Received: 6 July 2021 | Accepted: 3 November 2021

DOI: 10.1112/topo.12218

RESEARCH ARTICLE

Journal of Topology

Instanton Floer homology, sutures, and Heegaard diagrams

Zhenkun Li[1] | **Fan Ye**[2]

[1] Department of Mathematics, Stanford University, Stanford, California, USA

[2] Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, UK

Correspondence
Fan Ye, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, UK.
Email: fy260@cam.ac.uk

Funding information
NSF, Grant/Award Number: 1808794

Abstract

This paper establishes a new technique that enables us to access some fundamental structural properties of instanton Floer homology. As an application, we establish, for the first time, a relation between the instanton Floer homology of a 3-manifold or a null-homologous knot inside a 3-manifold and the Heegaard diagram of that 3-manifold or knot. We further use this relation to compute the instanton knot homology of some families of $(1,1)$-knots, including all torus knots in S^3, which were mostly unknown before. As a second application, we also study the relation between the instanton knot homology $KHI(Y,K)$ and the framed instanton Floer homology $I^\sharp(Y)$. In particular, we prove the inequality $\dim_{\mathbb{C}} I^\sharp(Y) \leqslant \dim_{\mathbb{C}} KHI(Y,K)$ for all rationally null-homologous knots $K \subset Y$ and we constructed a new decomposition of the framed instanton Floer homology of Dehn surgeries along K that corresponds to the decomposition along torsion spin^c decompositions in monopole and Heegaard Floer theory.

MSC 2020
57K10 (primary), 57K14, 57K16, 57K18, 57K31 (secondary)

Contents

1. INTRODUCTION . . . 40
2. INSTANTON FLOER HOMOLOGY AND BALANCED SUTURED MANIFOLDS . . . 47
3. INSTANTON FLOER HOMOLOGY AND HEEGAARD DIAGRAMS . . . 57
4. INSTANTON FLOER HOMOLOGY AND THE DECOMPOSITION . . . 84
5. SOME REMARKS AND FUTURE DIRECTIONS . . . 104
ACKNOWLEDGEMENTS . . . 105
REFERENCES . . . 106

1 | INTRODUCTION

The instanton homology of closed 3-manifolds and knots in 3-manifolds was introduced by Floer [12, 13], which became a powerful tool in the study of 3-dimensional topology. Some related constructions were made by Kronheimer and Mrowka [34–36], the first author [42], and Daemi and Scaduto [11]. Apart from instanton Floer homology, there are three Floer homologies of closed 3-manifolds, knots, and balanced sutured manifolds: Heegaard Floer homology by Ozsváth and Szabó [52, 53], Rasmussen [55], and Juhász [26], monopole Floer homology by Kronheimer and Mrowka [32, 34], and embedded contact homology (ECH) by Hutchings [25], Colin, Ghiggini, Honda, and Hutchings [10]. For closed 3-manifolds, all these three Floer homologies are isomorphic by work of Kutluhan, Lee, and Taubes [38], or Taubes [61] combined with Colin, Ghiggini, and Honda [9]. However, instanton Floer homology remains isolated from the rest. The following conjecture is still open.

Conjecture 1.1 [34, Conjecture 7.24]. *For a balanced sutured manifold (M,γ), we have*

$$SHI(M,\gamma)\cong SFH(M,\gamma)\otimes\mathbb{C}.$$

In particular, for a knot K in a closed 3-manifold Y, there are isomorphisms

$$I^{\sharp}(Y)\cong\widehat{HF}(Y)\otimes\mathbb{C}\text{ and }KHI(Y,K)\cong\widehat{HFK}(Y,K)\otimes\mathbb{C}.$$

Here SHI is sutured instanton Floer homology [34], SFH is sutured (Heegaard) Floer homology [26], $I^{\sharp}$ is framed instanton Floer homology [35], $\widehat{HF}$ is the hat version of Heegaard Floer homology [53], KHI is instanton knot homology [34], and $\widehat{HFK}$ is (Heegaard) knot Floer homology [52, 55].

Instanton Floer homology is closely related to the representations of the fundamental groups and many other topological properties of 3-manifolds and knots. For example, it is the essential ingredient in proving the property P conjecture [31] and the fact that Khovanov homology detects the unknot [35]. Despite those remarkable applications, many fundamental structural properties of instanton Floer homology remain unknown. We propose a few of them here:

(1) Instanton Floer homology serves as a topological invariant for 3-manifolds and knots. On the other hand, Heegaard diagrams are one of the most important ways to describe 3-manifolds and knots and are the basis for Heegaard Floer homology as well. So is it possible to relate instanton Floer homology with the Heegaard diagrams of 3-manifolds and knots?

(2) The monopole and Heegaard Floer homology of closed 3-manifolds both decompose along spin^c structures. Non-torsion spin^c structures have their correspondence in instanton theory by looking at the simultaneous generalized eigenspace decompositions. However, the simultaneous generalized eigenspace decomposition of instanton Floer homology does not distinguish the torsions of the homology group of 3-manifolds and hence is trivial for all rational homology spheres. So is it possible to obtain a decomposition of instanton Floer homology corresponding to the torsions? This new decomposition would also be the prerequisite for the fourth problem.
(3) Can we understand the Euler characteristic of instanton Floer homology? Can we relate it to some other topological invariants of the 3-manifold?
(4) Can we relate the instanton Floer homology of knots and 3-manifolds? In particular, can we derive a surgery formula for the instanton Floer homology of Dehn surgeries along knots?

This paper develops a new technique that enables us to access those fundamental questions listed above. In this paper, we mainly deal with the first and the second and provide some answers to the fourth questions. First, toward answering the first question, we establish the following.

Theorem 1.2. *Suppose Y is a rational homology sphere, and $K \subset Y$ is a knot. Suppose $(\Sigma, \alpha, \beta, z, w)$ is a doubly pointed Heegaard diagram of (Y, K). Then there is a balanced sutured handlebody (H, γ) constructed from $(\Sigma, \alpha, \beta, z, w)$ (cf. Subsection 3.1), so that the followings hold*

$$\dim_{\mathbb{C}} I^{\sharp}(-Y) \leqslant \dim_{\mathbb{C}} KHI(-Y, K) \leqslant \dim_{\mathbb{C}} SHI(-H, -\gamma).$$

Remark 1.3. For most arguments in this paper, there are minus signs before the manifold and the suture, which means that we take the reverse orientation. This is because the proofs are based on contact gluing maps for sutured instanton homology (*cf.* Subsection 2.3).

The proof of Theorem 1.2 makes use of rationally null-homologous tangles in balanced sutured manifolds. In particular, we proved the following proposition.

Proposition 1.4. *Suppose (M, γ) is a balanced sutured manifold and T is a connected vertical tangle in (M, γ) (cf. Definition 3.1). Suppose $M_T = M \backslash N(T)$ and $\gamma_T = \gamma \cup m_T$, where m_T is the meridian of T. If $[T] = 0 \in H_1(M, \partial M; \mathbb{Q})$, then we have*

$$\dim_{\mathbb{C}} SHI(-M, -\gamma) \leqslant \dim_{\mathbb{C}} SHI(-M_T, -\gamma_T). \tag{1}$$

By Proposition 1.4, we also prove a generalization of the first inequality in Theorem 1.2, which generalizes the result for null-homologous knots by Wang [62, Proposition 1.18].

Proposition 1.5. *Suppose Y is a closed 3-manifold and $K \subset Y$ is a knot such that*

$$[K] = 0 \in H_1(Y; \mathbb{Q}).$$

Then we have

$$\dim_{\mathbb{C}} I^{\sharp}(-Y) \leqslant \dim_{\mathbb{C}} KHI(-Y, K).$$

In Theorem 1.2, we bound the dimensions of $I^\sharp(-Y)$ and $KHI(-Y,K)$ by the dimension of sutured instanton homology $SHI(-H,-\gamma)$, which is still difficult to compute in general. However, in the case where H is a handlebody, an upper bound of $\dim_{\mathbb{C}} SHI(H,\gamma)$ can be calculated via bypass exact triangles (for bypass exact triangle, *cf.* [5, Theorem 1.21], and for the algorithm to obtain an upper bound, *cf.* [15, Section 4]). In particular, we apply this theorem to $(1,1)$-knots in lens spaces, whose Heegaard diagrams are described in Proposition 3.28, and obtain the following theorem.

Theorem 1.6. *Suppose Y is a lens space, and $K \subset Y$ is a $(1,1)$-knot. Then we have*

$$\dim_{\mathbb{C}} KHI(Y,K) \leqslant \mathrm{rk}_{\mathbb{Z}} \widehat{HFK}(Y,K).$$

Prior to the current paper, there are two main approaches to estimate the dimension of KHI. The first is via the spectral sequence from Khovanov homology to instanton knot homology established by Kronheimer and Mrowka [35]. This bound is sharp for all alternating knots and many other knots. However, Khovanov homology is only defined for knots in S^3, so we cannot have any information for knots in other 3-manifolds. The second way is to study a set of explicit generators of the instanton knot homology and its variances for some special families of knots, and the number of generators bounds the dimension of homology. This idea has been exploited by Hedden, Herald, and Kirk [21] and Daemi and Scaduto [11]. Our Theorem 1.2 and Theorem 1.6 then offers a totally new way to obtain an upper bound of $\dim_{\mathbb{C}} KHI$, and the following corollary indicates that this bound is sharp for many examples.

Corollary 1.7. *Suppose $K \subset S^3$ is a $(1,1)$-knot that is also a Heegaard Floer L-space knot. Then*

$$\dim_{\mathbb{C}} KHI(S^3,K) = \mathrm{rk}_{\mathbb{Z}} \widehat{HFK}(S^3,K).$$

Proof. Suppose the Alexander polynomial of K is $\Delta_K(t) = \sum_{i\in\mathbb{Z}} c_i t^i$. From Ozsváth and Szabó [54, Theorem 1.2], we have

$$\mathrm{rk}_{\mathbb{Z}} \widehat{HFK}(S^3,K) = \sum_{i\in\mathbb{Z}} |c_i|.$$

In instanton theory, the main result of Kronheimer and Mrowka [33], or Lim [47], states that the Euler characteristic of the ith grading of $KHI(S^3,K)$ equals $\pm c_i$. As a result, we have

$$\dim_{\mathbb{C}} KHI(S^3,K) \geqslant \sum_{i\in\mathbb{Z}} |c_i|.$$

Hence Theorem 1.6 applies and we conclude the desired equality. $\square$

Corollary 1.7 would provide many examples whose related spectral sequences from Khovanov homology to instanton knot homology have some nontrivial intermediate pages. In particular, for torus knots, previously there were only partial computations of KHI from the related specatral sequences (cf. [37, 48]), while Corollary 1.7 applies to torus knots directly since torus knots admit lens spaces surgeries (*cf.* Moser [49]).

Corollary 1.8. *For a torus knot $K = T_{(p,q)}$, suppose its Alexander polynomial is*

$$\Delta_K(t) = t^{-\frac{(p-1)(q-1)}{2}} \frac{(t^{pq}-1)(t-1)}{(t^p-1)(t^q-1)} = \sum_{i=-\frac{(p-1)(q-1)}{2}}^{\frac{(p-1)(q-1)}{2}} c_i t^i.$$

Then we have

$$\dim_{\mathbb{C}} KHI(S^3, K, i) = |c_i|,$$

where i denotes the Alexander grading of $KHI(S^3, K)$.

In general lens spaces, we may obtain lower bounds from graded Euler characteristics of SHI and provide more examples of Conjecture 1.1. For example, the Heegaard Floer homology of constrained knots in lens spaces, introduced by the second author [64], is determined by their graded Euler characteristics. However, in the manifold other than S^3, it is not known if graded Euler characteristics of $\widehat{HFK}$ and KHI are equal. We will deal with the graded Euler characteristics in a forthcoming paper [45].

A (*Heegaard*) *Floer simple knot* is a knot K in a rational homology sphere Y such that

$$\mathrm{rk}_{\mathbb{Z}} \widehat{HFK}(Y, K) = \mathrm{rk}_{\mathbb{Z}} \widehat{HF}(Y) = |H_1(Y; \mathbb{Z})|, \tag{1.1}$$

where $|H_1(Y; \mathbb{Z})|$ is the order of the first homology group of Y. A (*Heegaard Floer*) *L-space* is a rational homology sphere Y satisfies the latter equality in (1.1). Examples of Floer simple knots include simple knots in lens spaces (*cf.* Definition 3.29; see also [57, Section 2.1]). Rasmussen and Rasmussen [58] proved that for Floer simple knots, there is an interval in $\mathbb{Q} \cup \{\infty\}$ so that a Dehn surgery gives an L-space if and only if the surgery slope is in the interval. In the interior of the interval, the dual knots are also Floer simple knots.

A similar result may hold for instanton Floer homology. A knot K in a rational homology sphere Y is called an *instanton Floer simple knot* if

$$\dim_{\mathbb{C}} KHI(Y, K) = \dim_{\mathbb{C}} I^\sharp(Y) = |H_1(Y; \mathbb{Z})|. \tag{1.2}$$

An *instanton L-space* is a rational homology sphere Y that satisfies the latter equality in (1.2).

Proposition 1.9. *Simple knots in lens spaces are instanton Floer simple knots.*

Proof. Suppose K is a simple knot in $Y = L(p, q)$. Combining Theorem 1.2, Theorem 1.6 and a direct calculation of knot Floer homology, we have

$$\dim_{\mathbb{C}} I^\sharp(Y) \leqslant \dim_{\mathbb{C}} KHI(Y, K) \leqslant \mathrm{rk}_{\mathbb{Z}} \widehat{HFK}(Y, K) = p. \tag{4}$$

By Scaduto [59, Corollary 1.4], we have

$$\dim_{\mathbb{C}} I^\sharp(Y) \geqslant |H_1(Y)| = p. \tag{4}$$

Hence we conclude the desired equality. □

Inspired by the result about Floer simple knots by Rasmussen and Rasmussen [58], we prove the following theorem for simple knots in lens spaces.

Theorem 1.10. *Suppose K is a simple knot in a lens space Y. Fixing a framing on $\partial Y(K)$ by picking an arbitrary longitude of K, then there exists an integer $N \geqslant 0$ so that for any $r \in \mathbb{Q}$ with $|r| \geqslant N$, the manifold obtained by a surgery of slope r along K is an instanton L-space, and the dual knot is also an instanton Floer simple knot.*

Remark 1.11. In general, framed instanton Floer homology is very difficult to compute. Only some special families of 3-manifolds were studied (cf. [1, 7, 46, 60]). Theorem 1.10 provides many new examples whose framed instanton Floer homology can be computed.

Towards answering the second question regarding the decomposition of instanton Floer homology, we establish the following.

Theorem 1.12. *Suppose $\widehat{Y}$ is a closed 3-manifold, and $\widehat{K} \subset \widehat{Y}$ is a rationally null-homologous knot. Let $\widehat{Y}(\widehat{K}) = \widehat{Y}\backslash\mathrm{int}(N(\widehat{K}))$ be the knot complement and let S be a Seifert surface of $\widehat{K}$. Suppose further that $\lambda = \partial S \subset \partial\widehat{Y}(\widehat{K})$ is connected and $|\hat{\mu}\cdot\lambda| = q$, where $\hat{\mu}$ is the meridian of $\widehat{K}$ on $\partial\widehat{Y}(\widehat{K})$ and the dot $\cdot$ denotes the algebraic intersection number. Then there is a decomposition associated to $\widehat{K}$ up to isomorphism:*

$$I^{\sharp}(\widehat{Y}) \cong \bigoplus_{i=0}^{q-1} I^{\sharp}(\widehat{Y}, i).$$

When $H_1(\widehat{Y}) = \mathbb{Z}_q$ and $\widehat{K}$ represents a generator of $H_1(\widehat{Y})$, we can regard the decomposition in Theorem 1.12 as an analog of the torsion spin^c decompositions

$$\widetilde{HM}(\widehat{Y}) = \bigoplus_{\mathfrak{s}\in\mathrm{Spin}^c(\widehat{Y})} \widetilde{HM}(\widehat{Y},\mathfrak{s}) \text{ and } \widehat{HF}(\widehat{Y}) = \bigoplus_{\mathfrak{s}\in\mathrm{Spin}^c(\widehat{Y})} \widehat{HF}(\widehat{Y},\mathfrak{s}).$$

Here $\widetilde{HM}$ is the tilde version of monopole Floer homology [8].

To provide some evidence, we will prove the following result in a forthcoming paper [45].

Proposition 1.13. *Under the hypothesis and the statement of Theorem 1.12, there is a well-defined $\mathbb{Z}_2$ grading on $I^{\sharp}(\widehat{Y}, i)$. Suppose $\mu \subset \partial\widehat{Y}(\widehat{K})$ is a simple closed curve so that $|\mu\cdot\lambda| = 1$ and suppose Y is the manifold obtained by Dehn filling along μ. For any integer $i \in [0, q-1]$, we have*

$$\chi(I^{\sharp}(\widehat{Y}, i)) = \chi(I^{\sharp}(Y)).$$

In particular, if $\widehat{Y}$ is an instanton L-space and $Y = S^3$, then for any integer $i \in [0, q-1]$, we have

$$I^{\sharp}(\widehat{Y}, i) \cong \mathbb{C}.$$

Remark 1.14. The defect of the decomposition of $I^{\sharp}(\widehat{Y})$ in Theorem 1.12 is that currently, we do not know if it is independent of the choice of $\widehat{K}$ inside $\widehat{Y}$.

For integral surgeries on a null-homologous knot, we obtain more results inspired by the large surgery formula in Heegaard Floer homology (cf. [52, Theorem 4.1]). For a null-homologous knot K in a closed 3-manifold Y, let the basis of $\partial Y(K)$ be formed by the meridian of K and the boundary of a Seifert surface.

Proposition 1.15. *Suppose Y is a closed 3-manifold and $K \subset Y$ is a null-homologous knot. Suppose n is an integer satisfying $n \geqslant 2g(K)+1$ and suppose $\widehat{Y}_n$ is obtained from Y by performing the n surgery along K. For any integer $i \in [0, n-2g(K)-1] \cup \{n-1\}$, we have*

$$I^{\sharp}(\widehat{Y}_n, i) \cong I^{\sharp}(Y).$$

For any integer $i \in [0, n-1]$, we have

$$I^{\sharp}(\widehat{Y}_n, i) \cong I^{\sharp}(\widehat{Y}_{n+1}, i+1).$$

Thus, we have

$$\dim_{\mathbb{C}} I^{\sharp}(\widehat{Y}_{n+1}) - \dim I^{\sharp}(\widehat{Y}_n) = \dim_{\mathbb{C}} I^{\sharp}(Y).$$

The proofs of Theorem 1.12 and Proposition 1.15 make use of $SHI(\widehat{Y}(\widehat{K}), \gamma)$ for some special suture $\gamma \subset \partial \widehat{Y}(\widehat{K})$. In [42, Section 3], the first author constructed a grading, *that is*, a decomposition of SHI associated to a properly embedded surface with some admissible conditions. The Seifert surface plays role of this properly embedded surface, which decomposes $SHI(\widehat{Y}(\widehat{K}), \gamma)$ for any suture γ. Then we are able to identify some direct summands of the decomposition with $I^{\sharp}(\widehat{Y})$. Explicitly, we can construct the decomposition by the following proposition.

Proposition 1.16. *Suppose Y is a closed 3-manifold and $K \subset Y$ is a null-homologous knot. Suppose $\widehat{Y}$ is obtained from Y by performing the q/p-surgery along K with $q > 0$. Then there is a set $\mathcal{G}$ of sutures on the boundary of the knot complement $Y(K)$, determined by the integer q and the genus $g(K)$ of K so that the followings hold.*

(1) *For any suture $\gamma \in \mathcal{G}$, sutured instanton Floer homology carries a $\mathbb{Z}$ grading, that is,*

$$SHI(-Y(K), \gamma) = \bigoplus_{i \in \mathbb{Z}} SHI(-Y(K), \gamma, i).$$

(2) *For any suture $\gamma \in \mathcal{G}$, there is an integer i_γ so that there is an isomorphism*

$$f_{K,\gamma} : I^{\sharp}(-\widehat{Y}) \xrightarrow{\cong} \bigoplus_{j=i_\gamma}^{i_\gamma+q-1} SHI(-Y(K), \gamma, j).$$

(3) *For any two sutures $\gamma_1, \gamma_2 \in \mathcal{G}$, there is an isomorphism with respect to gradings* (that is, *sending the grading i_{γ_1} to i_{γ_2} and so on)*

$$g_{K,\gamma_1,\gamma_2} : \bigoplus_{j=i_{\gamma_1}}^{i_{\gamma_1}+q-1} SHI(-Y(K), \gamma_1, j) \xrightarrow{\cong} \bigoplus_{j=i_{\gamma_2}}^{i_{\gamma_2}+q-1} SHI(-Y(K), \gamma_2, j).$$

(4) *For any two sutures $\gamma_1, \gamma_2 \in \mathcal{G}$, there is a commutative diagram*

$$\begin{tikzcd}
\bigoplus_{j=i_{\gamma_1}}^{i_{\gamma_1}+q-1} SHI(-Y(K),\gamma_1,j) \arrow[rr, "g_{K,\gamma_1,\gamma_2}"] & & \bigoplus_{j=i_{\gamma_2}}^{i_{\gamma_2}+q-1} SHI(-Y(K),\gamma_2,j) \\
& I^{\sharp}(-\widehat{Y}) \arrow[ul, "f_{K,\gamma_1}"] \arrow[ur, "f_{K,\gamma_2}"'] &
\end{tikzcd}$$

Remark 1.17. Roughly speaking, in Proposition 1.16, the suture γ in the set $\mathcal{G}$ consists of two parallel copies of simple closed curves with some large slope. The isomorphism f_K, γ is constructed from cobordism maps associated to surgeries on curves in the interior of $Y(K)$. The isomorphism g_{K,γ_1,γ_2} is constructed from bypass maps. See the end of Subsection 4.1 for a sketch of the proof.

Remark 1.18. Theorem 1.12 is straightforward from term (2) of Proposition 1.16. However, there are many different choices of the grading i_γ. If we regard the decomposition in Theorem 1.12 as a $\mathbb{Z}_q$-grading on $I^{\sharp}(\widehat{Y})$, then different choices of i_γ lead to shifts of the $\mathbb{Z}_q$-grading. So we get a relative $\mathbb{Z}_q$-grading but it is straightforward to upgrade it to a canonical $\mathbb{Z}_q$-grading by fixing the choice of i_γ in term (2) of Proposition 1.16. In Heegaard Floer theory, if $\widehat{Y}$ is obtained from the q-surgery on a knot K in an integral homology sphere, then there is a canonical way to identify the set of spin^c structures on $\widehat{Y}$ with $\mathbb{Z}_q$. In our setup for instanton theory, with some efforts, one could fix a suitable i_γ so that $I^{\sharp}(\widehat{Y}, i)$ indeed corresponds to $\widehat{HF}(\widehat{Y}, [i])$. Here $\widehat{HF}(\widehat{Y}, [i])$ is the hat version of Heegaard Floer homology, of the 3-manifold $\widehat{Y}$ and the spin^c structure on $\widehat{Y}$ identified with $[i] \in \mathbb{Z}_q$.

We would like to make a remark on the developments up to date in answering those questions: In [2], Baldwin together with the authors proved a more general inequality that bounds the dimension of instanton Floer homology from above by the number of generators of any Heegaard Floer chain complex of the same manifold. In [44, 45], based on the techniques developed in the current paper, the authors fully answer the third question and relate the Euler characteristic of instanton Floer homology with the Turaev torsion of 3-manifolds. In [43], based on the answer to the second question in this paper, the authors develop a large surgery formula for instanton Floer homology of Dehn surgeries along knots.

