Proposition 3.4 to the case that W is equipped with a family of metrics.

6. Alexander polynomial constraints

In this section, we give a proof of the following result.

Theorem 6.1. *Suppose that K is a knot in S^3 and ± 2 is a cosmetic pair. Then $\Delta_K = 1$.*

Let K be a knot in a homology sphere Y. Let $X_\pm(K)$ denote the unique 2-sheeted cover of $Y_{\pm 2}(K)$. If ± 2 is a cosmetic pair for K, then $X_+(K)$ is orientation-preserving diffeomorphic to $X_-(K)$. We will use the following well-known surgery description of $X_\pm(K)$. Writing $\Sigma_2(K)$ for the branched double cover of K, the preimage $\widetilde{K} \subset \Sigma_2(K)$ remains null-homologous, so ± 1 surgery on $\widetilde{K}$ makes sense.

Lemma 6.2. *The manifold $X_\pm(K)$ is obtained from $\Sigma_2(K)$ by ± 1 surgery on $\widetilde{K}$.*

Proof. Since $Y_{\pm 2}(K)$ is obtained from Dehn filling the exterior of K, the double cover is described by a suitable Dehn filling of a double cover of $Y - K$. Since $H_1(Y - K)$ surjects onto $H_1(Y_{\pm 2}(K))$, we are Dehn filling the non-trivial double cover. The slope is simply the lift of the slope downstairs, which is $\pm 2\mu_K + \lambda_K$. The preimage of μ_K is $\mu_{\widetilde{K}}$ and the preimage of λ_K is two copies of $\lambda_{\widetilde{K}}$. Hence, the preimage of $\pm 2\mu_K + \lambda_K$ is two copies of $\pm\mu_{\widetilde{K}} + \lambda_{\widetilde{K}}$ and the result follows. $\square$

Proof of Theorem 6.1. Let K be a non-trivial knot in S^3 for which ± 2 is a cosmetic pair. We have $g(K) = 2$ by [Han23, Theorem 2(ii)], so the Alexander polynomial takes the form

$$\Delta_K(t) = at^2 + bt + c + bt^{-1} + at^{-2},$$

with a, b possibly zero. We will also use that the preimage $\widetilde{K}$ has Alexander polynomial

$$\Delta_{\widetilde{K}}(t) = \Delta_K(t^{1/2})\Delta_K(-t^{1/2}) = a^2t^2 + (2ac - b^2)t + (2a^2 - 2b^2 + c^2) + (2ac - b^2)t^{-1} + a^2t^{-2};$$

see e.g. [HKL10, Corollary 4.2]. Now [BL90, p182, Consequence 2] states that whenever a knot $J \subset Y$ admits cosmetic surgeries, its Alexander polynomial has $\Delta''_J(1) = 0$. Applying this to K gives $8a + 2b = 0$ and so $b = -4a$. Since $|\Delta_K(1)| = 1$, we also have $c = 6a \pm 1$.

Next, we study the Casson–Walker invariants of $X_\pm(K)$. Appealing to Lemma 6.2 and [BL90, p182, Consequence 2] once more, we see that $\Delta''_{\widetilde{K}}(1) = 0$ as well. As

$$\Delta''_{\widetilde{K}}(1) = 2a^2 + 2(2ac - b^2) + 6a^2 = 8a^2 + 4a(6a \pm 1) - 2(-4a)^2 = \pm 4a,$$

so combining these means $a = 0$ and hence $\Delta_K = 1$. $\square$

References

[APS75] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Spectral asymmetry and Riemannian geometry. II*, Math. Proc. Cambridge Philos. Soc. **78** (1975), no. 3, 405–432. MR397798

[BD95] P. J. Braam and S. K. Donaldson, *Floer's work on instanton homology, knots and surgery*, The Floer memorial volume, 1995, pp. 195–256. MR1362829

[BL90] Steven Boyer and Daniel Lines, *Surgery formulae for Casson's invariant and extensions to homology lens spaces*, J. Reine Angew. Math. **405** (1990), 181–220. MR1041002

[CB15] William Crawley-Boevey, *Decomposition of pointwise finite-dimensional persistence modules*, J. Algebra Appl. **14** (2015), no. 5, 1550066, 8. MR3323327

[CDX20] Lucas Culler, Aliakbar Daemi, and Yi Xie, *Surgery, polygons and* SU(N)*-Floer homology*, J. Topol. **13** (2020), no. 2, 576–668. MR4092776

[Dae20] Aliakbar Daemi, *Chern-Simons functional and the homology cobordism group*, Duke Math. J. **169** (2020), no. 15, 2827–2886. MR4158669

[Det21] Renaud Detcherry, *A quantum obstruction to purely cosmetic surgeries*, 2021. arXiv:2102.11323.