Conventions. If not mentioned, homology groups and cohomology groups are with $\mathbb{Z}$ coefficients. (Sutured) Heegaard Floer homology is with $\mathbb{Z}$ coefficient, while (sutured) instanton Floer homology is with $\mathbb{C}$ coefficient. We write $\mathbb{Z}_n$ for $\mathbb{Z}/n\mathbb{Z}$.

If it is not mentioned, all manifolds are smooth, connected, and oriented. Suppose M is an oriented manifold. Let $-M$ denote the same manifold with the reverse orientation, called the *mirror manifold* of M. If not mentioned, we do not consider orientations of knots. Suppose K is a knot in a 3-manifold M. Then $(-M, K)$ is the *mirror knot* in the mirror manifold. In S^3, the mirror knot is also denoted by $\bar{K}$.

For a manifold M, let $\mathrm{int}(M)$ denote its interior. For a submanifold A in a manifold Y, let $N(A)$ denote the tubular neighborhood. The knot complement of K in Y is denoted by $Y(K) = Y \backslash \mathrm{int}(N(K))$.

For a simple closed curve on a surface, we do not distinguish between its homology class and itself. The algebraic intersection number of two curves α and β on a surface is denoted by $\alpha \cdot \beta$, while the number of intersection points between α and β is denoted by $|\alpha \cap \beta|$. A basis (m, l) of

$H_1(T^2;\mathbb{Z})$ satisfies $m \cdot l = -1$. The *surgery* means the Dehn surgery and the slope q/p in the basis (m,l) corresponds to the curve $qm + pl$.

A *rational homology sphere* is a closed 3-manifold whose homology groups with rational coefficients are isomorphic to those of S^3. An *integral homology sphere* is defined similarly. A knot $K \subset Y$ is called *null-homologous* if it represents the trivial homology class in $H_1(Y;\mathbb{Z})$, while it is called *rationally null-homologous* if it represents the trivial homology class in $H_1(Y;\mathbb{Q})$.

For $r \in \mathbb{R}$, let $\lceil r \rceil$ and $\lfloor r \rfloor$ be the minimal integer and the maximal integer satisfying $\lceil r \rceil \geqslant r$ and $\lfloor r \rfloor \leqslant r$, respectively. An argument holds for *large enough* n if there exists a fixed $N \in \mathbb{Z}$ so that the argument holds for any integer $n > N$. An argument holds for *small enough* n if there exists a fixed $N \in \mathbb{Z}$ so that the argument holds for any integer $n < N$.

Organization. The paper is organized as follows.

In Section 2, we review some backgrounds, including instanton Floer homology of closed manifolds (Subsection 2.1), Heegaard Floer homology and instanton Floer homology of balanced sutured manifolds (Subsection 2.2), and bypass attachments on balanced sutured manifolds (Subsection 2.3).

In Section 3, we construct the sutured handlebody $(-H,-\gamma)$ for Theorem 1.2 in Subsection 3.1, and prove a generalization of Proposition 1.4 in Subsection 3.2, which leads to the proofs of Theorem 1.2 and Proposition 1.5. Then we deal with $(1,1)$-knots and prove Theorem 1.6 in Subsection 3.3 and Theorem 1.10 in Subsection 3.4.

In Section 4, we state basic setups and sketch the proofs of Proposition 1.16 and Theorem 1.12 in Subsection 4.1, leaving the essential parts to the next two subsections. Then we prove Proposition 1.15 in Subsection 4.5.

Finally, in Section 5, we discuss some possible future directions.

2 | INSTANTON FLOER HOMOLOGY AND BALANCED SUTURED MANIFOLDS

2.1 | Instanton Floer homology

Definition 2.1. Suppose Y is a closed 3-manifold and ω is a closed 1-submanifold in Y. Suppose further that there is a closed oriented surface $\Sigma \subset Y$ of genus at least one such that the algebraic intersection number $\omega \cdot \Sigma$ is odd. Then the pair (Y,ω) is called an *admissible pair*.

For admissible pairs, Floer constructed a homology group from $SO(3)$ connections.

Theorem 2.2 [13]. *Suppose (Y,ω) is an admissible pair. Then there is a finite-dimensional complex vector space $I^{\omega}(Y)$ called the* instanton Floer homology *of (Y,ω).*

Suppose (Y,ω) and (Y',ω') are two admissible pairs. Suppose W is a cobordism from Y to Y', that is, *$\partial W = -Y \sqcup Y'$, and suppose $\nu \subset W$ is a 2-submanifold with $\partial \nu = (-\omega) \sqcup \omega'$. Then there exists a complex-linear map*

$$I(W,\nu) : I^{\omega}(Y) \to I^{\omega'}(Y'),$$

called the cobordism map *associated to (W,ν).*

Remark 2.3. For a fixed 3-manifold Y, $I^\omega(Y)$ only depends on the class of ω in $H_1(Y;\mathbb{Z}_2)$.

For an admissible pair (Y,ω), any homology class $\alpha \in H_*(Y)$ induces a complex-linear action on the instanton Floer homology:

$$\mu(\alpha) : I^\omega(Y) \to I^\omega(Y).$$

For any two homology classes $\alpha_1, \alpha_2 \in H_*(Y)$, we have

$$\mu(\alpha_1+\alpha_2) = \mu(\alpha_1)+\mu(\alpha_2) \text{ and } \mu(\alpha_1)\mu(\alpha_2) = (-1)^{\deg(\alpha_1)\deg(\alpha_2)}\mu(\alpha_2)\mu(\alpha_1).$$

If $b_2(Y) > 0$, we can pick a basis $\beta_1,\dots,\beta_n$ of $H_2(Y;\mathbb{Q})$ and consider the simultaneous generalized eigenspaces of all the actions $\mu(\beta_1),\dots,\mu(\beta_n)$. The simultaneous eigenvalues, as a tuple $(\lambda_1,\dots,\lambda_n)$, can be viewed as a linear map from $H_2(Y;\mathbb{Q})$ to $\mathbb{Q}$. This linear map is the analog of the evaluation of the first Chern classes of spin^c structures in Heegaard Floer homology. We have the following definition.

Definition 2.4 [34, Definition 7.3]. Suppose (Y,ω) is an admissible pair, R is a closed surface of genus at least one, and $\omega \cdot R$ is odd. Let $I^\omega(Y|R)$ be the $(2g(R)-2,2)$-generalized eigenspaces of the pair of actions $((\mu(R),\mu(\mathrm{pt}))$ on $I^\omega(Y)$, where pt is any fixed basepoint on Y.

However, if $\alpha = 0 \in H_2(Y;\mathbb{Q})$, then $\mu(\alpha) = 0$. Hence for a rational homology sphere Y, all μ-actions associated to second homology classes are trivial, and we cannot obtain an effective decomposition from μ.

Suppose M is a compact 3-manifold with torus boundary. Suppose $\omega \subset M$ is a closed 1-submanifold such that there exists a closed surface Σ of genus at least one with $\omega \cdot \Sigma$ odd. Let $i : \partial M \to M$ be the inclusion, and let

$$i_* : H_1(\partial M;\mathbb{Q}) \to H_1(M;\mathbb{Q}) \tag{2.1}$$

be the induced map on homology. Let $\gamma_1,\gamma_2,\gamma_3$ be three simple closed curves on ∂M with

$$\gamma_1\cdot\gamma_2 = \gamma_2\cdot\gamma_3 = \gamma_3\cdot\gamma_1 = -1.$$

For $i \in \{1,2,3\}$, let Y_i be the closed 3-manifold obtained by Dehn filling along γ_i:

$$Y_i = M \bigcup_{\gamma_i = \{1\}\times\partial D^2} S^1 \times D^2.$$

Then clearly for $i \in \{1,2,3\}$, (Y_i,ω) are all admissible pairs. Floer proved the following theorem.

Theorem 2.5 [13]. *There is an exact triangle*

$$\begin{tikzcd} I^\omega(Y_1) \arrow[rr, "f_1"] & & I^\omega(Y_2) \arrow[dl, "f_2"] \\ & I^\omega(Y_3) \arrow[ul, "f_3"] & \end{tikzcd} \tag{2.2}$$

Furthermore, all maps in the exact triangle (2.2) are induced by cobordism maps.

Remark 2.6. In original construction of Floer [13] or Scaduto [59, Section 2], one has to add some extra component to ω in one of Y_1, Y_2, and Y_3 to make the exact triangle hold. However, from [7, Section 2.2], Baldwin and Sivek showed that one could wisely choose some other 1-submanifold ω' to start with. After adding the extra component coming from the original exact triangle, we finally arrive at a 1-submanifold representing the same homology class as ω in $H_1(Y;\mathbb{Z}_2)$ for all three 3-manifolds.

2.2 | Balanced sutured manifolds

Definition 2.7 [26, Definition 2.2]. A *balanced sutured manifold* (M,γ) consists of a compact 3-manifold M with non-empty boundary together with a closed 1-submanifold γ on ∂M. Let $A(\gamma) = [-1,1]\times\gamma$ be an annular neighborhood of $\gamma \subset \partial M$ and let $R(\gamma) = \partial M\backslash \mathrm{int}(A(\gamma))$, such that they satisfy the following properties.

(1) Neither M nor $R(\gamma)$ has a closed component.
(2) If $\partial A(\gamma) = -\partial R(\gamma)$ is oriented in the same way as γ, then we require this orientation of $\partial R(\gamma)$ induces the orientation on $R(\gamma)$, which is called the *canonical orientation*.
(3) Let $R_+(\gamma)$ be the part of $R(\gamma)$ for which the canonical orientation coincides with the induced orientation on ∂M from M, and let $R_-(\gamma) = R(\gamma)\backslash R_+(\gamma)$. We require that $\chi(R_+(\gamma)) = \chi(R_-(\gamma))$. If γ is clear in the contents, we simply write $R_\pm = R_\pm(\gamma)$, respectively.

For a balanced sutured manifold (M,γ), Juhász constructed sutured (Heegaard) Floer homology, and Kronheimer and Mrowka constructed sutured instanton Floer homology.

Definition 2.8 [26, Definition 2.11]. A *balanced diagram* is a triple (Σ,α,β) such that the followings hold.

(1) Σ is a compact surface with boundary.
(2) $\alpha = \{\alpha_1,\dots,\alpha_n\}$ and $\beta = \{\beta_1,\dots,\beta_n\}$ are two sets of pair-wise disjoint simple closed curves in the interior of Σ. We do not distinguish between the set and the union of curves.
(3) The maps $\pi_0(\partial\Sigma)\to\pi_0(\Sigma\backslash\alpha)$ and $\pi_0(\partial\Sigma)\to\pi_0(\Sigma\backslash\beta)$ are surjective.

Let M be the 3-manifold obtained from $\Sigma\times[-1,1]$ by attaching 3-dimensional 2-handles along $\alpha_i\times\{-1\}$ and $\beta_i\times\{1\}$ for each integer $i\in[1,n]$ and let $\gamma = \partial\Sigma\times\{0\}$. A balanced diagram (Σ,α,β) is called *compatible* with a balanced sutured manifold (N,ν) if the sutured manifold (M,γ) is diffeomorphic to (N,ν). A sutured manifold (N,ν) is called a *product sutured manifold* if it is compatible with $(\Sigma,\emptyset,\emptyset)$ for some Σ.

Theorem 2.9 [26]. *Suppose (M,γ) is a balanced sutured manifold. Then there is a balanced diagram compatible with (M,γ). We can construct a $\mathbb{Z}$-module $SFH(M,\gamma)$ from compatible balanced diagrams, which is independent of the choices of balanced diagrams and called the sutured (Heegaard) Floer homology of (M,γ).*

Remark 2.10. Sutured Heegaard Floer homology generalizes Heegaard Floer homology [53] and knot Floer homology [52, 55]. Suppose Y is a closed 3-manifold and $K\subset Y$ is a knot. Let $Y(1)$ be obtained from Y by removing a 3-ball and let δ be a simple closed curve on $\partial Y(1)$. Let γ consist of two meridians of K with opposite orientations. Then there are canonical isomorphisms

$$SFH(Y(1),\delta)\cong\widehat{HF}(Y) \text{ and } SFH(Y(K),\gamma)\cong\widehat{HFK}(Y,K). \tag{6}$$

Theorem 2.11 [34, Section 7.4]. *For a balanced sutured manifold (M,γ), one can associate a triple (Y,R,ω), called a* closure *of (M,γ), such that the followings hold.*

(1) *Y is a closed 3-manifold such that M is a submanifold of Y.*
(2) *$R\subset Y$ is a closed surface of genus at least one such that $R_+(\gamma)$ is a submanifold of R and $R\cap\mathrm{int}(M)=\emptyset$.*
(3) *$\omega\subset Y$ is a simple closed curve such that it intersects R transversely at one point and $\omega\cap\mathrm{int}(M)=\emptyset$.*

Moreover, the isomorphism class of $I^\omega(Y|R)$ as in Definition 2.4 is independent of the choices of the triple (Y,R,ω) and is a topological invariant of (M,γ).

Definition 2.12. For a balanced sutured manifold (M,γ), the vector space $I^\omega(Y|R)$ for a closure (Y,R,ω) of (M,γ) is called the *sutured instanton Floer homology* of (M,γ). It is also denoted by $SHI(M,\gamma)$ to stress the independence of choices of closures as claimed in Theorem 2.11.

One important property of these two sutured Floer homologies is that they detect the tautness of balanced sutured manifolds.

Definition 2.13 [26, Definition 2.6]. A balanced sutured manifold (M,γ) is called *taut* if M is irreducible and $R(\gamma)$ is incompressible and Thurston norm-minimizing in $[R(\gamma)]\in H_2(M,\gamma)$.

Theorem 2.14 ([27, Theorem 1.4] for *SFH* and [34, Theorem 7.12] for *SHI*). *Suppose (M,γ) is a balanced sutured manifold with M irreducible. Then the followings are equivalent.*

- *(M,γ) is taut.*
- *$SFH(M,\gamma)\neq 0$.*
- *$SHI(M,\gamma)\neq 0$.*

Another important property is about the product manifold.

Theorem 2.15 [27, Corollary 9.6] for *SFH* and [34, Theorem 7.18] for *SHI*, both are based on [50, Theorem 1.1]. *Suppose (M,γ) is a balanced sutured manifold and a homology product* (cf. *[27, Definition 9.1]). Then the followings are equivalent.*

- *(M,γ) is a product sutured manifold* (cf. *Definition 2.8).*
- *$SFH(M,\gamma)\cong\mathbb{Z}$.*
- *$SHI(M,\gamma)\cong\mathbb{C}$.*

In Theorem 2.11, only the isomorphism class of SHI is well-defined. Later, Baldwin and Sivek improved the naturality of SHI, making it possible to discuss elements in SHI. Similar work is done by Juhász, Thurston and Zemke [28] for SFH over $\mathbb{Z}_2$, and Kutluhan, Sivek, and Taubes [39] for sutured ECH.

Theorem 2.16 [3, Section 9]. *For a balanced sutured manifold (M,γ) and any two closures (Y_1,R_1,ω_1) and (Y_2,R_2,ω_2) of (M,γ), there is an isomorphism*

$$\Phi_{1,2} : I^{\omega_1}(Y_1|R_1) \xrightarrow{\cong} I^{\omega_2}(Y_2|R_2),$$

which is well-defined up to multiplication by a unit, that is, *a non-zero complex number. Furthermore, the isomorphism Φ satisfies the following two conditions.*

(1) *If $(Y_1,R_1,\omega_1) = (Y_2,R_2,\omega_2)$, then*

$$\Phi_{1,2} \doteq \mathrm{id},$$

where $\doteq$ means equal up to multiplication by a unit.

(2) *If there is a third closure (Y_3,R_3,ω_3), then we have*

$$\Phi_{1,3} \doteq \Phi_{2,3}\circ\Phi_{1,2} : I^{\omega_1}(Y_1|R_1) \to I^{\omega_3}(Y_3|R_3).$$

From Theorem 2.16, for a balanced sutured manifold (M,γ), Baldwin and Sivek [3, Section 9] constructed a projective transitive system based on the vector spaces $I^{\omega}(Y|R)$ coming from different closures of (M,γ) and the canonical maps Φ between them. This projective transitive system is denoted by

$$\underline{\mathrm{SHI}}(M,\gamma).$$

We can regard it as a complex vector space well-defined up to multiplication by a unit. From now on, we will write $\underline{\mathrm{SHI}}(M,\gamma)$ for the sutured instanton Floer homology of (M,γ).

Definition 2.17 [34, Section 7.6]. Suppose Y is a closed 3-manifold and $K\subset Y$ is a knot. Let $(Y(1),\delta)$ and $(Y(K),\gamma)$ be balanced sutured manifolds defined as in Remark 2.10. The *framed instanton Floer homology* of Y is defined by

$$I^{\sharp}(Y) := I^{S^1}(Y\sharp(S^1\times T^2)|\{1\}\times T^2).$$

It is isomorphic to $\underline{\mathrm{SHI}}(Y(1),\delta)$ (*cf.* [34, Section 7.4]), so we do not distinguish them. The *instanton knot homology* of (Y,K) is defined by

$$KHI(Y,K) := \underline{\mathrm{SHI}}(Y(K),\gamma). \tag{7}$$

Remark 2.18. In [3], in order to make the definition of KHI independent of different choices of knot complements and the position of the meridional suture, Baldwin and Sivek also added a basepoint to the data. However, since in this paper we only care about the dimension of KHI, we overlook this ambiguity and omit the basepoint from our notation. Also, the definition of $\underline{\mathrm{SHI}}(Y(1),\delta)$ depends on a choice of basepoint. For the same reason, we omit the basepoint.

The surgery exact triangle in Theorem 2.5 can easily be generalized for KHI. Suppose M is a compact 3-manifold with torus boundary and $K\subset M$ is a knot. Let $\gamma_1,\gamma_2,\gamma_3$ be three simple closed curves on ∂M with

$$\gamma_1\cdot\gamma_2 = \gamma_2\cdot\gamma_3 = \gamma_3\cdot\gamma_1 = -1.$$

For $i \in \{1,2,3\}$, let Y_i be a closed 3-manifold obtained by Dehn filling along γ_i and let K_i be the knot induced by K:

$$(Y_i, K_i) = (M, K) \underset{\gamma_i = \{1\}\times \partial D^2}{\cup} S^1 \times D^2.$$

Theorem 2.19. *There is an exact triangle*

$$\begin{tikzcd}
KHI(Y_1, K_1) \arrow[rr, "f_1"] & & KHI(Y_2, K_2) \arrow[dl, "f_2"] \\
& KHI(Y_3, K_3) \arrow[ul, "f_3"] &
\end{tikzcd} \tag{2.3}$$

Furthermore, all maps in the exact triangle (2.3) are induced by cobordism maps between corresponding closures of balanced sutured manifolds induced by (Y_i, K_i).

Suppose (M, γ) is a balanced sutured manifold and $S \subset M$ is a properly embedded surface. We state results about the decomposition of $\underline{\mathrm{SHI}}(M, \gamma)$ associated to S.

Definition 2.20 [15, Definition 2.25]. Suppose (M, γ) is a balanced sutured manifold and $S \subset (M, \gamma)$ is a properly embedded surface in M. The surface S is called an *admissible surface* if the followings hold.

(1) Every boundary component of S intersects γ transversely and non-trivially.
(2) We require that $\frac{1}{2}|S \cap \gamma| - \chi(S)$ is an even integer.

For an admissible surface $S \subset (M, \gamma)$, there is a well-defined $\mathbb{Z}$ grading on $\underline{\mathrm{SHI}}(M, \gamma)$.

Theorem 2.21 [42]. *Suppose (M, γ) is a balanced sutured manifold and $S \subset (M, \gamma)$ is an admissible surface with $n = \frac{1}{2}|S \cap \gamma|$. Then there exists a closure (Y, R, ω) of (M, γ) so that S extends to a closed surface $\bar{S} \subset Y$ with $\chi(\bar{S}) = \chi(S) - n$. Let $SHI(M, \gamma, S, i)$ denote the $(2i)$-generalized eigenspace of $\mu(\bar{S})$ acting on $SHI(M, \gamma) = I^{\omega}(Y|R)$. Then $SHI(M, \gamma, S, i)$ is preserved by the canonical maps in Theorem 2.16. Thus, the vector space*

$$\underline{\mathrm{SHI}}(M, \gamma, S, i)$$

is well defined up to multiplication by a unit. Furthermore, the followings hold.

(1) *If $|i| > \frac{1}{2}(n - \chi(S))$, then $\underline{\mathrm{SHI}}(M, \gamma, S, i) = 0$.*
(2) *If there is a sutured manifold decomposition $(M, \gamma) \overset{S}{\rightsquigarrow} (M', \gamma')$ (cf. [14, Section 3] and [27, Definition 2.7]), then we have*

$$\underline{\mathrm{SHI}}(M, \gamma, S, \frac{1}{2}(n - \chi(S))) \cong \underline{\mathrm{SHI}}(M', \gamma').$$

(3) *For any $i \in \mathbb{Z}$, we have*

$$\underline{\mathrm{SHI}}(M, \gamma, S, i) = \underline{\mathrm{SHI}}(M, \gamma, -S, -i).$$

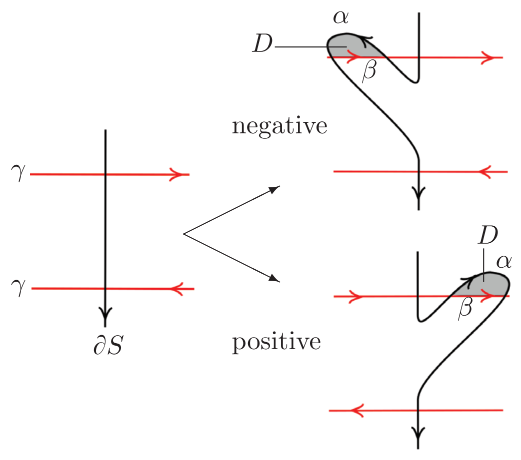

FIGURE 1 The positive and negative stabilizations of S

Remark 2.22. In [42], the grading was only constructed for an admissible surface with a connected boundary. When generalizing it to admissible surfaces with multiple boundary components, more choices arise in the construction of the grading. This new ambiguity was reduced to a combinatorial problem as discussed in [42, Section 3.3] and was then resolved in [30].

Remark 2.23. Term (1) of Theorem 2.21 comes from the adjunction inequality of instanton Floer homology (*cf.* [34, Proposition 7.5]). Term (2) of Theorem 2.21 is a restatement of [34, Proposition 7.11].

Suppose (M, γ) is a balanced sutured manifold, and $S \subset M$ is a properly embedded surface. If S is not admissible, then we isotop S to make it admissible.

Definition 2.24. Suppose (M, γ) is a balanced sutured manifold, and S is a properly embedded surface. A *stabilization* of S is a surface S' obtained from S by isotopy in the following sense. This isotopy creates a new pair of intersection points:

$$\partial S' \cap \gamma = (\partial S \cap \gamma) \cup \{p_+, p_-\}.$$

We require that there are arcs $\alpha \subset \partial S'$ and $\beta \subset \gamma$, oriented in the same way as $\partial S'$ and γ, respectively, and the followings hold.

(1) $\partial\alpha = \partial\beta = \{p_+, p_-\}$.
(2) α and β cobound a disk D with $\mathrm{int}(D) \cap (\gamma \cup \partial S') = \emptyset$.

The stabilization is called *negative* if ∂D is the union of α and β as an oriented curve. It is called *positive* if $\partial D = (-\alpha) \cup \beta$. See Figure 1. We denote by $S^{\pm k}$ the surface obtained from S by performing k positive or negative stabilizations, respectively.

The following lemma is straightforward.

Lemma 2.25. *Suppose (M,γ) is a balanced sutured manifold, and S is a properly embedded surface. Suppose S^+ and S^- are obtained from S by performing a positive and a negative stabilization, respectively. Then we have the following.*

(1) *If we decompose (M,γ) along S or S^+* (cf. *[14, Section 3] and [27, Definition 2.7]), then the resulting two balanced sutured manifolds are diffeomorphic.*
(2) *If we decompose (M,γ) along S^-, then the resulting balanced sutured manifold (M',γ') is not taut, as $R_{\pm}(\gamma')$ both become compressible.*

Remark 2.26. The definition of stabilizations of a surface depends on the orientations of the suture and the surface. If we reverse the orientation of the suture or the surface, then positive and negative stabilizations switch between each other.

The following theorem relates the gradings associated to different stabilizations of the same surface.

Theorem 2.27 [42, Proposition 4.3] and [62, Proposition 4.17]. *Suppose (M,γ) is a balanced sutured manifold and S is a properly embedded surface in M that intersects γ transversely. Suppose all the stabilizations mentioned below are performed on a distinguished boundary component of S. Then, for any $p,k,l\in\mathbb{Z}$ such that the stabilized surfaces S^p and S^{p+2k} are both admissible, we have*

$$\underline{\mathrm{SHI}}(M,\gamma,S^p,l)=\underline{\mathrm{SHI}}(M,\gamma,S^{p+2k},l+k).$$

Note that S^p is a stabilization of S as introduced in Definition 2.24, and, in particular, $S^0=S$.

Remark 2.28. The original form of Theorem 2.27 in [42] was stated for a Seifert surface in the case of a knot complement. However, it is straightforward to generalize the proof to the case of a general admissible surface in a general balanced sutured manifold, given the condition that the decompositions along S and $-S$ are both taut. This extra condition on taut decompositions was then dropped due to the work in [62].

2.3 | Bypass attachments

In this subsection, we review bypass maps for sutured instanton homology.

Definition 2.29 [22, Section 3.4]. Suppose (M,γ) is a balanced sutured manifold. An arc $\alpha\subset\partial M$ is called a *bypass arc* if the arc intersects the suture γ transversely at three points, including two endpoints.

For a bypass arc α, let P_0, P_1, and P_2 be its three intersection points with γ, ordered by any orientation of α. For $i=0,1,2$, let γ_i be the component of γ containing P_i. If $\gamma_0=\gamma_1\neq\gamma_2$ or $\gamma_1=\gamma_2\neq\gamma_0$, then α is called a *wave bypass*. If $\gamma_0=\gamma_2\neq\gamma_1$, then α is called an *anti-wave bypass*.

Remark 2.30. The names of wave and anti-wave follow from [18, Section 7], where waves and anti-waves are arcs whose endpoints are on the same curve. For an anti-wave bypass α, after removing the component of γ that only contains one intersection point, the arc α becomes a wave or an anti-wave. See Definition 3.35.

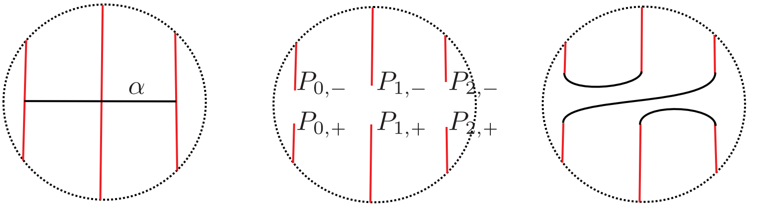

FIGURE 2 The bypass arc and the bypass attachment

Given a bypass arc α on a balanced sutured manifold (M, γ_1), we can change the suture locally as follows. Let $D \subset \partial M$ be a neighborhood of the arc $\alpha \subset \partial M$, which is a disk intersecting γ in three arcs. There are six endpoints after removing $\gamma \cap D$ from γ, labeled as follows. Suppose $P_{i,-}$ and $P_{i,+}$ are two endpoints corresponding to P_i, where the sign is chosen so that the oriented arc α, together with the arc-component of $\gamma \cap D$ from $P_{i,-}$ to $P_{i,+}$, gives an oriented framing of ∂M. Then we connect these six endpoints by the 'left-handed-principle': in D, a new suture γ_2 is obtained by connecting $P_{0,-}$ to $P_{1,-}$, connecting $P_{2,-}$ to $P_{0,+}$, and connecting $P_{2,+}$ to $P_{1,+}$. See Figure 2 for an example of a bypass arc and the corresponding new suture.

Proposition 2.31 [23, Section 2.3]. *Suppose (M, γ) is a balanced sutured manifold and α is a bypass arc. Suppose γ_2 is the new suture described as above. Then (M, γ_2) is still a balanced sutured manifold.*

If α is a wave bypass, the suture γ_2 is obtained from γ_1 via a 'mystery move' (cf. *[23, Figure 8]). If α is an anti-wave bypass, the suture γ_2 is obtained from γ_1 via a positive Dehn twist on ∂M. In both cases, the numbers of components of γ_1 and γ_2 are the same.*

Definition 2.32. The change from (M, γ_1) to (M, γ_2) is called a *bypass attachment* along α.

Remark 2.33. The definition of a bypass attachment is due to [22, Section 3.4]. Originally, a bypass attachment is a thickened half-disk attached to a contact 3-manifold M along an arc $\alpha \subset \partial M$, which carries some particular contact structure. The dividing set on the boundary ∂M can be thought of as equivalent to the suture. After the half-disk-attachment, the dividing set, or the suture, is changed in the way described in Definition 2.32. For our purpose, we do not require the balanced suture manifold (M, γ_1) to carry a contact structure, while we can still perform an abstract bypass attachment by modifying the suture in a neighborhood of α.

A bypass attachment induces a map

$$\psi_1 : \underline{\mathrm{SHI}}(-M, -\gamma_1) \to \underline{\mathrm{SHI}}(-M, -\gamma_2).$$

This map can be explained in the following two ways.

Contact handle decomposition. By Ozbagci [51, Section 3], the half-disk-attachment can be decomposed into two contact handle attachments. First, one can attach a contact 1-handle along two endpoints of the bypass arc α. Then one can attach a contact 2-handle along a circle that is the union of α and an arc on the attached contact 1-handle. Topologically, the 1-handle and the 2-handle form a canceling pair, so the diffeomorphism type of the 3-manifold does not change. However, the contact structure is changed, and the suture γ_1 is replaced by γ_2. In [4, Section 5], Baldwin and Sivek constructed contact handle attaching maps for $\underline{\mathrm{SHI}}$. Following Ozbagci's idea, they defined the map

$$\psi_1 : \underline{\mathrm{SHI}}(-M, -\gamma_1) \to \underline{\mathrm{SHI}}(-M, -\gamma_2)$$

to be the composition of contact handle attaching maps corresponding to the contact 1-handle and the 2-handle attaching.

Contact gluing maps. The half-disk attachment can be reinterpreted as follows. We start with (M, γ_1). Then pick $[1,2] \times \partial M$ to be a collar of the boundary, carrying a particular contact structure ξ specified by the bypass attachment, so that the boundary $\{1,2\} \times \partial M$ is convex and the dividing set is $(-\gamma_1) \sqcup \gamma_2$, with $\gamma_i \subset \{i\} \times \partial M$ for $i \in \{1,2\}$. Then we glue $[1,2] \times \partial M$ to M by the identification $\{1\} \times \partial M = \partial M$. The new 3-manifold is diffeomorphic to M, while the suture γ_1 is replaced by γ_2. In [24], Honda, Kazez, and Matić defined a gluing map for SFH

$$\Phi_\xi : SFH(-M, -\gamma_1) \to SFH(-M, -\gamma_2),$$

which was then re-visited by Juhász and Zemke [29]. Later, the first author [41] defined a similar gluing map for $\underline{\mathrm{SHI}}$, and we can define

$$\psi_1 = \Phi_\xi : \underline{\mathrm{SHI}}(-M, -\gamma_1) \to \underline{\mathrm{SHI}}(-M, -\gamma_2).$$

These two points of view are equivalent due to [41, Section 4]. We have some useful corollaries.

Lemma 2.34. *Suppose (M, γ) is a balanced sutured manifold and $\alpha, \beta \subset \partial M$ are two bypass arcs with $\alpha \cap \beta = \emptyset$. Let ψ_α and ψ_β be the bypass maps associated to α and β, respectively. Let (M, γ') be the resulting balanced sutured manifold after bypass attachments along both α and β. Then we have*

$$\psi_\alpha \circ \psi_\beta = \psi_\beta \circ \psi_\alpha : \underline{\mathrm{SHI}}(-M, -\gamma) \to \underline{\mathrm{SHI}}(-M, -\gamma').$$

Proof. Consider bypasses as the compositions of contact handle attachments. Since $\alpha \cap \beta = \emptyset$, the contact handles associated to α and β are attached to disjoint regions on ∂M. Then it follows immediately that the corresponding contact handle attaching maps commute with each other. □

Lemma 2.35. *Suppose (M, γ) is a balanced sutured manifold and $\alpha_0, \alpha_1 \subset \partial M$ are two bypass arcs. Suppose further that these two arcs are isotopic as bypass arcs,* that is, *there is a smooth family α_t of bypass arcs for $t \in [0,1]$. Then α_1 and α_2 lead to isotopic balanced sutured manifold (M, γ'), and the bypass maps ψ_{α_1} and ψ_{α_2} are the same:*

$$\psi_{\alpha_1} = \psi_{\alpha_2} : \underline{\mathrm{SHI}}(-M, -\gamma) \to \underline{\mathrm{SHI}}(-M, -\gamma').$$

Proof. It follows from the contact handle decomposition interpretation of the bypass attachments. □

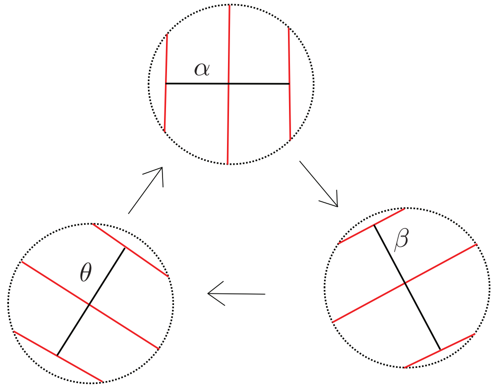

FIGURE 3 The bypass triangle

Remark 2.36. On the level of contact geometry, Honda has already proved Lemma 2.34 and Lemma 2.35 in [22]. Thus, these two lemmas can also be proved by combining Honda's results with the functoriality of gluing maps Φ_ξ in [41].

An important properties of bypass maps is the bypass exact triangle. Suppose (M,γ_1) is a balanced sutured manifold, and α is a bypass arc. Suppose D is a neighborhood of $\alpha\subset\partial M$ and (M,γ_2) is obtained by the bypass attachment along α. As shown in Figure 3, after attaching a bypass along α, there is an obvious bypass arc $\beta\subset D$. When we do the bypass attachment along β, we obtain the third balanced sutured manifold (M,γ_3). It still carries an obvious bypass arc θ. When we further do the bypass attachment along θ, we obtain (M,γ_1) again. Let ψ_1, ψ_2, and ψ_3 be the bypass maps associated to α, β, and θ, respectively. We have the following theorem.

Theorem 2.37 [5, Theorem 1.21]. *There exists an exact triangle*

$$\begin{tikzcd}
\underline{\mathrm{SHI}}(-M,-\gamma_1) \arrow[rr, "\psi_1"] & & \underline{\mathrm{SHI}}(-M,-\gamma_2) \arrow[dl, "\psi_2"] \\
& \underline{\mathrm{SHI}}(-M,-\gamma_3) \arrow[ul, "\psi_3"] &
\end{tikzcd}$$

3 | INSTANTON FLOER HOMOLOGY AND HEEGAARD DIAGRAMS

3.1 | Balanced sutured manifolds with tangles

Definition 3.1 [63, Definition 1.1]. Suppose (M,γ) is a balanced sutured manifold. A *tangle* $T\subset(M,\gamma)$ is a properly embedded 1-submanifold such that $T\cap A(\gamma)=\emptyset$. A tangle T is called *balanced* if

$$|T\cap R_+(\gamma)|=|T\cap R_-(\gamma)|.$$

A component a of T is called *vertical* if a is an arc from $R_+(\gamma)$ to $R_-(\gamma)$. A tangle T is called *vertical* if every component of T is vertical. Note that vertical tangles are balanced.

Suppose $T\subset(M,\gamma)$ is a vertical tangle, we construct a new balanced sutured manifold (M_T,γ_T), where $M_T=M\backslash N(T)$ and γ_T is the union of γ and one meridian for each component of T.

Theorem 3.2 [63]. *Suppose (M,γ) is a balanced sutured manifold and suppose $T \subset (M,\gamma)$ is a balanced tangle. Then there is a finite-dimensional complex vector space $SHI(M,\gamma,T)$, whose isomorphism class is a topological invariant of the triple (M,γ,T).*

We have the following theorem.

Theorem 3.3 [63, Lemma 7.10]. *For a vertical tangle $T \subset (M,\gamma)$, there is an isomorphism*

$$SHI(M,\gamma,T) \cong SHI(M_T,\gamma_T).$$

Then we introduce Heegaard diagrams of closed 3-manifolds and knots.

Definition 3.4. A (*genus g*) *diagram* is a triple (Σ,α,β), where:

(1) Σ is a closed surface of genus g;
(2) $\alpha = \{\alpha_1,\dots,\alpha_m\}$ and $\beta = \{\beta_1,\dots,\beta_n\}$ are two sets of pair-wise disjoint simple closed curves on Σ. We do not distinguish the set and the union of curves.

Let N_0 be the manifold obtained from $\Sigma\times[-1,1]$ by attaching 3-dimensional 2-handles along $\alpha_i\times\{-1\}$ and $\beta_j\times\{1\}$ for each integer $i\in[1,m]$ and each integer $j\in[1,n]$. Let N be the manifold obtained from N_0 by capping off spherical boundaries. A diagram (Σ,α,β) is called *compatible* with a 3-manifold M if $M\cong N$. In such case, we also write M is compatible with (Σ,α,β), or (Σ,α,β) is a diagram of M.

Definition 3.5. A (*genus g*) *Heegaard diagram* is a (genus g) diagram (Σ,α,β) satisfying the following conditions.

(1) $|\alpha| = |\beta| = g$, that is, there are g curves in either tuple.
(2) $\Sigma\backslash\alpha$ and $\Sigma\backslash\beta$ are connected.

Given a Heegaard diagram (Σ,α,β), the manifolds compatible with $(\Sigma,\alpha,\emptyset)$ and $(\Sigma,\emptyset,\beta)$ are called the *α-handlebody* and the *β-handlebody*, respectively.

Definition 3.6. A (*genus g*) *doubly-pointed Heegaard diagram* $(\Sigma,\alpha,\beta,z,w)$ is a (genus g) Heegaard diagram with two points z and w in $\Sigma\backslash\alpha\cup\beta$. Let $a\subset\Sigma\backslash\alpha$ and $b\subset\Sigma\backslash\beta$ be two arcs connecting z to w. Suppose a' and b' are obtained from a and b by pushing them into α-handlebody and β-handlebody, respectively. A doubly pointed Heegaard diagram $(\Sigma,\alpha,\beta,z,w)$ is called *compatible* with a knot K in a closed 3-manifold Y if (Σ,α,β) is compatible with Y and the union $a'\cup b'$ is isotopic to K.

Definition 3.7. Suppose (Σ,α,β) is a Heegaard diagram of a closed 3-manifold Y. A knot $K\subset Y$ is called the *core knot* of β_i for some $\beta_i\subset\beta$ if it is constructed as follows. Let M be the manifold compatible with the diagram $(\Sigma,\alpha,\beta\backslash\beta_i)$. It has a torus boundary and β_i induces a simple closed curve β_i' on ∂M. Dehn filling M along $\beta_i'\subset\partial M$ gives Y. Let K be the image of $S^1\times 0\subset S^1\times D^2$ under the filling map, where $S^1\times D^2$ is the filling solid torus.

The following is a basic fact in 3-dimensional topology.

Proposition 3.8 [52, Section 2.2]. *For any closed 3-manifold Y and any knot $K \subset Y$, there is a doubly pointed Heegaard diagram compatible with (Y, K).*

In the rest of this subsection, we provide the construction of the balanced sutured handlebody (H, γ) used in Theorem 1.2.

Construction 3.9. Suppose Y is a closed 3-manifold and $K \subset Y$ is a knot. Suppose $(\Sigma, \alpha, \beta, z, w)$ is a genus $(g-1)$ doubly pointed Heegaard diagram compatible with (Y, K). Consider the manifold M obtained from $\Sigma \times [-1, 1]$ by attaching a 3-dimensional 1-handle along $\{z, w\} \times \{1\}$. Let Σ' be the component of ∂M with genus g. Let $\alpha_g \subset \Sigma'$ be the curve obtained by running from z to w and then back over the 1-handle. Let $\beta_g \subset \Sigma'$ be a small circle around z. Set

$$\alpha' = \alpha \times \{1\} \cup \{\alpha_g\} \text{ and } \beta' = \beta \times \{1\} \cup \{\beta_g\}.$$

Then $(\Sigma', \alpha', \beta')$ is a genus g Heegaard diagram compatible with Y. Since β_g is a meridian of K, the knot K is the core knot of β_g.

Construction 3.10. Suppose Y is a closed 3-manifold and $(\Sigma', \alpha', \beta' = \{\beta_1, \dots, \beta_g\})$ is a genus g Heegaard diagram compatible with Y. Let $Y(1)$ be obtained from Y by removing a 3-ball. The manifold $Y(1)$ can be obtained from the α'-handlebody by attaching 3-dimensional 2-handles along β_i for each integer $i \in [1, g]$. Note that a 3-dimensional 2-handle can be thought of as $[-1, 1] \times D^2$ attached along $[-1, 1] \times \partial D^2$. Let $\theta_i = [-1, 1] \times \{0\}$ be the co-core of the 2-handle attached along β_i. We have a properly embedded tangle in $Y(1)$:

$$T = \theta_1 \cup \cdots \cup \theta_g.$$

Pick a simple closed curve $\delta \subset \partial Y(1)$ such that for any i, two endpoints of θ_i lie on two different sides of δ. From the construction, the manifold $Y(1)_T = Y(1) \backslash N(T)$ is the α'-handlebody and the suture δ_T consists of all β_i curves and a curve β_{g+1} induced by δ, that is,

$$\delta_T = \beta_1 \cup \cdots \cup \beta_g \cup \beta_{g+1}.$$

Hence $R_+(\delta_T)$ and $R_-(\delta_T)$ can be obtained from $\Sigma \backslash \beta$ by cutting along β_g, which are both spheres with $(g+1)$ punctures.

Construction 3.11. Suppose Y is a closed 3-manifold and $K \subset Y$ is a knot. Suppose $(\Sigma, \alpha, \beta = \{\beta_1, \dots, \beta_{g-1}\}, z, w)$ is a genus $(g-1)$ doubly pointed Heegaard diagram of (Y, K). We apply Construction 3.9 to obtain a genus g Heegaard diagram $(\Sigma', \alpha', \beta' = \{\beta_1, \dots, \beta_g\})$ of Y, and then apply Construction 3.10 to obtain a balanced sutured handlebody

$$(H, \gamma) = (Y(1)_T, \delta_T = \beta_1 \cup \cdots \cup \beta_{g+1}).$$

Note that the diagram $(\Sigma', \alpha', \beta)$ is compatible with the knot complement $Y(K)$. Suppose β''_g and β''_{g+1} are curves on $\partial Y(K)$ induced by β_g and β_{g+1}, respectively. Since $\beta''_g \cap \beta''_{g+1} = \emptyset$ and $\partial Y(K) \cong T^2$, the curve β''_{g+1} is parallel to β''_g. Since β''_g is a meridian of K and $(Y(K), \beta''_g \cup \beta''_{g+1})$ is a balanced sutured manifold, the curve β''_{g+1} must be another meridian of K with the orientation opposite to that of β''_g.

We provide an explicit construction of the curve $\beta_{g+1} \subset \partial H$ in Construction 3.11.

Construction 3.12. Suppose $(\Sigma', \alpha', \beta' = \{\beta_1, \dots, \beta_g\})$ is a genus g Heegaard diagram compatible with a closed 3-manifold Y. Let H be the α'-handleboby. For any integer $i \in [1, g]$, let β_i be oriented arbitrarily and let $\beta'_i \subset \partial H$ be the curve obtained by pushing off β_i to the right with respect to the orientation. Suppose β'_i is oriented reversely. Let β_{g+1} be the curve obtained from β'_i by band sums with respect to orientations so that β_{g+1} is disjoint from $\beta_1, \dots, \beta_g$. Set

$$\gamma = \beta_1 \cup \cdots \cup \beta_{g+1}.$$

It is straightforward to check that (H, γ) is the one obtained in Construction 3.11.

We can also obtain the original 3-manifold Y from the sutured handlebody (H, γ) as follows.

Construction 3.13. Suppose H is a handlebody, and γ is a suture on ∂H such that $R_+(\gamma)$ and $R_-(\gamma)$ are both spheres with $(g+1)$ punctures. Let $\Sigma = \partial H$. Suppose Σ has genus g. Let $\alpha_1,\dots,\alpha_g$ be boundaries of g compressing disks $D_1, \dots, D_g$ so that $H \backslash (D_1 \cup \cdots \cup D_g)$ is a 3-ball. Since $R_+(\gamma)$ and $R_-(\gamma)$ are both spheres with $(g+1)$ punctures, the suture γ has $(g+1)$ components. We can take arbitrary g of them to form β. Then (Σ, α, β) is a Heegaard diagram. Let Y be a closed 3-manifold compatible with (Σ, α, β). Since different choices of such g curves from γ are related to each other by a finite sequence of handle slides, the manifold Y is well defined up to diffeomorphism.

Let δ be the remaining component of γ and let T be the union of co-cones of β_i curves as in Construction 3.10. It is straightforward to check that $(Y(1)_T, \delta_T) = (H, \gamma)$.

3.2 | A dimension inequality for tangles

In this subsection, we prove a generalization of Proposition 1.4.

Proposition 3.14. *Suppose (M, γ) is a balanced sutured manifold and T is a vertical tangle. Suppose $a \subset T$ is a component of T so that $[a] = 0 \in H_1(M, \partial M; \mathbb{Q})$. Let $T' = T \backslash a$. Then we have*

$$\dim_{\mathbb{C}} SHI(-M, -\gamma, T') \leqslant \dim_{\mathbb{C}} SHI(-M, -\gamma, T).$$

Proof. We prove this proposition in followings steps. By Theorem 3.3, it suffices to prove

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_T, -\gamma_T) \leqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T'}, -\gamma_{T'}),$$

where (M_T, γ_T) and $(M_{T'}, \gamma_{T'})$ are constructed as in Definition 3.1.

Step 1. We construct an auxiliary manifold M_{T_0} with a family of sutures Γ_n for $n \in \mathbb{N} \cup \{-, +\}$ on ∂M_{T_0}.