[DLVVW22] Aliakbar Daemi, Tye Lidman, David Shea Vela-Vick, and C.-M. Michael Wong, *Ribbon homology cobordisms*, Adv. Math. **408** (2022), Paper No. 108580, 68. MR4467148

[DM91] Michel Domergue and Yves Mathieu, *Nœuds qui ne sont pas déterminés par leur complément dans les 3-variétés à bord*, Bulletin de la Société Mathématique de France **119** (1991), no. 3, 327–341 (fr). MR92i:57003

[Don02] S. K. Donaldson, *Floer homology groups in Yang-Mills theory*, Cambridge Tracts in Mathematics, vol. 147, Cambridge University Press, Cambridge, 2002. With the assistance of M. Furuta and D. Kotschick. MR1883043

[ELZ02] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian, *Topological persistence and simplification*, 2002, pp. 511–533. Discrete and computational geometry and graph drawing (Columbia, SC, 2001). MR1949898

[Flo88] Andreas Floer, *An instanton-invariant for* 3*-manifolds*, Comm. Math. Phys. **118** (1988), no. 2, 215–240. MR956166

[Flo90] ———, *Instanton homology, surgery, and knots*, Geometry of low-dimensional manifolds, 1 (Durham, 1989), 1990, pp. 97–114. MR1171893

[FPS22] David Futer, Jessica S. Purcell, and Saul Schleimer, *Effective bilipschitz bounds on drilling and filling*, Geom. Topol. **26** (2022), no. 3, 1077–1188. MR4466646

[FPS24] ———, *Excluding cosmetic surgeries on hyperbolic 3-manifolds*, 2024. arXiv:2403.10448.

[FQ90] Michael H. Freedman and Frank Quinn, *Topology of 4-manifolds*, Princeton Mathematical Series, vol. 39, Princeton University Press, Princeton, NJ, 1990. MR1201584

[Frø02] Kim A. Frøyshov, *Equivariant aspects of Yang-Mills Floer theory*, Topology **41** (2002), no. 3, 525–552. MR1910040

[Fre84] Michael H. Freedman, *The disk theorem for four-dimensional manifolds*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983), 1984, pp. 647–663. MR804721

[Fur90] Mikio Furuta, *Homology cobordism group of homology* 3*-spheres*, Invent. Math. **100** (1990), no. 2, 339–355. MR1047138

[Gab87] David Gabai, *Foliations and the topology of* 3*-manifolds. II*, J. Differential Geom. **26** (1987), no. 3, 461–478. MR910017

[GL89] C. McA. Gordon and J. Luecke, *Knots are determined by their complements*, J. Amer. Math. Soc. **2** (1989), no. 2, 371–415. MR965210

[Gor91] Cameron McA. Gordon, *Dehn surgery on knots*, Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990), 1991, pp. 631–642. MR1159250

[Han23] Jonathan Hanselman, *Heegaard Floer homology and cosmetic surgeries in* S^3, J. Eur. Math. Soc. (JEMS) **25** (2023), no. 5, 1627–1669. MR4592856

[HK11] Matthew Hedden and Paul Kirk, *Chern-Simons invariants,* SO(3) *instantons, and* $\mathbb{Z}/2$ *homology cobordism*, Chern-Simons gauge theory: 20 years after, 2011, pp. 83–114. MR2809448

[HKL10] Chris Herald, Paul Kirk, and Charles Livingston, *Metabelian representations, twisted Alexander polynomials, knot slicing, and mutation*, Math. Z. **265** (2010), no. 4, 925–949. MR2652542

[IJ24] Kazuhiro Ichihara and In Dae Jong, *Large alternating montesinos knots do not admit purely cosmetic surgeries*, 2024.