Since $[a] = 0 \in H_1(M, \partial M; \mathbb{Q})$, there exist $q, k \in \mathbb{Z}$ and arcs $b_1, \ldots, b_k \subset \partial M$ such that there exists a surface S in M with ∂S consisting of $b_1, \ldots, b_k$ and q copies of a. Suppose components of T are $a_1, a_2, \ldots, a_m$, with $a_1 = a$.

As in Definition 3.1, we form a new balanced sutured manifold (M_T, γ_T) as follows. Since T has m components, $\partial N(T)$ intersects each of $R_+(\gamma)$ and $R_-(\gamma)$ in m disks and intersects $\mathrm{int}(M)$ in m cylinders. Let $C_1, \ldots, C_m$ be the cylinders corresponding to $a_1, \ldots, a_m$, respectively. Let $M_T = M \backslash \mathrm{int}(N(T))$. For any integer $i \in [1, n]$, let $\gamma_i \subset C_i$ be a simple closed curve representing the generator of $H_1(C_i)$. Let a_i be oriented from $R_+(\gamma)$ to $R_-(\gamma)$. Then γ_i has an induced orientation from a_i. Let

$$\gamma_T = \gamma \cup \gamma_1 \cup \cdots \cup \gamma_m.$$

The surface S is modified into a properly embedded surface S_T in M_T as follows. First, for the part of ∂S consists of q copies of a, we isotop them to be on C_1. Then b_j for any integer $j \in [1, k]$ can be viewed as an arc on $\partial M_T \backslash C_1$. We can isotop S to make it intersect a_i transversely for any integer $i \in [2, m]$. Let S_T be obtained from S by removing disks in $N(T)$. Hence $S_T \cap C_1$ consists of q arcs, each intersecting γ_1 transversely at one point. For any integer $i \in [2, m]$, the intersection $S_T \cap C_i$ is a (possibly empty) collection of circles that are parallel to γ_i.

Note that b_1 is disjoint from all γ_i, but intersects the original suture γ. Since the arc $b_1 \subset \partial S_T \subset \partial M_T$ has one endpoint in $R_+(\gamma_T)$ and the other in $R_-(\gamma_T)$, the intersection number of γ and b_1 must be odd. Let b_1 be oriented from $R_+(\gamma_T)$ to $R_-(\gamma_T)$ and let the intersection points between b_1 and γ be $p_1, \ldots, p_l$ with l odd, ordered by the orientation of b_1. Let b_1' be a perturbation of b_1 such that b_1' and b_1 meet at endpoints. Suppose the intersection points between b_1' and γ are $q_1, \ldots, q_l$ so that q_i is near p_i for integer $i \in [1, l]$.

If $l = 1$, then (M_T, γ_T) is enough for the proof. If $l > 1$, we have to perform the following modification on (M_T, γ_T). We attach a contact 1-handle to (M, γ) along q_1 and q_{l-1} in the sense of Baldwin and Sivek [4, Section 3.2], or equivalently, attach a product 1-handle in the sense of Kronheimer and Mrowka [34, proof of Proposition 6.9]. They both proved that the balanced sutured manifolds before and after attaching such a 1-handle have exactly the same closure. Thus, after attaching the 1-handle, the sutured instanton Floer homology does not change. We still use (M, γ) and (M_T, γ_T) to denote sutured manifolds after attaching the 1-handle. Now we can choose an arc ζ satisfying the following conditions.

(1) Endpoints of ζ are contained in ∂C_1.
(2) The arc ζ intersects γ transversely at one point.
(3) The arc ζ is disjoint from S_T.

The arc ζ can be obtained by first going along b_1' until reaching q_1, then going along the 1-handle, and going back to b_1' at the point q_{l-1} and then keeping going along b_1'. Finally, we slightly perturb this arc to make it disjoint from S_T. See the middle subfigure of Figure 4.

Let a_0 be the arc obtained by pushing a neighborhood of q_l in ζ into the interior of M_T. Suppose the endpoints of a_0 are still in ζ and a_0 is disjoint from S_T. The arc a_0 is a vertical tangle in M_T and hence also a vertical tangle in the original manifold M. Let $T_0 = T \cup a_0$, and $T_0' = T' \cup a_0$. Let (M_{T_0}, γ_{T_0}) be obtained similarly as (M_T, γ_T). Since $M_{T_0} = M_T \backslash N(a_0)$, the cylinder C_0 and the suture γ_0 are defined similarly as C_i and γ_i for any integer $i \in [1, m]$. A sketch is shown in the left-subfigure of Figure 5. Let $\zeta_\pm$ be two parts of ζ contained in $R_\pm(\gamma_{T_0})$, respectively.

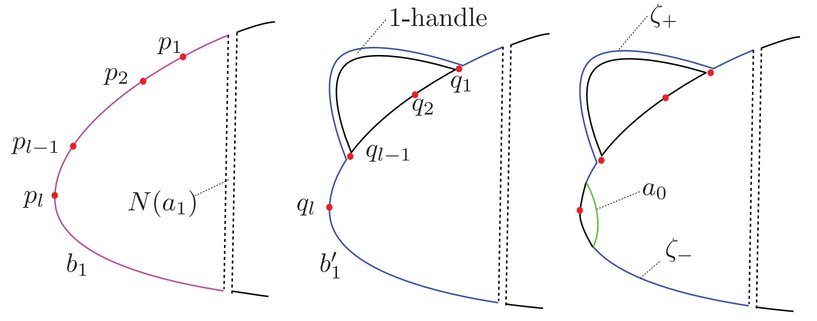

FIGURE 4 Left, the arc b_1; middle, the 1-handle on β_1' and the arc ζ; right, the arcs ζ_+, a_0, and ζ_-

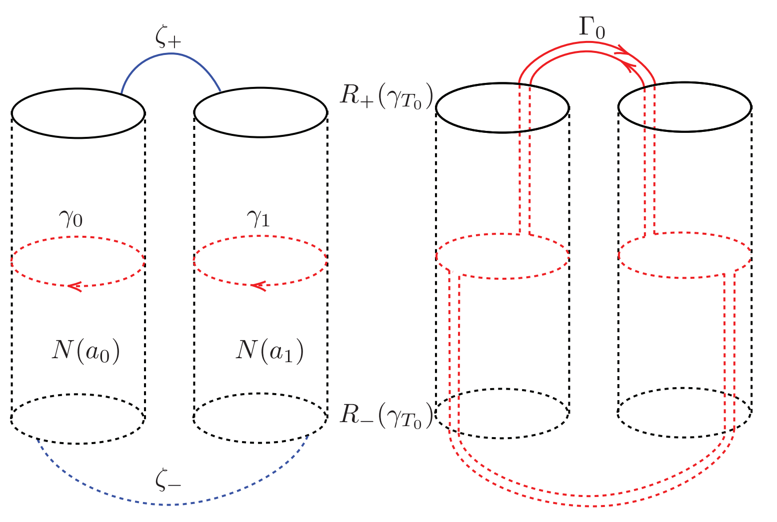

FIGURE 5 Left, a sketch of the balanced sutured manifold (M_{T_0}, γ_{T_0}) with the arcs ζ_+ and ζ_-; right, the suture Γ_0

Now, we pick a family of sutures Γ_n on M_{T_0} such that Γ_n is the suture obtained from γ_{T_0} by replacing γ_0 and γ_1 by other two curves. We describe the two new curves as follows. For Γ_0, the new curves are depicted in the right subfigure of Figure 5. Note that the part of the new curves in $R_\pm(\gamma_{T_0})$ consists of two parallel copies of $\zeta_\pm$, respectively. The suture Γ_n is obtained from Γ_0 by Dehn twists along $-\gamma_1$ for n times, as shown in the left subfigure of Figure 6.

There are two obvious bypass arcs in the left-subfigure of Figure 6, denoted by η_+ and η_-, respectively. By Theorem 2.37, these two bypass arcs induce two bypass exact triangles:

$$\begin{tikzcd}
\underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_{n-1}) \arrow[rr, "\psi^{n-1}_{+,n}"] & & \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n) \arrow[dl, "\psi^n_{+,+}"] \\
& \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_+) \arrow[ul, "\psi^+_{+,n-1}"] &
\end{tikzcd} \tag{3.1}$$

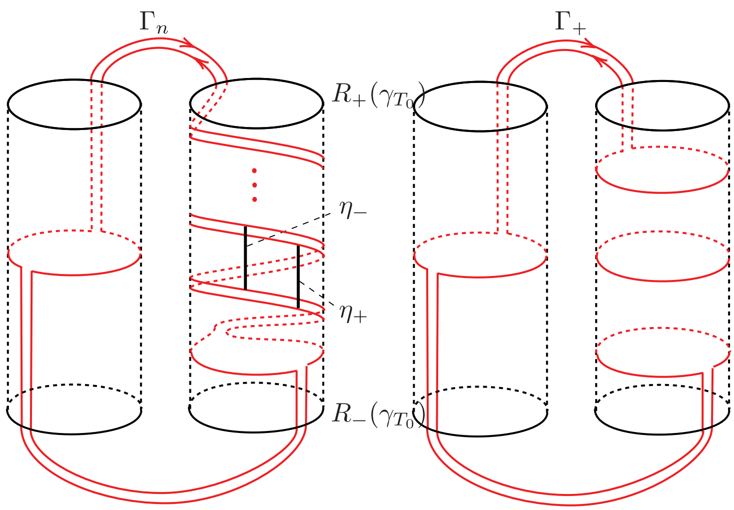

FIGURE 6 Left, the suture Γ_n, and the bypass arcs η_+ and η_-; right, the suture Γ_+

and

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_{n-1}) \arrow[rr, "\psi^{n-1}_{-,n}"] & & \underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_n) \arrow[dl, "\psi^{n}_{-,-}"] \\ & \underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_-) \arrow[ul, "\psi^{-}_{-,n-1}"] & \end{tikzcd} \tag{3.2}$$

respectively, where Γ_- is the same as the original suture γ_{T_0}, and Γ_+ is the suture as depicted in the right subfigure of Figure 6.

Note that the bypasses are attached to η_+ and η_- from the exterior of the 3-manifold M_{T_0}, though the point of view in Figure 6 is from the interior of the manifold. Hence readers have to take extra care when performing these bypass attachments.

Step 2. We use bypass maps and bypass triangles to derive a dimension inequality about M_{T_0}.

Recall we have a properly embedded surface $S_T \subset M_T$. Since $a_0 \cap S_T = \emptyset$, we can regard S_T as a properly embedded surface in M_{T_0}. Let S_T be oriented so that the orientation of ∂S_T coincides with that of $a\ (= a_1)$. Note that ζ_+ and ζ_- are both disjoint from S_T. We can perform stabilizations on S_T so that the followings hold.

(1) S_T is admissible with respect to the suture $\Gamma_-\ (= \gamma_{T_0})$.
(2) For any $n \in \mathbb{N} \cup \{-,+\}$, S_T has minimal possible number of intersection points with the part of the suture $\Gamma_n \backslash (\gamma \cup \gamma_2 \cup \cdots \cup \gamma_m)$.
(3) S_T is disjoint from $\zeta_+ \cup \zeta_-$.

The surface after stabilizations is still denoted by S_T. After obtaining the surface S_T satisfying the above three conditions, we further perform stabilizations as follows. For any $n \in \mathbb{N} \cup \{-,+\}$, if S_T is already admissible with respect to Γ_n, then let S_n be the surface S_T without any further change. If S_T is not admissible with respect to Γ_n, then we perform a negative stabilization on S_T within C_1 to make it admissible, and write S_n for the resulting surface. Equivalently, define a map

$$\tau : \mathbb{N} \cup \{-,+\} \mapsto \{0,-1\}$$

$$\tau(n) = \begin{cases} 0 & \text{If } |S_T \cap \Gamma_n \cap C_1| \text{ is odd,} \\ -1 & \text{If } |S_T \cap \Gamma_n \cap C_1| \text{ is even.} \end{cases}$$

Then we take $S_n = S_T^{\tau(n)}$, where the superscript follows Definition 2.24. Note that all the stabilizations are with respect to Γ_n rather than $-\Gamma_n$.

By Theorem 2.21, for any $n \in \mathbb{N} \cup \{-,+\}$, there is a closure (Y_n, R_n, ω_n) for (M_{T_0}, Γ_n) so that S_n extends to a closed surface $\bar{S}_n$. Let

$$i^n_{\max} = -\frac{1}{2}\chi(\bar{S}_n), \text{ and } i^n_{\min} = \frac{1}{2}\chi(\bar{S}_n) - \tau(n).$$

Lemma 3.15. *If $i > i^n_{max}$ or $i < i^n_{min}$, then* $\underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n, S_n, i) = 0$.

Proof. If $\tau(n) = 0$, then the lemma follows directly from term (1) of Theorem 2.21. If $\tau(n) = -1$, then term (1) of Theorem 2.21 only implies that for $i < i^n_{min} - 1$,

$$\underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n, S_n, i) = 0.$$

For the remaining case where $i = i^n_{min} - 1$, from term (3) of Theorem 2.21, we have

$$\underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n, S_n, i) = \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n, -S_n, -i) = \underline{\mathrm{SHI}}\left(-M_{T_0}, -\Gamma_n, -S_n, -\frac{\chi(\bar{S}_n)}{2}\right).$$

From the construction of S_n, we know that $-S_n$ is obtained from S_T by a negative stabilization with respect to the suture $-\Gamma_n$. Hence we can apply Lemma 2.25 and term (2) of Theorem 2.21 to obtain the vanishing result. □

Remark 3.16. A priori, we do not know if $\underline{\mathrm{SHI}}$ is non-vanishing at the gradings $i^n_{\max}$ and $i^n_{\min}$, though this does not make any difference in the proof of Proposition 3.14.

Next, we will derive a graded version of bypass exact triangles (3.1) and (3.2). To do so, we will discuss more about the surface S_n. Since ∂S_T contains q copies of a_1, for any $n \in \mathbb{N}$, we have

$$|S_T \cap \Gamma_+ \cap C_1| = q, \ |S_T \cap \Gamma_- \cap C_1| = 3q, \text{ and } |S_T \cap \Gamma_n \cap C_1| = (2n+1)q.$$

From Theorem 2.21, we know that for any $n \in \mathbb{N}$,

$$\chi(\bar{S}_-) = \chi(\bar{S}_+) - q + \tau(-) \text{ and } \chi(\bar{S}_n) = \chi(\bar{S}_+) - nq + \tau(n). \tag{3.3}$$

From Lemma 3.15, we have

$$\lim_{n\to+\infty} i^n_{\max} = +\infty \text{ and } \lim_{n\to+\infty} i^n_{\min} = -\infty. \tag{3.4}$$

To present the grading shifting property better, we make the following definition.

Definition 3.17. Suppose (M,γ) is a balanced sutured manifold and S is an admissible surface in (M,γ). For any $i, j \in \mathbb{Z}$, define

$$\underline{\mathrm{SHI}}(M,\gamma,S,i)[j] = \underline{\mathrm{SHI}}(M,\gamma,S,i-j).$$

Lemma 3.18. *For any $n \in \mathbb{N}$, we have two exact triangles*

$$\begin{tikzcd}
\underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_n,S_n)[i^{n+1}_{\min}-i^n_{\min}] \arrow[r, "\psi^n_{+,n+1}"] & \underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_{n+1},S_{n+1}) \arrow[dl, "\psi^{n+1}_{+,+}"] \\
\underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_+,S_+)[i^{n+1}_{\max}-i^+_{\max}] \arrow[u, "\psi^+_{+,n}"] &
\end{tikzcd}$$

and

$$\begin{tikzcd}
\underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_n,S_n)[i^{n+1}_{\max}-i^n_{\max}] \arrow[r, "\psi^n_{-,n+1}"] & \underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_{n+1},S_{n+1}) \arrow[dl, "\psi^{n+1}_{-,-}"] \\
\underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_-,S_-)[i^{n+1}_{\min}-i^-_{\min}] \arrow[u, "\psi^-_{-,n}"] &
\end{tikzcd}.$$

Furthermore, all maps in the above two exact triangles are grading preserving.

Proof. We only prove the grading shifting behavior of the map $\psi^{n+1}_{+,+}$ in the triangle (3.1), and the proof for any other map is similar. For simplicity, we also assume that $\tau(n) = \tau(+) = 0$ (Note that $\tau(+) = 0$ by definition), and other cases are similar. From Subsection 2.3, we know bypass maps are constructed via contact handle maps and ultimately via cobordisms maps associated to Dehn surgeries (*cf.* [4, Section 3]). It is obvious that the bypass arc is disjoint from ∂S_{n+1}. By construction of the grading in Theorem 2.21, this implies that the following map is grading preserving:

$$\psi^{n+1}_{+,+} : \underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_{n+1},S_{n+1}) \to \underline{\mathrm{SHI}}(-M_{T_0},-\Gamma_+,S_{n+1}).$$

From Figure 7, it is straightforward to check that S_{n+1} is obtained from S_+ by

$$\frac{1}{2}(|S_{n+1}\cap\Gamma_+\cap C_1| - |S_+\cap\Gamma_+\cap C_1|) = 2(i^{n+1}_{\max} - i^+_{\max})$$

negative stabilizations, with respect to the suture Γ_+. Hence they become positive stabilizations with respect to $-\Gamma_+$. Then the grading shift follows from Theorem 2.27. □

Remark 3.19. The statement of Lemma 3.18 can be illustrated in Figure 8, where sutured instanton homologies are denoted by the related sutures and maps are denoted by horizontal arrows.

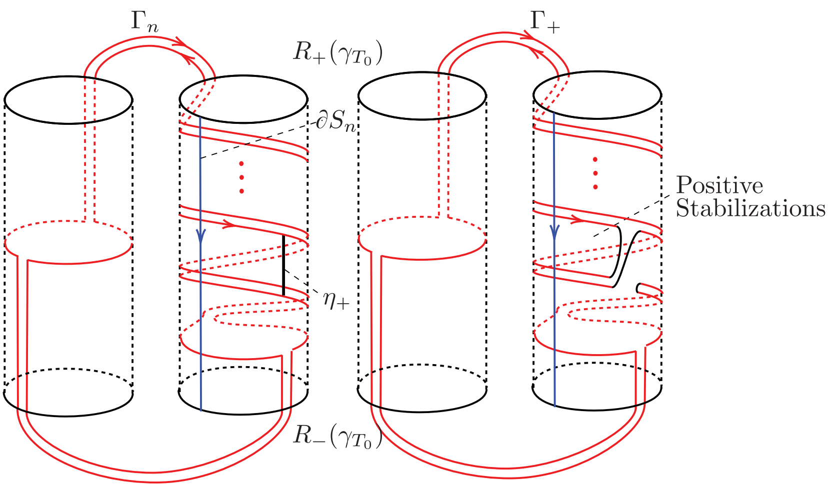

FIGURE 7 Left, the suture Γ_n, the surface S_n, and the bypass η_+; right, the suture Γ_+ after the bypass attachment along η_+

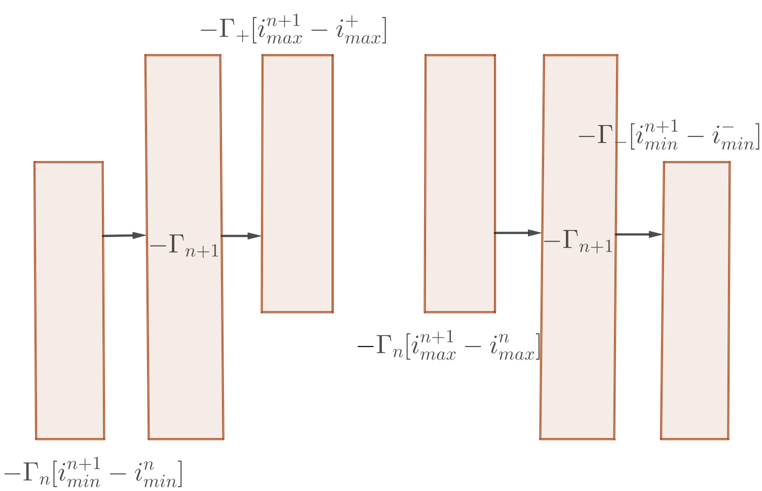

FIGURE 8 Illustration of Lemma 3.18

The heights of the blocks depend on $i_{\max} - i_{\min}$. This illustration is also useful for statements in Section 4.

Equipped with Lemma 3.18, we are able to prove the following lemma. For any $i \in \mathbb{Z}, n \in \mathbb{N}$, let $\psi^{n,i}_{\pm,n+1}$ be the restriction of $\psi^n_{\pm,n+1}$ on the ith grading associated to S_n.

Lemma 3.20. *The map*

$$\psi^{n,j}_{+,n+1} : \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n, S_n, j) \to \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_{n+1}, S_{n+1}, j - (i^n_{min} - i^{n+1}_{min}))$$

is an isomorphism if

$$j < i^{n+1}_{max} + (i^n_{min} - i^{n+1}_{min}) - (i^+_{max} - i^+_{min}).$$

Similarly, the map

$$\psi^{n,j}_{-,n+1}:\underline{SHI}(-M_{T_0},-\Gamma_n,S_n,j)\to\underline{SHI}(-M_{T_0},-\Gamma_{n+1},S_{n+1},j+(i^{n+1}_{max}-i^{n}_{max}))$$

is an isomorphism if

$$j>i^{n+1}_{min}-(i^{n+1}_{max}-i^{n}_{max})+(i^{-}_{max}-i^{-}_{min}).$$

Proof. We only prove the first statement. The proof of the second argument is similar. Suppose

$$i=j-(i^{n+1}_{\min}-i^{n+1}_{\min}).$$

Then we know that there is a map

$$\psi^{n+1,i}_{+,+}:\underline{SHI}(-M_{T_0},-\Gamma_{n+1},S_{n+1},i)\to\underline{SHI}(-M_{T_0},-\Gamma_{+},S_{+},i-i^{n+1}_{\max}+i^{+}_{\max}).$$

By assumption, we have

$$\begin{aligned}
i-i^{n+1}_{\max}+i^{+}_{\max}&=j-(i^{n}_{\min}-i^{n+1}_{\min})-i^{n+1}_{\max}+i^{+}_{\max}\\
&<i^{n+1}_{\max}+(i^{n}_{\min}-i^{n+1}_{\min})-(i^{+}_{\max}-i^{+}_{\min})-(i^{n}_{\min}-i^{n+1}_{\min})-i^{n+1}_{\max}+i^{+}_{\max}\\
&=i^{+}_{\mathrm{minmin}}.
\end{aligned}$$

Hence it follows from Lemma 3.15 that $\psi^{n+1,i}_{+,+}=0$. By Lemma 3.18, the map $\psi^{n,j}_{+,n+1}$ is surjective. The proof of injectivity is similar. □

Lemma 3.21. *For any $n\in\mathbb{N}$, there is an exact triangle*

$$\begin{tikzcd}
\underline{SHI}(-M_{T_0},-\Gamma_n) \arrow[rr] & & \underline{SHI}(-M_{T_0},-\Gamma_{n+1}) \arrow[dl, "F_{n+1}"] \\
& \underline{SHI}(-M_{T_0'},-\gamma_{T_0'}) \arrow[ul, "G_n"] &
\end{tikzcd} \tag{3.5}$$

Furthermore, we have two commutative diagrams related to $\psi^n_{+,n+1}$ and $\psi^n_{-,n+1}$, respectively

$$\begin{tikzcd}
\underline{SHI}(-M_{T_0},-\Gamma_n) \arrow[rr, "\psi^n_{\pm,n+1}"] & & \underline{SHI}(-M_{T_0},-\Gamma_{n+1}) \\
& \underline{SHI}(-M_{T_0'},-\gamma_{T_0'}) \arrow[ul, "G_n"] \arrow[ur, "G_{n+1}"'] &
\end{tikzcd}$$

Proof. Let γ_1' be the curve obtained by pushing γ_1 into the interior of M_{T_0}, with the framing from ∂M_{T_0}. The $(0,1,\infty)$-surgery triangle associated to γ_1' is the following.

$$\begin{tikzcd}
((-M_{T_0})_1,-\Gamma_{n+1}) \arrow[rr] & & ((-M_{T_0})_\infty,-\Gamma_{n+1}) \arrow[dl] \\
& ((-M_{T_0})_0,-\Gamma_{n+1}) \arrow[ul] &
\end{tikzcd}$$

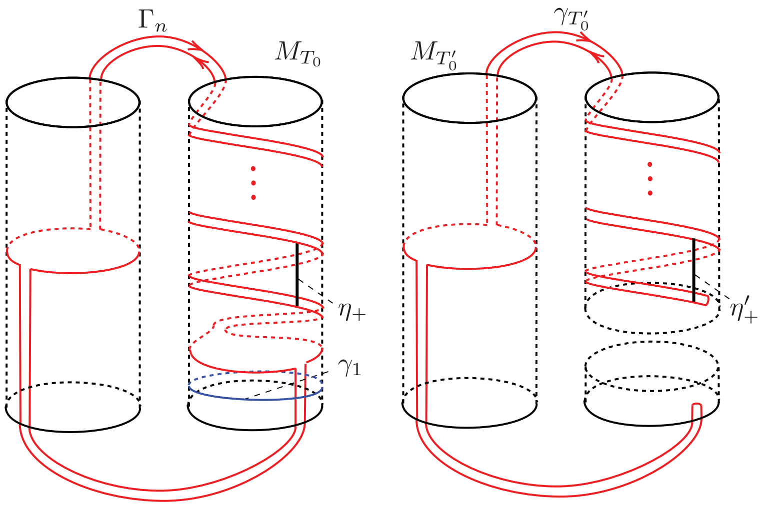

FIGURE 9 Left, the suture Γ_n, the meridian γ_1, and the bypass η_+; right, the balanced sutured manifold $(M_{T_0'}, \gamma_{T_0'})$ after attaching a contact 2-handle along γ_1

Since γ_1' is in the interior of M_{T_0}, the surgeries do not influence the procedure of constructing closures of balanced sutured manifolds. Hence from Theorem 2.5, we have an exact triangle

$$\begin{tikzcd}
\underline{\mathrm{SHI}}((-M_{T_0})_1, -\Gamma_{n+1}) \arrow[rr] & & \underline{\mathrm{SHI}}((-M_{T_0})_\infty, -\Gamma_{n+1}) \arrow[dl] \\
& \underline{\mathrm{SHI}}((-M_{T_0})_0, -\Gamma_{n+1}) \arrow[ul] &
\end{tikzcd}$$

The ∞-surgery does not change anything, so

$$((-M_{T_0})_\infty, -\Gamma_{n+1}) \cong (-M_{T_0}, -\Gamma_{n+1}).$$

The 1-surgery is equivalent to a Dehn twist along γ_1'. It does not change the underlying 3-manifold, while the suture Γ_{n+1} is replaced by Γ_n:

$$((-M_{T_0})_1, -\Gamma_{n+1}) \cong (-M_{T_0}, -\Gamma_n).$$

Finally, for the 0-surgery, from [4, Section 3.3], we know that on the level of closures, performing a 0-surgery is equivalent to attaching a contact 2-handle along $\gamma_1 \subset \partial M_{T_0}$. Attaching such a contact 2-handle changes (M_{T_0}, Γ_{n+1}) to $(M_{T_0'}, \gamma_{T_0'})$. Hence we obtain the desired exact triangle.

To prove two commutative diagrams, first note that the curve γ_1' is disjoint from the bypass arc η_+. As a result, the related maps commute with each other:

$$\psi^n_{+,n+1} \circ G_n = G_{n+1} \circ \psi_{\eta_+'},$$

where η_+' is the bypass arc as shown in the right subfigure of Figure 9. It is straightforward to check that the bypass along η_+' is a trivial bypass, and hence from [23, Section 2.3] it does not change the contact structure. From Subsection 2.3, the bypass maps can be reinterpreted as contact gluing maps, and by the functoriality of instanton contact gluing maps in [41], we know that the bypass map ψ_{η_+} corresponding to the trivial bypass η_+ is the identity map. Hence we conclude that

$$\psi^n_{+,n+1} \circ G_n = G_{n+1} \circ \psi_{\eta'_+} = G_{n+1} \circ \mathrm{id} = G_{n+1}.$$

The other commutative diagram involving $\psi^n_{-,n+1}$ can be proved similarly. □

Lemma 3.22. *For a large enough integer n, the map G_n in Lemma 3.21 is zero.*

Proof. We assume the lemma does not hold and derive a contradiction. For any n, there exists $x \in \underline{\mathrm{SHI}}(-M_{T'_0}, -\gamma_{T'_0})$ such that

$$y = G_n(x) \neq 0 \in \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n).$$

Suppose

$$y = \sum_{j \in \mathbb{Z}} y_j, \text{ where } y_j \in \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n, S_n, j),$$

$$j_{\max} = \max_{y_j \neq 0} j \text{ and } j_{\min} = \min_{y_j \neq 0} j.$$

By assumption, $j_{\max}$ and $j_{\min}$ both exist and $j_{\max} \geqslant j_{min}$. Suppose

$$z = G_{n+1}(x) \in \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_{n+1}),$$

and similarly

$$z = \sum_{j \in \mathbb{Z}} z_j, \text{ where } z_j \in \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_{n+1}, S_{n+1}, j).$$

From facts (3.3) and (3.4), we know that for a large enough integer n, we have

$$i^{n+1}_{\max} + (i^n_{\min} - i^{n+1}_{\min}) - (i^+_{\max} - i^+_{\min}) > i^{n+1}_{\min} - (i^{n+1}_{\max} - i^n_{\max}) + (i^-_{\max} - i^-_{\min}).$$

Hence at least one of the following two statements must be true.

(1) $j_{\max} > i^{n+1}_{\min} - (i^{n+1}_{\max} - i^n_{\max}) + (i^-_{\max} - i^-_{\min})$.
(2) $j_{\min} < i^{n+1}_{\max} + (i^n_{\min} - i^{n+1}_{\min}) - (i^+_{\max} - i^+_{\min})$.

We only work with the case where the first statement is true, and the other case is similar. From Lemma 3.21, we have

$$z = \psi^n_{+,n+1}(y) = \psi^n_{-,n+1}(y).$$

Suppose

$$i = j_{\max} = j_{\max} + (i^n_{\max} - i^{n+1}_{\max}),$$

and

$$j' = i + (i^n_{\min} - i^{n+1}_{\min}).$$

By Lemma 3.18, we have

$$\psi^{n,j'}_{+,n+1}(y_{j'}) = z_i = \psi^{n,j_{\max}}_{-,n+1}(y_{j_{\max}}).$$

Since $j' > j_{\max}$, we have $z_i = 0$. By Lemma 3.20, the first statement implies $\psi^{n,j_{\max}}_{-,n+1}$ is an isomorphism. Hence $y_{j_{\max}} = 0$, which contradicts the assumption of $j_{\max}$. □

Suppose n is large enough. By the exact triangle (3.5), the fact that G_n is zero implies

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T_0'}, -\gamma_{T_0'}) = \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_{n+1}) - \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n).$$

From the exact triangle (3.2) and the fact that $\Gamma_- = \gamma_{T_0}$, we have

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T_0}, -\gamma_{T_0}) \leqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_{n+1}) - \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T_0}, -\Gamma_n).$$

Step 3. We obtain the desired inequality from the above equality and inequality.

Note that a_0 is an arc obtained by pushing a neighborhood of q_l in ζ into the interior of M_T, and M_{T_0} is obtained from M_T by removing $N(a_0)$. There is an embedded disk D in M_{T_0} whose boundary is the union of a_0 and the neighborhood of q_l in ζ. Moreover, ∂D intersects γ_{T_0} at two points, one of which is q_l and the other is in γ_0. By the proof of [34, Proposition 6.9], decomposing the sutured manifold (M_{T_0}, γ_{T_0}) along D does not change the isomorphism class of the sutured instanton Floer homology (D is a product disk in the sense of [27, Definition 2.8]). It is straightforward to check the sutured manifold after the sutured manifold decomposition is exactly (M_T, γ_T). Then we have

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T_0}, -\gamma_{T_0}) = \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_T, -\gamma_T).$$

Similarly, we have

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T_0'}, -\gamma_{T_0'}) = \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T'}, -\gamma_{T'}).$$

Thus, we conclude

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_T, -\gamma_T) \leqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-M_{T'}, -\gamma_{T'}).$$ □

Proof of Theorem 1.2. Suppose $(H, \gamma) = (Y(1)_T, \delta_T)$ and $(Y(K), \beta''_g \cup \beta''_{g+1})$ are obtained from Construction 3.11. Note that $\beta''_g \cup \beta''_{g+1}$ are parallel copies of the meridian of K. Then we have

$$KHI(-Y, K) = \underline{\mathrm{SHI}}(-Y(K), -(\beta''_g \cup \beta''_{g+1}))$$

by Definition 2.17. Since Y is a rational homology sphere, we have

$$H_1(Y(1), \partial Y(1); \mathbb{Q}) = 0.$$

In particular, any component of T has trivial rational homology class. Then the theorem follows from Proposition 3.14 and Theorem 3.3. □

Remark 3.23. Suppose (Σ, α, β) is a Heegaard diagram of a rational homology sphere Y and K is the core knot of β_i for some $\beta_i \subset \beta$. Suppose $(H, \gamma) = (Y(1)_T, \delta_T)$ is obtained from Construction 3.12. Then the proof of Theorem 1.2 applies without change, and we conclude the same inequality.

Proof of Proposition 1.5. Similar to the proof of Theorem 1.2, since the knot K has trivial rational homology class, the corresponding tangle has trivial homology class in $H_1(Y(1), \partial Y(1); \mathbb{Q})$. □

Corollary 3.24. *Suppose (Σ, α, β) is a genus g Heegaard diagram of a rational homology sphere Y. Let (H, γ) be the sutured handlebody obtained by Construction 3.12. Suppose*

$$n = \frac{1}{2} \sum_{i=1}^{g} |\alpha_i \cap \gamma|.$$

Then we have

$$\dim_{\mathbb{C}} I^{\sharp}(-Y) \leqslant 2^{n-g}.$$

Proof. By Theorem 1.2, we know that

$$\dim_{\mathbb{C}} I^{\sharp}(Y) \leqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma).$$

Hence it suffices to prove that

$$\underline{\mathrm{SHI}}(-M, -\gamma) \leqslant 2^{n-g}.$$

First note that if $n < g$, then there exists an integer $i \in [1, g]$ such that $\alpha_i \cap \gamma = \emptyset$. Hence $(-H, -\gamma)$ is not taut. By Theorem 2.14, $\underline{\mathrm{SHI}}(-M, -\gamma) = 0$.

When $g = n$, then either there exists an integer $i \in [1, g]$ such that $\alpha_i \cap \gamma = \emptyset$, and then $\underline{\mathrm{SHI}}(-M, -\gamma) = 0$ as above, or for each $i \in \{1, \ldots, g\}$, α_i intersects γ precisely at two points. Thus the disk $D_i \subset H$ bounded by α_i is a product disk. Let $D = D_1 \cup \cdots \cup D_g$. We can perform a sutured manifold decomposition $(-H, -\gamma) \stackrel{D}{\rightsquigarrow} (D^3, \delta)$, where D^3 is a 3-ball and δ is a suture on ∂D^3. From Theorem 2.15, we know that

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma) = \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(D^3, \delta) = 1.$$

We prove other cases by induction on n. Assume that the statement has been proved for $g \leqslant n = k - 1$. For the case of $n = k$, we proceed as follows. Since $k > g$, there is a curve α_i satisfying $\alpha_i \cap \gamma \geqslant 4$. Suppose $\eta \subset \alpha_i$ is an arc such that $\partial \eta \subset \gamma$, and the interior of η intersects with γ transversely once. By Theorem 2.37, we have a bypass exact triangle from the bypass attachment along η:

$$\begin{tikzcd}
& \underline{\mathrm{SHI}}(-H, -\gamma) \arrow[dr] & \\
\underline{\mathrm{SHI}}(-H, -\gamma_2) \arrow[ur] & & \underline{\mathrm{SHI}}(-H, -\gamma_1) \arrow[ll]
\end{tikzcd}$$

It is straightforward to check that

$$\sum_{i=1}^{n} |\gamma_1 \cap \alpha_i| \leqslant 2k - 2 \text{ and } \sum_{i=1}^{n} |\gamma_2 \cap \alpha_i| \leqslant 2k - 2.$$

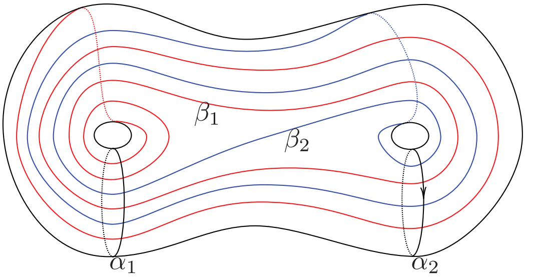

FIGURE 10 A Heegaard diagram of Y

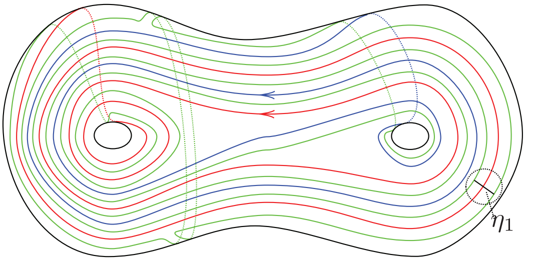

FIGURE 11 The diagram of (H, γ)

By the induction hypothesis and the exactness, we conclude that

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma) \leqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma_1) + \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma_2) \leqslant 2^{k-g}.$$

Thus, we complete the induction. $\square$

Example 3.25. Suppose we have a Heegaard diagram (Σ, α, β) of Y as in Figure 10. It is straightforward to check that

$$H_1(Y) = \mathbb{Z}_2 \oplus \mathbb{Z}_2.$$

We can apply Construction 3.12 to obtain a sutured handlebody (H, γ). See Figure 11. By Theorem 1.2, we know that

$$\dim_{\mathbb{C}} I^{\sharp}(-Y) \leqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma). \tag{3.6}$$

It remains to bound $\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma)$. If we apply Corollary 3.24 directly, we obtain

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma) \leqslant 64.$$

However, we can improve this bound by examining the bypass exact triangles more carefully in the following steps.

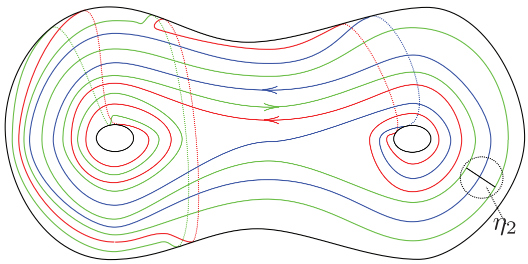

FIGURE 12 The diagram of (H, γ_1)

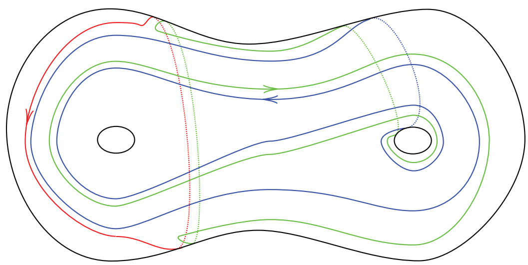

FIGURE 13 The diagram of (H, γ_2)

Step 1. We can do a bypass attachment along the arc η_1 as shown in Figure 11. By Theorem 2.37, there exists an exact triangle

$$\begin{tikzcd}
 & \underline{\mathrm{SHI}}(-H,-\gamma) \arrow[dr] & \\
\underline{\mathrm{SHI}}(-H,-\gamma_2) \arrow[ur] & & \underline{\mathrm{SHI}}(-H,-\gamma_1) \arrow[ll]
\end{tikzcd} \tag{3.7}$$

The sutures γ_1 and γ_2 are depicted in Figures 12 and 13, respectively.

Step 2. We compute $\underline{\mathrm{SHI}}(-H,-\gamma_2)$. The curve α_2 bounds a disk $D \subset H$. It then induces a grading on $\underline{\mathrm{SHI}}(-H,-\gamma_2)$. Note that D is not admissible in the sense of Definition 2.20, so we perform a negative stabilization on D as in Definition 2.24, and write $D^-$for the resulting disk. By Theorem 2.21, $\underline{\mathrm{SHI}}(-H,-\gamma_2,D^-,i)=0$, for $|i|>1$. We can perform a sutured manifold decomposition

$$(-H,-\gamma_2) \overset{D^-}{\rightsquigarrow} (V,\gamma_2'),$$

where V is a solid torus and γ_1' is depicted as in Figure 14. From Theorem 2.21 and Theorem 2.14, we know that

$$\underline{\mathrm{SHI}}(-H,-\gamma_2,D^-,1) \cong \underline{\mathrm{SHI}}(V,\gamma_2') = 0.$$

By Lemma 2.25, Theorem 2.21, and Theorem 2.14, we know

$$\begin{aligned}
\underline{\mathrm{SHI}}(-H,-\gamma_2,D^-,-1) &= \underline{\mathrm{SHI}}(-H,-\gamma_2,-D^-,1) \\
&= \underline{\mathrm{SHI}}(-H,-\gamma_2,(-D)^+,1) \\
&= 0.
\end{aligned}$$

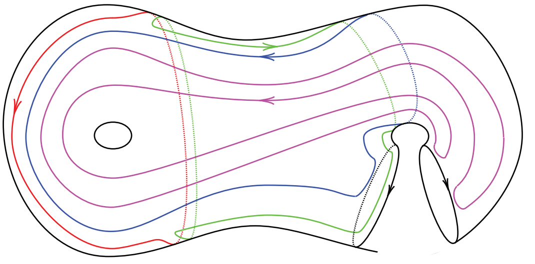

FIGURE 14 The diagram of (V, γ_1')

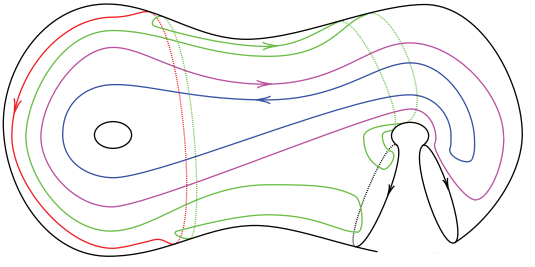

FIGURE 15 The diagram of (V, γ_2'')

By Theorem 2.21, Theorem 2.27, and Theorem 2.14, we know

$$\begin{aligned}\underline{\mathrm{SHI}}(-H, -\gamma_2, D^-, 0) &= \underline{\mathrm{SHI}}(-H, -\gamma_2, -D^-, 0)\\ &= \underline{\mathrm{SHI}}(-H, -\gamma_2, (-D)^+, 0)\\ &= \underline{\mathrm{SHI}}(-H, -\gamma_2, (-D)^-, 1)\\ &= \underline{\mathrm{SHI}}(V, \gamma_2'').\end{aligned}$$

Here (V, γ_2'') is obtained from $(-H, -\gamma_2)$ by decomposing along $(-D)^-$. V is a solid torus and γ_2'' is depicted as in Figure 15. From [42, Proposition 1.4], we know that

$$\underline{\mathrm{SHI}}(V, \gamma_2'') \cong \mathbb{C}^2.$$

Hence we conclude that

$$\underline{\mathrm{SHI}}(-H, -\gamma_2) \cong \mathbb{C}^2. \tag{3.8}$$

Step 3. We can perform a second bypass along the arc η_2 as shown in Figure 12 on (H, γ_1) and obtain an exact triangle

$$\begin{tikzcd}[column sep=small] & \underline{\mathrm{SHI}}(-H,-\gamma_1) \arrow[dr] & \\ \underline{\mathrm{SHI}}(-H,-\gamma_4) \arrow[ur] & & \underline{\mathrm{SHI}}(-H,-\gamma_3) \arrow[ll, "\psi_4^3"] \end{tikzcd} \tag{3.9}$$

It is straightforward to check that both γ_3 and γ_4 intersect the disk D at four points, so we can compute in the same way as we did for $\underline{\mathrm{SHI}}(-H,-\gamma_2)$

$$\underline{\mathrm{SHI}}(-H,-\gamma_3)\cong\mathbb{C}^3,\ \text{and}\ \underline{\mathrm{SHI}}(-H,-\gamma_4)\cong\mathbb{C}^5. \tag{3.10}$$

From (3.6), (3.7), (3.8), (3.9), and (3.10), we know that

$$\dim_{\mathbb{C}} I^{\sharp}(-Y)\leqslant 10.$$

3.3 | The instanton knot homology of $(1,1)$-knots

In this subsection, we use Theorem 1.2 to prove Theorem 1.6.

Definition 3.26. Suppose $p,q\in\mathbb{Z}$ satisfy $p\geqslant 1, 0\leqslant q<p$ and $\gcd(p,q)=1$. Let $\tilde{\alpha}$ and $\tilde{\beta}$ be two straight lines in $\mathbb{R}^2$ passing the origin with slopes 0 and p/q, respectively, and let $r:\mathbb{R}^2\to T^2$ be the quotient map induced by $(x,y)\to(x+m,y+n)$ for $m,n\in\mathbb{Z}$. Suppose $\alpha=r(\tilde{\alpha})$ and $\beta=r(\tilde{\beta})$. Then the manifold compatible with the Heegaard diagram (T^2,α,β) is called a *lens space* and is denoted by $L(p,q)$. Furthermore, the Heegaard diagram (T^2,α,β) is called the *standard diagram* of the lens space. In particular, we regard S^3 as a lens space $L(1,0)$.

The lens space is oriented so that the orientation on the α-handlebody is induced from the standard embedding of $S^1\times D^2$ in $\mathbb{R}^3$. With this convention, the lens space $L(p,q)$ comes from the p/q-surgery on the unknot in S^3.

Definition 3.27. A proper embedded arc η in a handlebody H is called a *trivial arc* if there is an embedded disk $D\subset H$ satisfying $\partial D=\eta\cup(D\cap\partial H)$. The disk D is called the *cancelling disk* of η. A knot K in a closed 3-manifold Y admits a (1,1)-*decomposition* if the followings hold.

(1) Y admits a splitting $Y=H_1\cup_{T^2}H_2$ so that $H_1\cong H_2\cong S^1\times D^2$.
(2) $K\cap H_i$ is a properly embedded trivial arc in H_i for $i\in\{1,2\}$.

In this case, Y is either a lens space or $S^1\times S^2$. A knot K admitting a (1,1)-decomposition is called a $(1,1)$-*knot*.

Proposition 3.28 [56, Section 6.2; 17, Section 2]. *For $p,q,r,s\in\mathbb{N}$ satisfying $2q+r\leqslant p$ and $s<p$, a (1,1)-decomposition of a knot determines and is determined by a doubly-pointed diagram. After isotopy, such a diagram becomes (T^2,α,β,z,w) in Figure 16, where p is the total number of intersection points, q is the number of strands around either basepoint, r is the number of strands in the middle band, and the ith point on the right-hand side is identified with the $(i+s)$-th point on the left-hand side.*

Definition 3.29. A simple closed curve β on (T^2,α,z,w) is called *reduced* if the number of intersection points between α and β is minimal. The doubly pointed diagram in Figure 16 is called the (1,1)-*diagram* of type (p,q,r,s), which is denoted by $W(p,q,r,s)$. Strands around basepoints are called *rainbows* and strands in the bands are called *stripes*.