[Ito23] Tetsuya Ito, *A remark on the finiteness of purely cosmetic surgeries*, Algebr. Geom. Topol. **23** (2023), no. 5, 2213–2219. MR4621429

[Kir97] *Problems in low-dimensional topology*, Geometric topology (Athens, GA, 1993), 1997, pp. 35–473. MR1470751

[KM10] Peter Kronheimer and Tomasz Mrowka, *Knots, sutures, and excision*, J. Differential Geom. **84** (2010), no. 2, 301–364. MR2652464

[KMOS07] P. Kronheimer, T. Mrowka, P. Ozsváth, and Z. Szabó, *Monopoles and lens space surgeries*, Ann. of Math. (2) **165** (2007), no. 2, 457–546. MR2299739

[LLP23] Adam Simon Levine, Tye Lidman, and Lisa Piccirillo, *New constructions and invariants of closed exotic 4-manifolds*, arXiv e-prints (July 2023), arXiv:2307.08130, available at `2307.08130`.

[LPCZ23] Tye Lidman, Juanita Pinzón-Caicedo, and Raphael Zentner, *Toroidal integer homology three-spheres have irreducible* $SU(2)$*-representations*, J. Topol. **16** (2023), no. 1, 344–367. MR4575869

[MMR94] John W. Morgan, Tomasz Mrowka, and Daniel Ruberman, *The* L^2*-moduli space and a vanishing theorem for Donaldson polynomial invariants*, Monographs in Geometry and Topology, vol. II, International Press, Cambridge, MA, 1994. MR1287851

[NST24] Yuta Nozaki, Kouki Sato, and Masaki Taniguchi, *Filtered instanton Floer homology and the homology cobordism group*, J. Eur. Math. Soc. (JEMS) **26** (2024), no. 12, 4699–4761. MR4780492

[NW15] Yi Ni and Zhongtao Wu, *Cosmetic surgeries on knots in* S^3, J. Reine Angew. Math. **706** (2015), 1–17. MR3393360

[OS04a] Peter Ozsváth and Zoltán Szabó, *Holomorphic disks and genus bounds*, Geom. Topol. **8** (2004), 311–334. MR2023281

[OS04b] ———, *Holomorphic disks and knot invariants*, Adv. Math. **186** (2004), no. 1, 58–116. MR2065507

[OS05] Peter Ozsváth and Zoltán Szabó, *On the Heegaard Floer homology of branched double-covers*, Adv. Math. **194** (2005), no. 1, 1–33. MR2141852

[OS08] Peter S. Ozsváth and Zoltán Szabó, *Knot Floer homology and integer surgeries*, Algebr. Geom. Topol. **8** (2008), no. 1, 101–153. MR2377279

[OS11] ———, *Knot Floer homology and rational surgeries*, Algebr. Geom. Topol. **11** (2011), no. 1, 1–68. MR2764036

[Rav18] Huygens C. Ravelomanana, *Exceptional cosmetic surgeries on homology spheres*, Topology Appl. **240** (2018), 249–258. MR3784409

[Sca15] Christopher W. Scaduto, *Instantons and odd Khovanov homology*, J. Topol. **8** (2015), no. 3, 744–810. MR3394316

[Sch85] Martin Scharlemann, *Smooth spheres in* $\mathbf{R}^4$ *with four critical points are standard*, Invent. Math. **79** (1985), no. 1, 125–141. MR774532

[Sei08] Paul Seidel, *Fukaya categories and Picard-Lefschetz theory*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2008. MR2441780

[SS21] András I. Stipsicz and Zoltán Szabó, *Purely cosmetic surgeries and pretzel knots*, Pacific J. Math. **313** (2021), no. 1, 195–211. MR4313433

[Tao22] Ran Tao, *Knots admitting purely cosmetic surgeries are prime*, Topology Appl. **322** (2022), Paper No. 108270, 11. MR4498360

[Wan06] Jiajun Wang, *Cosmetic surgeries on genus one knots*, Algebraic & Geometric Topology **6** (2006), no. 3, 1491 –1517.

[Wan22] Joshua Wang, *The cosmetic crossing conjecture for split links*, Geom. Topol. **26** (2022), no. 7, 2941–3053. MR4540900

[Wu11] Zhongtao Wu, *Cosmetic surgery in L-space homology spheres*, Geom. Topol. **15** (2011), no. 2, 1157–1168. MR2831258

[ZC05] Afra Zomorodian and Gunnar Carlsson, *Computing persistent homology*, Discrete Comput. Geom. **33** (2005), no. 2, 249–274. MR2121296

Department of Mathematics and Statistics, Washington University in St. Louis
Email address: adaemi@wustl.edu

Department of Mathematics, North Carolina State University, Raleigh, NC 27607
Email address: tlid@math.ncsu.edu

Department of Mathematics, University of Vermont, Burlington, VT 05405
Email address: Mike.Miller-Eismeier@uvm.edu