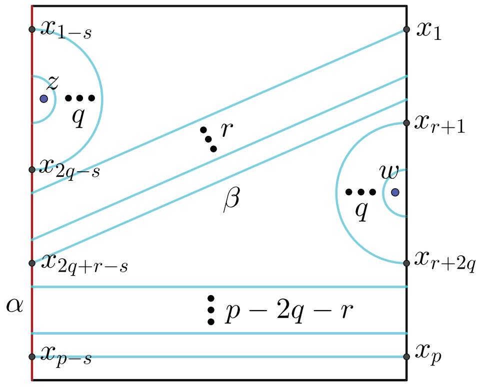

FIGURE 16 (1,1)-diagram

If the (1,1)-*diagram* of $W(p,q,r,s)$ is a Heegaard diagram for some parameters (p,q,r,s), or equivalently, β has one component and represents a non-trivial homology class in $H_1(T^2)$, then the corresponding knot is also denoted by $W(p,q,r,s)$.

A $(1,1)$-knot whose (1,1)-diagram does not have rainbows is called a *simple knot* (*c.f.* [57, Section 2.1]). For simple knots, let $K(p,q,k) = W(p,0,k,q)$.

Proposition 3.30. *The mirror knot of a* $(1,1)$*-knot* $W(p,q,r,s)$ *is*

$$W(p,q,p-2q-r,p-s+2q).$$

Proof. The Heegaard diagram of the mirror knot of $W(p,q,r,s)$ is obtained by the (1,1)-diagram of $W(p,q,r,s)$ by vertical reflection. We redraw the Heegaard diagram so that the lower band becomes the middle band and the middle band becomes the lower band. This proposition follows from the definition. $\square$

According to [17, Section 3] (also [52, Section 6]), for the $\widehat{HFK}$ of a $(1,1)$-knot, the generators of the chain complexes are intersection points of α and β in the (1,1)-diagram and there is no differential. Thus, the following proposition holds.

Proposition 3.31. *For a* $(1,1)$*-knot* $K = W(p,q,r,s)$ *in* Y*, we have*

$$\widehat{HFK}(Y,K) \cong \mathbb{Z}^p.$$

We restate Construction 3.12 more carefully.

Construction 3.32. Suppose (T^2,α,β,z,w) is the (1,1)-diagram of $W(p,q,r,s)$. We construct a sutured handlebody (H,γ) as follows, called the (1,1)-*sutured-handlebody* of $W(p,q,r,s)$.

(1) Let Σ be the genus-two boundary of the manifold obtained from $[-1,1]\times T^2$ by attaching a 3-dimensional 1-handle along $\{1\}\times\{z,w\}$. For simplicity, when drawing the diagram, the attached 1-handle will still be denoted by two basepoints z and w.

(2) Let α_1 and β_1 denote the curves on Σ induced from α and β, respectively. Let β be oriented so that the innermost rainbow around z is oriented clockwise, which induces an orientation of β_1. If there is no rainbow, let β be oriented so that each stripe goes from left to right in Figure 16.
(3) Consider the straight arc connecting z to w in Figure 16. It induces a simple closed curve α_2 on Σ by going along the 1-handle. Let β_2 be the curve on Σ induced by a small circle around z, oriented counterclockwise.
(4) Let γ_1 and γ_2 be obtained by pushing off β_1 and β_2 to the right with respect to the orientation. Suppose they are oriented reversely with respect to β_1 and β_2, respectively. Let a_0 be a straight arc connecting the innermost rainbow of β around z to the above small circle. It induces an arc connecting γ_1 to γ_2, still denoted by a_0. Let γ_3 be obtained by a band sum of γ_1 and γ_2 along a_1, with the induced orientation.
(5) Let H be the handlebody compatible with the diagram $(\Sigma, \{\alpha_1, \alpha_2\}, \emptyset)$ and let

$$\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3.$$

Rainbows and stripes are defined similarly for sutures.

The main goal is to prove the following theorem.

Theorem 3.33. *Suppose (H, γ) is the (1,1)-sutured-handlebody of $W(p, q, r, s)$ constructed in Construction 3.32. Then we have*

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma) \leqslant p.$$

Before proving this theorem, we first use it to derive Theorem 1.6.

Proof of Theorem 1.6. Combining Theorem 1.2, Proposition 3.31, and Theorem 3.33, for a $(1,1)$-knot $K = W(p, q, r, s)$ in a lens space Y, we have

$$\dim_{\mathbb{C}} KHI(-Y, K) \leqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma) \leqslant p = \mathrm{rk}_{\mathbb{Z}} \widehat{HFK}(-Y, K).$$

Then the theorem follows from Proposition 3.30, that is, the mirror knot of a $(1,1)$-knot is still a $(1,1)$-knot with the same intersection number p. □

Proof of Theorem 3.33. We prove the theorem by induction on p for any (1,1)-diagram of $W(p, q, r, s)$ where β has only one component. This includes the case that β represents a trivial homology class. The induction is based on the bypass exact triangle in Theorem 2.37. We will show three balanced sutured manifolds in the bypass exact triangle are all (1,1)-sutured handlebodies, where one is the (1,1)-sutured handlebody we want and the other two are (1,1)-sutured handlebodies with smaller number p. By straightforward algebra, if the dimension inequality holds for two terms in the bypass exact triangle, then it also holds for the third term.

For the base case, consider $p = 1$. The curves $\alpha_1, \alpha_2, \beta_1, \beta_2$ in Construction 3.32 satisfy

$$|\alpha_1 \cap \beta_1| = |\alpha_2 \cap \beta_2| = 1.$$

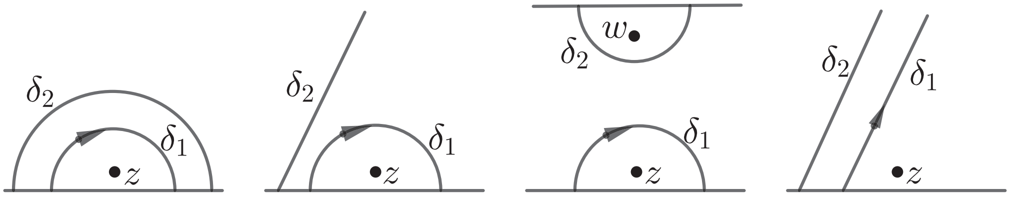

FIGURE 17 Several cases of δ_1 and δ_2

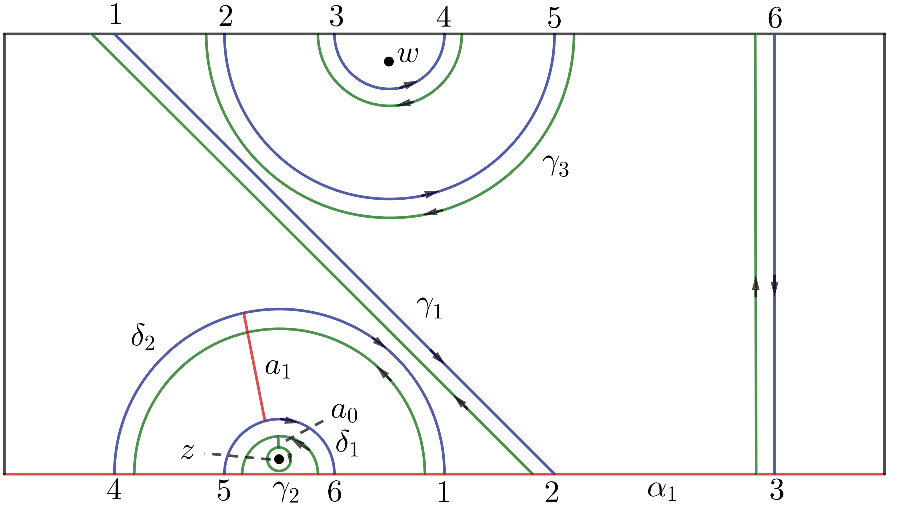

FIGURE 18 The suture related to $W(6,2,1,3)$ and the anti-wave bypass arc

It is straightforward to check (H, γ) is a product sutured manifold, so is $(-H, -\gamma)$. Then Theorem 2.15 implies

$$\dim_{\mathbb{C}} SHI(-H, -\gamma) = 1.$$

Now we deal with the case where $p > 1$. In Construction 3.32, the innermost rainbow around z, if exists, is oriented clockwise. Suppose δ_1 is either the innermost rainbow around z, or a stripe that is closest to z with z on its right-hand side. Suppose δ_2 is another rainbow or stripe that is closest to δ_1 and is to the left of δ_1. See Figure 17 for all possible cases. Compared to Figure 16, we have rotated the square counterclockwise by 90° for the purpose of a better display.

We consider two different cases about the orientation of δ_2.

Case 1. Suppose δ_1 and δ_2 are oriented parallelly.

We use $W(6,2,1,3)$ shown in Figure 18 as an example to carry out the proof, and the general case is similar. In this example, two innermost rainbows around z are oriented parallelly. By construction, the curve γ_3 is parallel (regardless of orientations) to γ_1 outside the neighborhood of the band-sum arc a_0. Thus, there exists a unique rainbow of γ_3 between δ_1 and δ_2 around z. Let a_1 be an anti-wave bypass arc cutting these three rainbows, as shown in Figure 18. Suppose γ' and γ'' are the other two sutures involved in the bypass triangle associated to a_2.

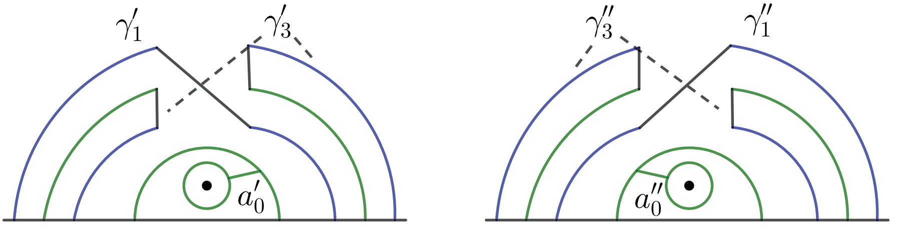

FIGURE 19 Local diagrams after bypass attachments

From Proposition 2.31, we can describe the sutures γ' and γ'' as follows. First, let γ_1^1 and γ_1^2 be two components of $\gamma_1 \backslash \partial a_1$. Suppose

$$\gamma_1' = \gamma_1^1 \cup a_1 \text{ and } \gamma_1'' = \gamma_1^2 \cup a_1$$

as shown in Figure 19. Second, γ' is obtained from γ by a Dehn twist along γ_1'', and γ'' is obtained from γ by a Dehn twist along γ_1'.

There is a more direct way to describe γ' and γ''. First, note that the suture γ_2 is disjoint from both Dehn-twist curves γ_1' and γ_1'', so γ_2 remains the same in γ' and γ''. Second, it is straightforward to check the result of γ_1 under the Dehn twist along γ_1'' is γ_1', and the result of γ_1 under the Dehn twist along γ_1' is γ_1''. Thus, γ_1' is a component of γ' and γ_1'' is a component of γ''.

To figure out the image γ_3' of γ_3 under the Dehn twist along γ_1'', we first observe that we can isotop the band-sum arc a_0 to a new position a_0' such that its endpoints $\partial a_0'$ lie on $\gamma_1' \cap \gamma_1$ and γ_2, as shown in the left subfigure of Figure 19. Thus, the facts that a_0' is disjoint from γ_1'' and that γ_1' is the image of γ_1 under the Dehn twist along γ_1'' imply that performing a Dehn twist along γ_1'' and performing the band sum along a_0' commute with each other. Thus, we conclude that γ_3' can be obtained from a band sum on γ_1' and γ_2 along the arc a_0'. Similarly we can describe the image γ_3'' of γ_3 under the Dehn twist along γ_1'. Thus, we have described the sutures

$$\gamma' = \gamma_1' \cup \gamma_2 \cup \gamma_3' \text{ and } \gamma'' = \gamma_1' \cup \gamma_2 \cup \gamma_3''$$

explicitly, and it follows that (H, γ') and (H, γ'') are both (1,1)-sutured handlebodies. Suppose they are associated to $W(p', q', r', s')$ and $W(p'', q'', r'', s'')$, respectively.

From the above description, both γ_1' and γ_1'' are reduced. We have

$$p' + p'' = |\gamma_1' \cap \alpha_1| + |\gamma_1'' \cap \alpha_1| = |\gamma_1 \cap \alpha_1| = p.$$

Thus, the induction applies.

Case 2. Suppose δ_1 and δ_2 are oriented oppositely.

An example $W(10, 3, 1, 5)$ is shown in Figure 20. By construction, there is a rainbow of γ_3 to the right of δ_2. Let a_2 be a wave bypass arc cutting δ_1, δ_2, and this rainbow as shown in Figure 20. Suppose γ' and γ'' are the other two sutures involved in the bypass triangle associated to a_2, respectively.

To describe the sutures γ' and γ'' more explicitly, note that the arc a_2 cuts γ_1 into two parts γ_1^1 and γ_1^2. Suppose that near a_2, γ_1^1 is to the left of a_2 and γ_1^2 is to the right of a_2. For γ', it consists of

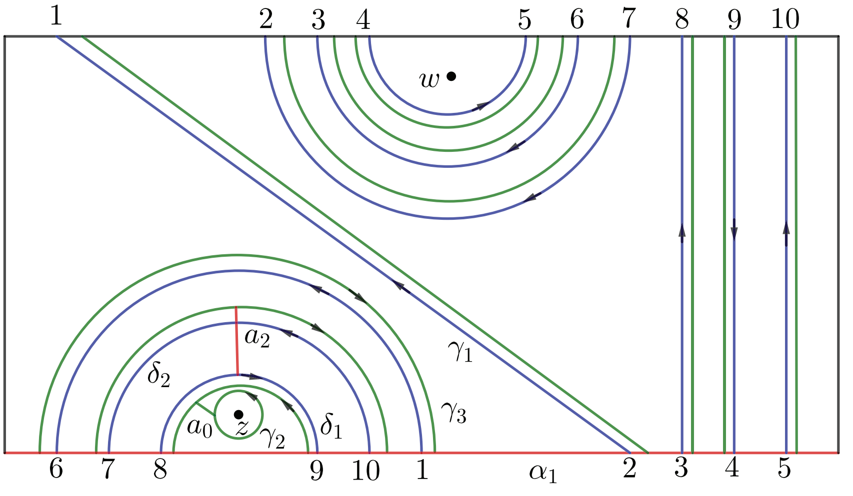

FIGURE 20 The suture related to $W(10,3,1,5)$ and the wave bypass arc

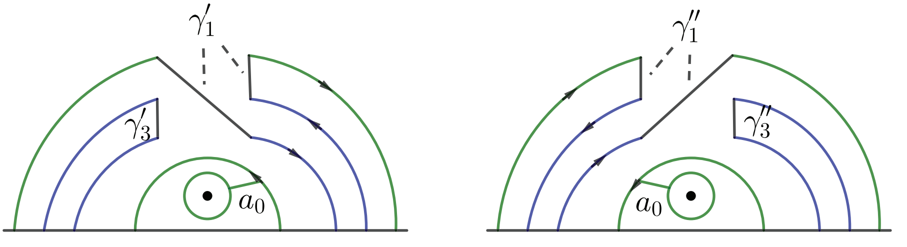

FIGURE 21 Local diagrams after bypass attachments

FIGURE 22 Local diagrams after isotopy

three components:

$$\gamma' = \gamma_1' \cup \gamma_2 \cup \gamma_3',$$

where γ_2 is as before, γ_1' is obtained by cutting γ_3 open by a_2 and gluing it to γ_1^2, and γ_3' is obtained by gluing a copy of a_2 to γ_1^1. They are depicted as in the left subfigure of Figure 21. Note that the curve γ_1' is not reduced. We can isotop the curve along the arc γ_1^2 into a reduced curve. The orientations of curves imply this reduced curve is depicted as in the left subfigure of Figure 22. Note that γ_3' is also not reduced. However, from Figure 22 it is straightforward to check that γ_3' can be thought of as obtained from γ_2 and γ_1' by a band sum along the arc a_0'. Also, it is clear that

$$|\gamma_1' \cap \alpha_1| = |\gamma_1^1 \cap \alpha_1|.$$

Similarly, γ'' consists of three components:

$$\gamma'' = \gamma_1'' \cup \gamma_2 \cup \gamma_3'',$$

where γ_2 is as before, γ_1'' is obtained by cutting γ_3 open by a_2 and gluing it to γ_1^1, and γ_3'' is obtained by gluing a copy of a_2 to γ_1^2. They are depicted as in the right subfigure of Figure 21. Considering the orientations, we can isotop γ_2'' along γ_2^1 to the position shown in the right subfigure of Figure 22. Then γ_3'' can be thought of as obtained from γ_1 and γ_2'' by a band sum along a_0''. Also,

$$|\gamma_1'' \cap \alpha_1| = |\gamma_1^2 \cap \alpha_1|.$$

Hence we conclude that (H, γ') and (H, γ'') are both (1,1)-sutured-handlebodies, and

$$|\gamma_2' \cap \alpha_1| + |\gamma_2'' \cap \alpha_1| = |\gamma_{2,1} \cap \alpha_1| + |\gamma_{2,2} \cap \alpha_1| = |\gamma_2 \cap \alpha_1|.$$

Thus, the induction applies. □

3.4 | Large surgeries on simple knots

In this subsection, we generalize the idea about the anti-wave bypass arc in Case 1 of the proof of Theorem 3.33 to prove Theorem 1.10.

For a simple knot $K = K(p, q, k) \subset Y$ defined as in Definition 3.29, consider its (1,1)-diagram $(T^2, \alpha_1, \beta_1, z, w)$ and the Heegaard diagram $(\Sigma, \{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\})$ of Y from Construction 3.9. Suppose $\alpha = \{\alpha_1, \alpha_2\}$ and $m = \beta_2$. Let l be an arc connecting z to w in $T^2 - \beta$, which induces a curve on Σ, still denoted by l. Then $(\Sigma, \alpha, \{\beta_1\})$ is a diagram of the knot complement $Y(K)$ and (m, l) forms a basis of $H_1(\partial Y(K))$.

Given $p, q \in \mathbb{Z}$ such that $\gcd(p, q) = 1$, let β_2' be the curve on Σ obtained by resolving intersection points of $|q|$ parallel copies of m and $|p|$ parallel copies of l. Then $(\Sigma, \alpha, \{\beta_1, \beta_2'\})$ is a Heegaard diagram of the manifold obtained from the q/p-surgery on K.

Remark 3.34. There are two ways of resolutions, namely the positive resolution and the negative resolution. The choice of the ways depends on orientations of Σ and curves and signs of p and q. However, the goal in Theorem 1.10 is for any large enough surgery slope without regard for the sign. So the choice here is not important.

Definition 3.35 [18, Section 7]. An arc a in a Heegaard diagram (Σ, α, β) is called an *anti-wave* if it satisfies the following conditions.

(1) It is properly embedded in a component R of $\Sigma \backslash (\alpha \cup \beta)$.
(2) Its endpoints lie on the interior of distinct arcs r_1, r_2 of ∂R, where r_1 and r_2 are subsets of the same curve $\alpha_i \subset \alpha$ or $\beta_i \subset \beta$ for some i.
(3) The local signs of intersection at two endpoints are the same.

For the simple knot K, there exist anti-waves in each component of $T^2\backslash\alpha\cup\beta$. It is possible to choose l so that $l\cap R=\emptyset$ for some component R of $T^2\backslash(\alpha\cup\beta)$. Let a be an anti-wave in R. It induces an anti-wave in $(\Sigma,\alpha,\{\beta_1,\beta_2'\})$ which is still denoted by a.

Lemma 3.36. *Given the surgery slope q/p, consider the Heegaard diagram $(\Sigma,\alpha,\{\beta_1,\beta_2'\})$ and the anti-wave a defined as above. Let β_1^1 and β_1^2 denote two arcs of $\beta_1\backslash\partial a$ and let $\beta_{1,i}=\beta_1^i\cup a$ for $i\in\{1,2\}$. Suppose Y_i is the manifold compatible with the Heegaard diagram $(\Sigma,\alpha,\{\beta_{1,i},\beta_2'\})$ and K_i is the core knot of β_2'. Suppose $(Y_0,K_0)=(Y,K)$. Then there exists a knot $K'\subset Y$ such that (Y_i,K_i), for $i=0,1,2$, satisfy the exact triangle associated to K' in Theorem 2.19.*

Proof. Consider neighborhoods $N(\beta_1)$ and $N(\beta_2')$ of β_1 and β_2' on Σ, respectively. The manifold

$$\Sigma\backslash\mathrm{int}(N(\beta_1))\cup\mathrm{int}(N(\beta_2'))$$

is diffeomorphic to $S^2\backslash\bigcup_{i=1}^4 D_i$, where D_i are pair-wise disjoint disks. Suppose ∂D_1 and ∂D_2 are images of $\partial N(\beta_2')$ under the diffeomorphism and a becomes an arc connecting ∂D_1 to ∂D_2. There exists a curve $\beta_3\subset S^2$ separating $D_1\cup D_2$ and $D_3\cup D_4$. It induces a null-homologous curve on Σ which is still denoted by β_3. By construction, β_3 is disjoint from β_1,β_2' and $\beta_{1,i}$ for $i\in\{1,2\}$.

The 3-manifold compatible with the diagram $(\Sigma,\alpha,\{\beta_3\})$ has two toroidal boundary components. It can be regarded as the complement of K and K' in Y for some knot K'. Equivalently, K is the core knot of β_2 and K' is the core knot of β_1 and the way how K and K' is linked is determined by β_3. After isotopy, $\beta_2,\beta_{2,1}$ and $\beta_{2,2}$ intersect pair-wise at one point. Then Theorem 2.19 applies to this case for K'. $\square$

Lemma 3.37. *There exists $N_0>0$ so that for any surgery slope q/p with $|q/p|\geqslant N_0$, the manifolds Y_i for $i=0,1,2$ in Lemma 3.36 corresponding to q/p satisfy*

$$|H_1(Y_0)|=|H_1(Y_1)|+|H_1(Y_2)|.$$

Proof. For $i\in\{1,2\}$, it is straightforward to check that $\beta_{1,i}$ is a reduced curve without rainbows (*cf.* Definition 3.29). Suppose orientations of curves are chosen so that

$$|\beta_{1,i}\cap\alpha_1|=\beta_{1,i}\cdot\alpha_1=p_i \text{ and } |\beta_{1,i}\cap\alpha_2|=\beta_{1,i}\cdot\alpha_2=k_i.$$

By construction, we have

$$p_1+p_2=p_0 \text{ and } k_1+k_2=k_0.$$

It is clear that K_i is the dual knot of some surgery on $K(p_i,q_i,k_i)$ for some q_i. Suppose

$$l\cdot\alpha_1=x \text{ and } l\cdot\alpha_2=y.$$

Suppose the orientation of m is chosen so that $m\cdot\alpha_2=1$. Then we have

$$\beta_2'\cdot\alpha_1=px, \text{and } \beta_2'\cdot\alpha_2=q+py.$$

Thus, for $i \in \{1,2\}$, we have

$$|H_1(Y_0)| = |pk_0 - (q+py)p_0| \text{ and } |H_1(Y_i)| = |pk_i - (q+py)p_i|.$$

One of these orders is the sum of other two. If $|q/p|$ is large enough, the order

$$|H_1(Y_0)| = p|k_0 - (q/p+y)p_0|$$

is greater than $|H_1(Y_i)|$ for $i \in \{1,2\}$. Hence we conclude the lemma. □

Proof of Theorem 1.10. We prove the theorem by induction on p_0 for simple knots $K(p_0, q_0, k_0)$.

For the base case, consider $p_0 = 1$. The simple knot is the unknot in S^3. The dual knot is the core knot of β for the standard diagram of a lens space, which is also a simple knot (*cf.* [57, Section 2.3]). The theorem follows from Proposition 1.9.

When $p_0 > 1$, consider knots K_i and the related simple knots $K(p_i, q_i, k_i)$ for $i \in \{1,2\}$ in Lemmas 3.36 and 3.37. By induction hypothesis, there exist N_i for $K(p_i, q_i, k_i)$ so that r surgery with $|r| \geqslant N$ induces an instanton Floer simple knot. Equivalently,

$$\dim_{\mathbb{C}} KHI(Y_1, K_1) = |H_1(Y_1)| \text{ and } \dim_{\mathbb{C}} KHI(Y_2, K_2) = |H_1(Y_2)|.$$

We have to discuss the basis of the homology at first. The basis (m, l_0) of $H_1(\partial Y(K_0))$ is chosen with respect to the anti-wave in the Heegaard diagram corresponding to K_0, which induces a basis on $H_1(\partial Y_i(K_i))$. However, in the proofs of Lemmas 3.36 and 3.37, another basis (m, l_i) is chosen with respect to the anti-wave in the Heegaard diagram corresponding to K_i. Suppose $l_i = x_i m + l$. Then

$$qm + pl = (q - px_i)m + pl_i.$$

Suppose $N = \max\{N_1 + |x_1|, N_2 + |x_2|, N_0\}$. Then Lemma 3.37 implies

$$|H_1(Y_0)| = |H_1(Y_1)| + |H_1(Y_2)|.$$

Combining Theorem 2.19 and Lemma 3.36, we have

$$\dim_{\mathbb{C}} KHI(Y_0, K_0) \leqslant \dim_{\mathbb{C}} KHI(Y_1, K_1) + \dim_{\mathbb{C}} KHI(Y_2, K_2).$$

Hence we have

$$\dim_{\mathbb{C}} KHI(Y_0, K_0) \leqslant |H_1(Y_0)|.$$

Combining Theorem 1.2 and [59, Corollary 1.4], we have

$$\dim_{\mathbb{C}} KHI(Y_0, K_0) \geqslant \dim_{\mathbb{C}} I^{\sharp}(Y_0) \geqslant |H_1(Y_0)|.$$

Thus,

$$\dim_{\mathbb{C}} KHI(Y_0, K_0) = |H_1(Y_0)|.$$

and the induction applies. □

Remark 3.38. For the above proof, the induction on $\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma)$ does not work any more because the inequality

$$\dim_{\mathbb{C}} KHI(-Y, K) \leqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-H, -\gamma)$$

might not always be sharp. That is the reason why we switch from bypass exact triangles to surgery exact triangles.

4 | INSTANTON FLOER HOMOLOGY AND THE DECOMPOSITION

4.1 | Basic setups

Suppose Y is a closed 3-manifold and $K \subset Y$ is a null-homologous knot. Let $Y(K)$ be the knot complement $Y \backslash \mathrm{int}(N(K))$. Any Seifert surface S of K gives rise to a framing on $\partial Y(K)$: the longitude λ can be picked as $S \cap \partial Y(K)$ with the induced orientation from S, and the meridian μ can be picked as the meridian of the solid torus $N(K)$ with the orientation so that $\mu \cdot \lambda = -1$. The 'half lives and half dies' fact for 3-manifolds implies that the following map has a 1-dimensional image:

$$\partial_* : H_2(Y(K), \partial Y(K); \mathbb{Q}) \to H_1(\partial Y(K); \mathbb{Q}).$$

Hence, any two Seifert surfaces lead to the same framing on $\partial Y(K)$. We write $g(K)$ for the minimal genus of the Seifert surface of K. If a Seifert surface of minimal surface is chosen, we also write it as $g(S)$.

Definition 4.1. The framing (μ, λ) defined as above is called the *canonical framing* of (Y, K). With respect to this canonical framing, let

$$\widehat{Y}_{q/p} = Y(K) \cup_{\phi} S^1 \times D^2$$

be the 3-manifold obtained from Y by a q/p surgery along K, that is,

$$\phi(\{1\} \times \partial D^2) = q\mu + p\lambda.$$

When the surgery slope is understood, we also write $\widehat{Y}_{q/p}$ simply as $\widehat{Y}$. Let $\widehat{K}$ be the dual knot, that is, the image of $S^1 \times \{0\} \subset S^1 \times D^2$ in $\widehat{Y}$ under the gluing map.

Convention. Throughout this section, we will always assume that $\gcd(p, q) = 1$ and $q > 0$ or $(p, q) = (1, 0)$ for a Dehn surgery. Especially, the original pair (Y, K) can be thought of as a pair $(\widehat{Y}, \widehat{K})$ obtained from (Y, K) by the $1/0$ surgery. Moreover, we will always assume that the knot complement $Y(K)$ is irreducible. This is because if $Y(K)$ is not irreducible, then $Y(K) \cong Y'(K')\sharp Y''$ for some closed 3-manifold Y', Y'' and a null-homologous knot $K' \subset Y'$. By the connected sum formula [40, Section 1.2], we have

$$\underline{\mathrm{SHI}}(Y(K),\gamma) \cong \underline{\mathrm{SHI}}(Y'(K'),\gamma) \otimes I^\sharp(Y'')$$

for any suture γ. Hence, all results hold after tensoring $I^\sharp(Y'')$.

Next, we describe various families of sutures on the knot complement. Suppose $K \subset Y$ is a null-homologous knot and the pair $(\widehat{Y},\widehat{K})$ is obtained from (Y,K) by a q/p surgery. Note we can identify the complement of $K \subset Y$ with that of $\widehat{K} \subset \widehat{Y}$, that is, $\widehat{Y}(\widehat{K}) = Y(K)$.

On $\partial Y(K)$, there are two framings: One comes from K, and we write longitude and meridian as λ and μ, respectively. The other comes from $\widehat{K}$. Note only the meridian $\hat{\mu}$ of $\widehat{K}$ is well-defined, and by definition, it is $\hat{\mu} = q\mu + p\lambda$.

Definition 4.2. If $p = 0$, then $q = 1$ and $\hat{\mu} = \mu$. We can take $\hat{\lambda} = \lambda$. If $(q,p) = (0,1)$, then we take $\hat{\lambda} = -\mu$. If $p, q \neq 0$, then we take $\hat{\lambda} = q_0\mu + p_0\lambda$, where (q_0, p_0) is the unique pair of integers so that the following conditions are true.

(1) $0 \leqslant |p_0| < |p|$ and $p_0 p \leqslant 0$.
(2) $0 \leqslant |q_0| < |q|$ and $q_0 q \leqslant 0$.
(3) $p_0 q - p q_0 = 1$.

In particular, if $(q,p) = (n,1)$, then $\hat{\lambda} = -\mu$.

For a homology class $x\lambda + y\mu$, let $\gamma_{x\lambda+y\mu}$ be the suture consisting of two disjoint simple closed curves representing $\pm(x\lambda + y\mu)$ on $\partial Y(K)$. Furthermore, for $n \in \mathbb{Z}$, define

$$\widehat{\Gamma}_n(q/p) = \gamma_{\hat{\lambda}-n\hat{\mu}} = \gamma_{(p_0-np)\lambda+(q_0-nq)\mu}, \text{ and } \widehat{\Gamma}_\mu(q/p) = \gamma_{\hat{\mu}} = \gamma_{p\lambda+q\mu}.$$

Suppose $(q_n, p_n) \in \{\pm(q_0 - nq, p_0 - np)\}$ such that $q_n \geqslant 0$.

When $\hat{\lambda}$ and $\hat{\mu}$ are understood, we omit the slope q/p and simply write $\widehat{\Gamma}_n$ and $\widehat{\Gamma}_\mu$. When $(q,p) = (1,0)$, we write Γ_n and Γ_μ instead.

Remark 4.3. Since the two components of the suture must be given opposite orientations, the notations $\gamma_{x\lambda+y\mu}$ and $\gamma_{-x\lambda-y\mu}$ represent the same suture on the knot complement $Y(K)$. Our choice makes $q_{n+1} \leqslant q_n$ for $n < -1$ and $q_{n+1} \geqslant q_n$ for $n \geqslant 0$.

Finally, we sketch the proofs of Proposition 1.16 and Theorem 1.12. The essential arguments are proved in the next two subsections.

Proof of Proposition 1.16. Suppose $\hat{\mu} = q\mu + p\lambda$. The set $\mathcal{G}$ of sutures consists of $-\widehat{\Gamma}_n$ for all $n \in \mathbb{N}$ satisfying $q_n > q + 2g(K)$, where $g(K)$ is the Seifert genus of K. For any $\gamma \in \mathcal{G}$. The grading in term (1) is from Theorem 2.21, where the admissible surface S is the Seifert surface of K with minimal genus (up to a stabilization, *cf.* Definition 4.10).

For $\gamma = -\widehat{\Gamma}_n$ in term (2), the image of $f_{K,\gamma}$ is a direct summand of 'middle gradings' of $\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n)$, which is denoted by $\mathcal{I}_+(-\widehat{Y},\widehat{K})$ in Definition 4.21. The isomorphism f_γ is

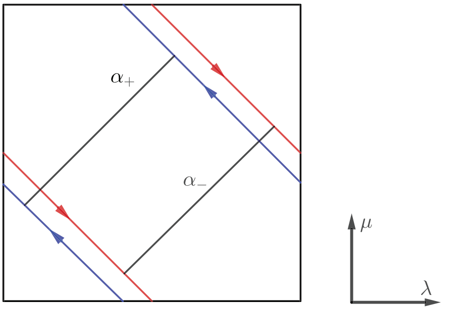

FIGURE 23 Bypass arcs on $\gamma_{(1,-1)}$

defined in Proposition 4.26. It is the restriction of F_n on the corresponding gradings, where F_n is defined in Lemma 4.9 (a special case of Lemma 3.21 for knots).

For $\gamma_1 = -\widehat{\Gamma}_{n_1}, \gamma_2 = -\widehat{\Gamma}_{n_2}$ in term (3), the isomorphism g_{K,γ_1,γ_2} is defined in Lemma 4.16 (see also Remark 4.22). It is the restrictions of bypass maps on corresponding gradings.

Term (4) is from commutative diagrams in Lemma 4.9. $\square$

Proof of Theorem 1.12. We prove this theorem for $-\widehat{Y}$. Suppose $\mu \subset \partial\widehat{Y}(\widehat{K})$ is a simple closed curve such that $|\mu \cdot \lambda| = 1$. Suppose Y is the manifold obtained by Dehn filling along μ and suppose K is the dual knot in Y. By the assumption of the Seifert surface S, we know that K is a null-homologous knot in Y. Moreover, we know that $(\widehat{Y}, \widehat{K})$ is obtained from Y by performing the q/p-surgery along (Y, K) with respect to the canonical framing induced by S. The choice of μ is not important since it will only change the integer p. Then we can apply the construction in Proposition 1.16. In particular, we can use the term (2) of Proposition 1.16 for any $\gamma \in \mathcal{G}$ to define the decomposition. Explicitly, we use $\mathcal{I}_+(\widehat{Y}, \widehat{K})$ to decompose $I^\sharp(\widehat{Y})$. By term (3) and term (4) of Proposition 1.16, this decomposition is well-defined up to isomorphism. $\square$

4.2 | Bypasses on knot complements

Suppose Y is a closed 3-manifold and $K \subset Y$ is a null-homologous knot. Let (μ, λ) be the canonical framing on $Y(K)$ in Definition 4.1. Suppose y_3/x_3 is a surgery slope with $y_3 \geqslant 0$. According to Honda [22, Section 4.3], there are two basic bypasses on the balanced sutured manifold $(Y(K), \gamma_{(x_3,y_3)})$, whose arcs are depicted as in Figure 23. The sutures involved in the bypass triangles were described explicitly in Honda [22, Section 4.4.4].

Definition 4.4. For a surgery slope y_3/x_3 with $y_3 \geqslant 0$, suppose its continued fraction is

$$\frac{y_3}{x_3} = [a_0, a_1, \dots, a_n] = a_0 - \cfrac{1}{a_1 - \cfrac{1}{\dots - \cfrac{1}{a_n}}},$$

where integers $a_i < -1$. If $y_3 > -x_3 > 0$, let

$$\frac{y_1}{x_1} = [a_0, \dots, a_{n-1}] \text{ and } \frac{y_2}{x_2} = [a_0, \dots, a_n + 1].$$

For simplicity, when $a_i = -2$ for integer $i \in (k, n]$ and $a_k \neq -2$, we can set

$$[a_0, \dots, a_n + 1] = [a_0, \dots, a_k + 1].$$

If $-x_3 > y_3 > 0$, we do the same thing for $x_3/(-y_3)$. If $y_3 > x_3 > 0$, we do the same thing for $y_3/(-x_3)$. If $x_3 > y_3 > 0$, we do the same thing for $x_3/(-y_3)$. If $y_3/x_3 = 1/0$, then set $y_1/x_1 = 0/1$ and $y_2/x_2 = 1/(-1)$. If $y_3/x_3 = 0/1$, then set $y_1/x_1 = 1/(-1)$ and $y_1/x_1 = 0/1$. We always require that $y_1 \geqslant 0$ and $y_2 \geqslant 0$.

Remark 4.5. It is straightforward to use induction to verify that for $y_3 > -x_3 > 0$,

$$x_3 = x_1 + x_2 \text{ and } y_3 = y_1 + y_2.$$

The bypass exact triangle in Theorem 2.37 becomes the following.

Proposition 4.6. *Suppose $K \subset Y$ is a null-homologous knot, and suppose the surgery slopes y_i/x_i for $i \in \{1, 2, 3\}$ are defined as in Definition 4.4. Suppose the indices are considered mod 3. Let $\psi^{y_i/x_i}_{+,y_{i+1}/x_{i+1}}$ and $\psi^{y_i/x_i}_{-,y_{i+1}/x_{i+1}}$ be bypass maps from two different bypasses, respectively. Then there are two exact triangles related to $\psi^{y_i/x_i}_{+,y_{i+1}/x_{i+1}}$ and $\psi^{y_i/x_i}_{-,y_{i+1}/x_{i+1}}$, respectively.*

$$\begin{tikzcd}
\underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x_2,y_2)}) \arrow[rr, "\psi^{y_2/x_2}_{\pm,y_3/x_3}"] & & \underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x_3,y_3)}) \arrow[dl, "\psi^{y_3/x_3}_{\pm,y_1/x_1}"] \\
& \underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x_1,y_1)}) \arrow[ul, "\psi^{y_1/x_1}_{\pm,y_2/x_2}"] &
\end{tikzcd}$$

Remark 4.7. Note that there are two different bypasses, which induce two different exact triangles. However, both of them involve the same set of balanced sutured manifolds.

Convention. We will use $\psi^*_{+,*}$ and $\psi^*_{-,*}$ to denote bypass maps with respect to some slopes.

Next, we describe the bypass exact triangles for $\widehat{\Gamma}_n$ and $\widehat{\Gamma}_\mu$ in Definition 4.2.

Proposition 4.8. *Suppose $K \subset Y$ is a null-homologous knot and suppose the pair $(\widehat{Y}, \widehat{K})$ is obtained from (Y, K) by a q/p surgery. Suppose further that the sutures $\widehat{\Gamma}_n$ and $\widehat{\Gamma}_\mu$ are defined as in Definition 4.2. Then there are two exact triangles related to $\psi^*_{+,*}$ and $\psi^*_{-,*}$, respectively.*

$$\begin{tikzcd}
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \arrow[rr, "\psi^{n}_{\pm,n+1}"] & & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) \arrow[dl, "\psi^{n+1}_{\pm,\mu}"] \\
& \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu) \arrow[ul, "\psi^{\mu}_{\pm,n}"] &
\end{tikzcd} \tag{4.1}$$

Proof. If $\widehat{\Gamma}_{n+1} = \gamma_{(x_3,y_3)}$ and $y_3 > -x_3 > 0$ and , then it is straightforward to check that

$$\gamma_{(x_1,y_1)} = \widehat{\Gamma}_\mu \text{ and } \gamma_{(x_2,y_2)} = \widehat{\Gamma}_n,$$

where (x_1, y_1) and $(x_2, y_2$ are defined as in Definition 4.4. Then the exact triangles follows from Proposition 4.6. The similar proof applies to other cases. $\square$

Similar to Lemma 3.21, we have the following proposition.

Lemma 4.9 [16, Section 3]. *Suppose $K \subset Y$ is a null-homologous knot and suppose the pair $(\widehat{Y}, \widehat{K})$ is obtained from (Y, K) by a q/p surgery. Suppose further that the sutures $\widehat{\Gamma}_n$ are defined as in Definition 4.2. Then, there is an exact triangle*

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \arrow[rr] & & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) \arrow[dl, "F_{n+1}"] \\ & \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta) \arrow[ul, "G_n"] & \end{tikzcd} \tag{4.2}$$

where the balanced sutured manifold $(\widehat{Y}(1), \delta)$ is defined as in Remark 2.10.

Furthermore, we have four commutative diagrams related to $\psi^n_{+,n+1}$ and $\psi^n_{-,n+1}$, respectively

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \arrow[rr, "\psi^n_{\pm,n+1}"] & & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) \\ & \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta) \arrow[ul, "G_n"] \arrow[ur, "G_{n+1}"'] & \end{tikzcd}$$

and

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \arrow[rr, "\psi^n_{\pm,n+1}"] \arrow[dr, "F_n"'] & & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) \arrow[dl, "F_{n+1}"] \\ & \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta) & \end{tikzcd}$$

The bypass maps in (4.1) behave well under the gradings on $\underline{\mathrm{SHI}}$ associated to the fixed Seifert surface of K. To provide more details, let us fix a minimal genus Seifert surface S of K.

Convention. We will always assume by default that the Seifert surface S has minimal possible intersections with any suture $\gamma_{(p,q)}$.

Definition 4.10. Suppose $K \subset Y$ is a null-homologous knot and $\gamma_{(x,y)}$ is a suture on $\partial Y(K)$ with $y \geqslant 0$. Suppose further that S is a minimal genus Seifert surface of K. Let $S^{\tau(y)}$ be a negative stabilization of S if y is even and be the original S if y is odd. When the suture $\gamma_{(x,y)}$ is understood, we simply write S^τ. More explicitly, we define a map $\tau : \mathbb{N} \mapsto \{0, -1\}$ as

$$\tau(y) = \begin{cases} 0 & y \text{ is odd} \\ -1 & y \text{ is even} \end{cases}$$

Remark 4.11. It is straightforward to check that $S^\tau \subset (M, \gamma_{(x,y)})$ is admissible. Note that the negative stabilization is with respect to the suture $\gamma_{(x,y)}$ rather than $-\gamma_{(x,y)}$. This is important because later we will incorporate this definition of S^τ with the bypass maps, where the orientations of the sutures are reverse (*cf.* Remark 2.26).

Convention. Note that in Subsection 3.2, we also define another function τ. Since the old definition will no longer be used, and the new tau function serves for the same purpose as the old, we keep using the same notation. From now on, we use the new definition of τ as in Definition 4.10.

Lemma 4.12. *Suppose $K \subset Y$ is a null-homologous knot and $\gamma_{(x,y)}$ is a suture on $\partial Y(K)$ with $y \geqslant 0$. Suppose further that S is a minimal genus Seifert surface of K. Then the maximal and minimal nontrivial gradings of $\underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, S^\tau)$ are*

$$i_{max} = \frac{1}{2}(y - 1 - \tau(y)) + g(S) = \lceil \frac{y-1}{2} \rceil + g(S)$$

and

$$i_{min} = -\frac{1}{2}(y - 1 + \tau(y)) - g(S) = \lceil -\frac{y-1}{2} \rceil - g(S).$$

Proof. The proof is similar to that of Lemma 3.15, though in the current lemma, we can also identify the top and bottom nontrivial gradings by making use of sutured manifold decompositions. Note that we have assumed that the knot complement $Y(K)$ is irreducible in the convention after Definition 4.1, and S is a minimal genus Seifert surface of K, so the decomposition of $(Y(K), \gamma)$ along S and $-S$ are both taut.

When y is odd, we have $S^\tau = S$. Then it follows directly from Theorem 2.21 that

$$i_{\max} = \frac{y-1}{2} + g(S) \text{ and } i_{\min} = -\frac{y-1}{2} - g(S).$$

When y is even, we have $S^\tau = S^-$. Then it follows directly from Theorem 2.21 that

$$i_{\max} = \frac{y}{2} + g(S).$$

To figure out the grading $i_{\min}$, note that

$$\begin{aligned}\underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, S^-, -i_{\max}) &= \underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, -(S^-), i_{\max}) \\ &= \underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, (-S)^+, i_{\max}) \\ &= 0.\end{aligned}$$

The last equality follows from Lemma 2.25 and the fact that the positive stabilization on $(-S)$ with respect to $\gamma_{(x,y)}$ becomes a negative one with respect to $-\gamma_{(x,y)}$ (*cf.* Remark 2.26).

We also have

$$\begin{aligned}\underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, S^-, 1 - i_{\max}) &= \underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, -(S^-), i_{\max} - 1) \\ &= \underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, (-S)^+, i_{\max} - 1) \\ &= \underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, (-S)^-, i_{\max})\end{aligned}$$

The last equality follows from Theorem 2.27. From Lemma 2.25 and Theorem 2.21, we have

$$\underline{\mathrm{SHI}}(-Y(K), -\gamma_{(x,y)}, (-S)^-, i_{\max}) \cong \underline{\mathrm{SHI}}(-M', -\gamma') \neq 0,$$

where (M', γ') is the taut balanced sutured manifold obtained from $(Y(K), \gamma_{(x,y)})$ by decomposing along $-S$. Hence we conclude that

$$i_{\min} = 1 - i_{\max} = 1 - \frac{y}{2} - g(S). \qquad \square$$

Definition 4.13. For any integer $y \in \mathbb{N}$, define

$$i^y_{\max} = \lceil \frac{y-1}{2} \rceil + g(S), \text{ and } i^y_{\min} = \lceil -\frac{y-1}{2} \rceil - g(S).$$

For the suture $\widehat{\Gamma}_n = \gamma_{(p_n, q_n)}$, define

$$\hat{i}^n_{max} = i^{q_n}_{max} \text{ and } \hat{i}^n_{min} = i^{q_n}_{min}.$$

Convention. Note that we use the similar notations as in Subsection 3.2, while from now on, we use the new definitions of $i^y_{\max}$ and $i^y_{\min}$ as in Definition 4.13. We will use $i^*_{\max}, \hat{i}^*_{\max}$ and $i^*_{min}, \hat{i}^*_{\min}$ to denote the maximal and minimal gradings for the slope specified by $*$.

In [42, Section 5], a graded version of the bypass exact triangles in Proposition 4.6 is proved, which is similar to Lemma 3.18.

Proposition 4.14 [42, Proposition 5.5]. *Suppose $K \subset Y$ is a null-homologous knot and suppose the pair $(\widehat{Y}, \widehat{K})$ is obtained from (Y, K) by a q/p surgery. Suppose further that the sutures $\widehat{\Gamma}_n$ and $\widehat{\Gamma}_\mu$ are defined as in Definition 4.2 and S is a minimal genus Seifert surface of K. Then the followings hold. Note that the grading shift notation comes from Definition 3.17.*

(1) *For $n \in \mathbb{Z}$ so that $q_{n+1} = q_n + q$, that is, $n \geqslant 0$, there are two bypass exact triangles:*

$$\begin{tikzcd}
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau)[\hat{i}^{n+1}_{\min} - \hat{i}^n_{\min}] \arrow[rr, "\psi^n_{+,n+1}"] & & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau) \arrow[dll, "\psi^{n+1}_{+,\mu}"] \\
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu, S^\tau)[\hat{i}^{n+1}_{\max} - \hat{i}^\mu_{\max}] \arrow[u, "\psi^\mu_{+,n}"] & &
\end{tikzcd}$$

and

$$\begin{tikzcd}
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau)[\hat{i}^{n+1}_{\max} - \hat{i}^n_{\max}] \arrow[rr, "\psi^n_{-,n+1}"] & & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau) \arrow[dll, "\psi^{n+1}_{-,\mu}"] \\
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu, S^\tau)[\hat{i}^{n+1}_{\min} - \hat{i}^\mu_{\min}] \arrow[u, "\psi^\mu_{-,n}"] & &
\end{tikzcd}$$

(2) *For $n \in \mathbb{Z}$ so that $q_{n+1} = q_n - q$, that is, $n < -1$, there are two bypass exact triangles:*

$$
\begin{tikzcd}
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau) \arrow[r, "\psi^n_{+,n+1}"] & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau)[\hat{\imath}^n_{\max} - \hat{\imath}^{n+1}_{\max}] \arrow[dl, "\psi^{n+1}_{+,\mu}"] \\
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu, S^\tau)[\hat{\imath}^n_{\min} - \hat{\imath}^\mu_{\min}] \arrow[u, "\psi^\mu_{+,n}"] &
\end{tikzcd}
$$

and

$$
\begin{tikzcd}
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau) \arrow[r, "\psi^n_{-,n+1}"] & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau)[\hat{\imath}^n_{\min} - \hat{\imath}^{n+1}_{\min}] \arrow[dl, "\psi^{n+1}_{-,\mu}"] \\
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu, S^\tau)[\hat{\imath}^n_{\max} - \hat{\imath}^\mu_{\max}] \arrow[u, "\psi^\mu_{-,n}"] &
\end{tikzcd}
\tag{4.2}
$$

(3) *For $n \in \mathbb{Z}$ so that $q_{n+1} + q_n = q$, that is, $n = -1$, there are two bypass exact triangles:*

$$
\begin{tikzcd}
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau)[\hat{\imath}^\mu_{\min} - \hat{\imath}^n_{\min}] \arrow[r, "\psi^n_{+,n+1}"] & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau)[\hat{\imath}^\mu_{\max} - \hat{\imath}^{n+1}_{\max}] \arrow[dl, "\psi^{n+1}_{+,\mu}"] \\
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu, S^\tau) \arrow[u, "\psi^\mu_{+,n}"] &
\end{tikzcd}
\tag{4.4}
$$

and

$$
\begin{tikzcd}
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau)[\hat{\imath}^\mu_{\max} - \hat{\imath}^n_{\max}] \arrow[r, "\psi^n_{-,n+1}"] & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau)[\hat{\imath}^\mu_{\min} - \hat{\imath}^{n+1}_{\min}] \arrow[dl, "\psi^{n+1}_{-,\mu}"] \\
\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu, S^\tau) \arrow[u, "\psi^\mu_{-,n}"] &
\end{tikzcd}
\tag{4.5}
$$

Furthermore, all maps involved in the above bypass exact triangles are grading preserving.

Remark 4.15. The above proposition can be understood by Remark 3.19. Alternatively, we can understand the above proposition by the following method, which is inspired by the curve invariant introduced by Hanselman, Rasmussen, and Waston [19, 20].

(1) Consider the lattice $\mathbb{Z}^2 \subset \mathbb{R}^2$. A surgery slope $y/x \in \mathbb{Q} \cup \{\infty\}$ corresponds to a straight arc connecting two lattice points in $\mathbb{Z}^2$.
(2) Suppose the sutures $\gamma_{(x_1,y_1)}, \gamma_{(x_2,y_2)}$, and $\gamma_{(x_3,y_3)}$ are defined as in Definition 4.4. Then it is easy to see the arcs corresponding to these three sutures bound a triangle containing no lattice point in the interior. There are two different triangles up to translation, which correspond to two different bypass triangles. All bypass maps are clockwise in $\mathbb{R}^2$. Rotation around the origin by 180° will switch the roles of $\psi^*_{+,*}$ and $\psi^*_{-,*}$.
(3) The height of the middle point of the straight arc indicates the grading before stabilization (so there are gradings of half integers). If the top endpoints of two arcs are the same, the grading shift is about $\hat{\imath}^*_{\min}$. If the bottom endpoints of two arcs are the same, the grading shift is about $\hat{\imath}^*_{\max}$.

4.3 | Decomposing framed instanton Floer homology

In this subsection, we prove term (2) of Proposition 1.16. Throughout this subsection, let $K \subset Y$ be a null-homologous knot and let the pair $(\widehat{Y}, \widehat{K})$ be obtained from (Y, K) by a q/p surgery with $q > 0$. Suppose the sutures $\widehat{\Gamma}_n = \gamma_{(p_n, q_n)}$ and $\widehat{\Gamma}_\mu$ are defined as in Definition 4.2 and suppose S is a minimal genus Seifert surface of K. The stabilization S^τ of S is chosen according to Definition 4.10. The maximal and minimal gradings of the involved sutured instanton Floer homology are described in Lemma 4.12. For any $i \in \mathbb{Z}$, let

$$\psi^{n,i}_{\pm,n+1} = \psi^n_{\pm,n+1}|_{\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, i)}$$

be the restriction.

Lemma 4.16. *Suppose $n \in \mathbb{Z}$ so that $q_{n+1} = q_n + q$, that is, $n \geqslant 0$. Then the map*

$$\psi^{n,i}_{+,n+1} : \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, i) \to \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau, i - \hat{i}^n_{min} + \hat{i}^{n+1}_{min})$$

is an isomorphism if $i \leqslant \hat{i}^n_{max} - 2g(S)$. Similarly, the map

$$\psi^{n,i}_{-,n+1} : \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, i) \to \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau, i - \hat{i}^n_{max} + \hat{i}^{n+1}_{max})$$

is an isomorphism if $i \geqslant \hat{i}^n_{min} + 2g(S)$.

Proof. The proof the lemma is similar to that of Lemma 3.20. □

Lemma 4.17. *For any $n \in \mathbb{Z}$ so that $q_{n+1} - q = q_n \geqslant 2g$, the map*

$$G_n : \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta) \to \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n)$$

defined as in (4.2) is the zero map.

Proof. The proof of the lemma is similar to that of Lemma 3.22. □

Corollary 4.18. *We have*

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu) \geqslant \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta).$$

Proof. Since for a large enough integer n, we have $G_n = 0$, we know that

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta) = \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) - \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n).$$

Then the corollary follows directly from Proposition 4.8. □

Corollary 4.19. *Suppose L_1 is a non-empty link in $\widehat{Y}$ that is disjoint from $\widehat{K}$. Let $L_2 = L_1 \cup \widehat{K}$. Consider the link complements $-\widehat{Y}(L_1)$ and $-\widehat{Y}(L_2)$. For the link complement, let $\widehat{\Gamma}_\mu$ be the suture* *consisting of two meridians for each component of the link. We have*

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-\widehat{Y}(L_2), -\widehat{\Gamma}_\mu) \geqslant 2 \cdot \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-\widehat{Y}(L_1), -\widehat{\Gamma}_\mu).$$

Proof. The same argument to prove Corollary 4.18 can be applied verbatim to verify

$$\dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-\widehat{Y}(L_2), -\widehat{\Gamma}_\mu) \geqslant \underline{\mathrm{SHI}}(-\widehat{Y}(L_1)(1), -\widehat{\Gamma}_\mu \cup -\delta),$$

where $\widehat{Y}(L_1)(1)$ is obtained from $-\widehat{Y}(L_1)$ by removing a 3-ball disjoint from L_1, and δ is a simple closed curve on the new spherical boundary component. From [4, Lemma 4.14], we have

$$\underline{\mathrm{SHI}}(-\widehat{Y}(L_1)(1), -\widehat{\Gamma}_\mu \cup -\delta) = 2 \cdot \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-\widehat{Y}(L_1), -\widehat{\Gamma}_\mu).$$ □

Lemma 4.20. *Suppose $n \in \mathbb{Z}$ satisfies $q_{n+1} - q = q_n \geqslant q + 2g(S)$, and suppose $i, j \in \mathbb{Z}$ with*

$$\hat{i}^n_{min} + 2g(S) \leqslant i, j \leqslant \hat{i}^n_{max} - 2g(S) \text{ and } i - j = q.$$

Then we have

$$\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, i) \cong \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, j).$$

Proof. Since $i \leqslant \hat{i}^n_{\max} - 2g(S)$, by Lemma 4.16, we know

$$\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, i) \cong \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau, i - \hat{i}^n_{\min} + \hat{i}^{n+1}_{\min}).$$

Similarly, since $j \geqslant \hat{i}^n_{\min} + 2g(S)$, we know that

$$\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, j) \cong \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau, j - \hat{i}^n_{\max} + \hat{i}^{n+1}_{\max}).$$

Note also that

$$\begin{aligned} i - \hat{i}^n_{\min} + \hat{i}^{n+1}_{\min} &= j - \hat{i}^n_{\max} + \hat{i}^{n+1}_{\max} + q + (\hat{i}^n_{\max} - \hat{i}^n_{\min}) - (\hat{i}^{n+1}_{\max} - \hat{i}^{n+1}_{\min}) \\ &= j - \hat{i}^n_{\max} + \hat{i}^{n+1}_{\max} + q + (q_n - 1 + 2g(S)) - (q_{n+1} - 1 + 2g(S)) \\ &= j - \hat{i}^n_{\max} + \hat{i}^{n+1}_{\max}. \end{aligned}$$

Hence we obtain the desired result. □

Definition 4.21. Suppose $n \in \mathbb{Z}$ satisfies $q_{n+1} - q = q_n \geqslant q + 2g(S)$. Define

$$\mathcal{I}_+(-\widehat{Y}, \widehat{K}, i) = \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, \hat{i}^n_{\max} - 2g(S) - i),$$

and

$$\mathcal{I}_+(-\widehat{Y}, \widehat{K}) = \bigoplus_{i=0}^{q-1} \mathcal{I}_+(-\widehat{Y}, \widehat{K}, i).$$

Remark 4.22. From Lemma 4.16, the definition of $\mathcal{I}_+(-\widehat{Y},\widehat{K})$ is independent of the choice of the integer n satisfying the required condition. Also, by Lemma 4.20, the definition of $\mathcal{I}_+(-\widehat{Y},\widehat{K})$ would be the same (up to a $\mathbb{Z}_q$ grading shift) if we consider arbitrary q consecutive gradings within the range $[\hat{i}^n_{\min}+2g(S),\hat{i}^n_{\max}-2g(S)]$.

Next, our goal is to show that there is an isomorphism

$$\mathcal{I}_+(-\widehat{Y},\widehat{K})\cong \underline{\mathrm{SHI}}(-\widehat{Y}(1),-\delta).$$

To do so, we first introduce some notations for performing computations.

Definition 4.23. Suppose $n\in\mathbb{Z}$. The direct sum of some consecutive gradings of

$$\underline{\mathrm{SHI}}(-Y(K),-\widehat{\Gamma}_n,S^\tau)$$

is called a *block*. For a block A, the number of gradings involved is called the *size* of A.

Example 4.24. Suppose $n\in\mathbb{Z}$ satisfies $q_n\geqslant q+2g(S)$. Let A,B,C and D be the blocks consisting of the top $2g(S)$ gradings, the next q gradings, the next $q_n-q-2g(S)$ gradings, and the last $2g(S)$ gradings of $\underline{\mathrm{SHI}}(-Y(K),-\widehat{\Gamma}_n,S^\tau)$, respectively. We write

$$\underline{\mathrm{SHI}}(-Y(K),-\widehat{\Gamma}_n,S^\tau)=\begin{pmatrix}A\\B\\C\\D\end{pmatrix}.$$

From Definition 4.21, we know that $\mathcal{I}_+(-\widehat{Y},\widehat{K})$ is itself a block and in fact

$$\mathcal{I}_+(-\widehat{Y},\widehat{K})=B.$$

Also, we can write

$$\underline{\mathrm{SHI}}(-Y(K),-\widehat{\Gamma}_n,S^\tau)=\begin{pmatrix}A\\E\\F\\D\end{pmatrix},$$

where E and F are of size $(q_n-q-2g(S))$ and q, respectively. By comparing the gradings, we have

$$\begin{pmatrix}B\\C\end{pmatrix}=\begin{pmatrix}E\\F\end{pmatrix}$$

A priori, we do not have $B=E$ and $C=F$ since they have different sizes. However, when putting together, the total size of B and C equals that of E and F.

Lemma 4.25. *Let $\mathcal{I}_+(-\widehat{Y},\widehat{K})$ be defined as in Definition 4.21. Then we have*

$$\dim_{\mathbb{C}}\mathcal{I}_+(-\widehat{Y},\widehat{K})=\dim_{\mathbb{C}}\underline{\mathrm{SHI}}(-\widehat{Y}(1),-\delta).$$

Proof. Suppose $n \in \mathbb{Z}$ satisfies $q_n \geqslant q + 2g(S)$. We can apply Proposition 4.14. Using blocks, we have the following. (There is no enough room for writing down the whole notation for $\underline{\mathrm{SHI}}$, so we use the sutures to denote them.)

$$
\begin{array}{cccccccc}
\text{size} & -\widehat{\Gamma}_\mu & \xrightarrow{\psi^{\mu}_{+,n}} & -\widehat{\Gamma}_n & \xrightarrow{\psi^{n}_{+,n+1}} & -\widehat{\Gamma}_{n+1} & \xrightarrow{\psi^{n+1}_{+,\mu}} & -\widehat{\Gamma}_\mu \\
q & G & & & & X_1 & & G \\
2g(S) & H & & A & & X_2 & & H \\
q_n - q - 2g(S) & & & E & & X_3 & & \\
q & & & F & & X_4 & & \\
2g(S) & & & D & & X_5 & &
\end{array}
$$

From the exactness, we know that

$$X_1 = G,\ X_3 = E,\ X_4 = F,\ \text{and}\ X_5 = D.$$

There is another bypass exact triangle, and similarly we have

$$
\begin{array}{cccccccc}
\text{size} & -\widehat{\Gamma}_\mu & \xrightarrow{\psi^{\mu}_{-,n}} & -\widehat{\Gamma}_n & \xrightarrow{\psi^{n}_{-,n+1}} & -\widehat{\Gamma}_{n+1} & \xrightarrow{\psi^{n+1}_{-,\mu}} & -\widehat{\Gamma}_\mu \\
2g(S) & & & A & & A & & \\
q & & & B & & B & & \\
q_n - q - 2g(S) & & & C & & C & & \\
2g(S) & I & & D & & X_6 & & I \\
q & J & & & & J & & J
\end{array}
$$

Comparing the two expressions of $\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau)$, we have

$$\begin{pmatrix} G \\ X_2 \\ E \\ F \\ D \end{pmatrix} = \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau) = \begin{pmatrix} A \\ B \\ C \\ X_6 \\ J \end{pmatrix}.$$

Taking sizes into consideration, we know that

$$\begin{pmatrix} G \\ X_2 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix},\ E = C,\ \text{and}\ \begin{pmatrix} F \\ D \end{pmatrix} = \begin{pmatrix} X_6 \\ J \end{pmatrix}.$$

Thus, we know that

$$\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}, S^\tau) = \begin{pmatrix} A \\ B \\ E \\ F \\ D \end{pmatrix}.$$

Comparing this expression with the expression of $\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau)$ in Example 4.24, we have

$$\begin{aligned}\dim_{\mathbb{C}} \mathcal{I}_+(-\widehat{Y}, \widehat{K}) &= \dim_{\mathbb{C}} B \\ &= \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) - \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \\ &= \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta),\end{aligned}$$

where the last equality follows from Lemma 4.17. □

Proposition 4.26. *Suppose $n \in \mathbb{Z}$ satisfies $q_{n+1} - q = q_n \geqslant q + 2g(S)$. Then the map F_n restricted to $\mathcal{I}_+(-\widehat{Y}, \widehat{K})$ is an isomorphism, that is,*

$$F_n|_{\mathcal{I}_+(-\widehat{Y},\widehat{K})} : \mathcal{I}_+(-\widehat{Y}, \widehat{K}) \xrightarrow{\cong} \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta).$$

Proof. It suffices to show that the restriction of F_n is surjective. Since $q_n \geqslant q + 2g(S)$, we have $q_{n-1} = q_n - q \geqslant 2g(S)$. By Lemma 4.17, we know that $G_{n-1} = 0$. By exactness in (4.2), the map F_n is surjective. Then it suffices to show that F_n remains surjective when restricted to $\mathcal{I}_+(-\widehat{Y}, \widehat{K})$. For any $x \in \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta)$, let $y \in \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n)$ be an element so that $F_n(y) = x$. Suppose

$$y = \sum_{j \in \mathbb{Z}} y_j, \text{ where } y_j \in \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, j).$$

For any y_j, we want to find $y_j' \in \mathcal{I}_+(-\widehat{Y}, \widehat{K})$ so that $F_n(y_j) = F_n(y_j')$.

To do this, we first assume that $j > \hat{i}^n_{\max} - 2g(S)$. Then there exists an integer m so that

$$\hat{i}^n_{\max} - 2g(S) - q + 1 \leqslant j - mq \leqslant \hat{i}^n_{\max} - 2g(S).$$

We can take

$$y_j' = (\psi^{n,j-mq}_{-,n+1})^{-1} \circ \cdots \circ (\psi^{n,\hat{i}^{n+m}_{\max} - \hat{i}^n_{\max} + j - mq}_{-,n+m})^{-1} \circ \psi^{n+m-1}_{+,n+m} \circ \ldots \circ \psi^n_{+,n+1}(y_j). \tag{4.6}$$

From Lemma 4.16, all the negative bypass maps involved in (4.6) are isomorphisms so the inverses exist. Also, we have

$$y_j' \in \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, j - mq) \subset \mathcal{I}_+(-\widehat{Y}, \widehat{K}).$$

Finally, from Lemma 4.9, we know that $F_n(y_j') = F_n(y_j)$.

For

$$j \in [\hat{i}^n_{\max} - 2g(S) - q - 1, \hat{i}^n_{\max} - 2g(S)],$$

we can simply take $y_j' = y_j$.

For $j \leqslant \hat{i}^n_{\max} - 2g(S) - q - 1$, we can pick y_j' similarly as in (4.6), while switching the roles of $\psi^*_{+,*}$ and $\psi^*_{-,*}$ in (4.6).

In summary, we can take

$$y' = \sum_{j\in\mathbb{Z}} y'_j \in \mathcal{I}_+(-\hat{Y}, \hat{K}) \text{ with } F_n(y') = F_n(y) = x.$$

Hence the restriction of F_n is still surjective, and we obtain the desired result. □

In Definition 4.21, we use a large enough integer n to define $\mathcal{I}_+(-\hat{Y}, \hat{K})$. We can also use a small enough integer n to define a vector space similar to $\mathcal{I}_+(-\hat{Y}, \hat{K})$. Recall

$$\hat{\Gamma}_n = \gamma_{(p_n, q_n)}$$

is defined as in Definition 4.2 and q_n is chosen to be always non-negative.

Definition 4.27. Suppose $n \in \mathbb{Z}$ satisfies $q_{n-1} - q = q_n \geqslant q + 2g(S)$. Define

$$\mathcal{I}_-(-\hat{Y}, \hat{K}, i) = \underline{\mathrm{SHI}}(-Y(K), -\hat{\Gamma}_n, S^\tau, \hat{i}^n_{\max} - 2g(S) - i),$$

and

$$\mathcal{I}_-(-\hat{Y}, \hat{K}) = \bigoplus_{i=0}^{q-1} \mathcal{I}_-(-\hat{Y}, \hat{K}, i).$$

The arguments for $\mathcal{I}_-(-\hat{Y}, \hat{K})$ are similar to those for $\mathcal{I}_+(-\hat{Y}, \hat{K})$. We sketch them as follows.

Lemma 4.28. *Suppose $n \in \mathbb{Z}$ satisfies $q_{n-1} - q = q_n$, that is, $n < -1$. Then the map*

$$\psi^{n-1, i+\hat{i}^{n-1}_{max} - \hat{i}^n_{max}}_{+,n} : \underline{\mathrm{SHI}}(-Y(K), -\hat{\Gamma}_{n-1}, S^\tau, i + \hat{i}^{n-1}_{max} - \hat{i}^n_{max}) \to \underline{\mathrm{SHI}}(-Y(K), -\hat{\Gamma}_n, S^\tau, i)$$

is an isomorphism if $i \leqslant \hat{i}^n_{max} - 2g(S)$. Similarly, the map

$$\psi^{n-1, i-\hat{i}^n_{min} + \hat{i}^{n-1}_{min}}_{-,n} : \underline{\mathrm{SHI}}(-Y(K), -\hat{\Gamma}_{n-1}, S^\tau, i - \hat{i}^n_{min} + \hat{i}^{n-1}_{min}) \to \underline{\mathrm{SHI}}(-Y(K), -\hat{\Gamma}_n, S^\tau, i)$$

is an isomorphism if $i \geqslant \hat{i}^n_{min} + 2g(S)$.

Proof. The proof is similar to the proof of Lemma 3.20. □

Lemma 4.29. *For any $n \in \mathbb{Z}$ so that $q_{n-1} - q = q_n \geqslant 2g$, the map*

$$F_n : \underline{\mathrm{SHI}}(-Y(K), -\hat{\Gamma}_n) \to \underline{\mathrm{SHI}}(-\hat{Y}(1), -\delta)$$

defined as in (4.2) is the zero map.

Proof. If it is not, then let $j_{\max} \in \mathbb{Z}$ be the maximal index j so that there exists

$$x \in \underline{\mathrm{SHI}}(-Y(K), -\hat{\Gamma}_n, S^\tau, j)$$

with $F_n(x) \neq 0$. Since $q_n \geqslant 2g$, by Lemma 4.12, we know that either

$$j_{\max} \leqslant \hat{i}^n_{\max} - 2g(S) \text{ or } j_{\max} \geqslant \hat{i}^n_{\min} + 2g(S).$$

Suppose, without loss of generality, that $j_{\max} \geqslant \hat{i}^n_{\min} + 2g(S)$ and

$$x \in \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, j_{\max})$$

satisfying $F_n(x) \neq 0$. By Lemma 4.28, $\psi_{+,n}^{n-1, j_{\max}+\hat{i}^{n-1}_{\max}-\hat{i}^n_{\max}}$ is an isomorphism, and we can take

$$y = \psi_{-,n}^{n-1, j_{\max}+\hat{i}^{n-1}_{\max}-\hat{i}^n_{\max}} \circ (\psi_{+,n}^{n-1, j_{\max}+\hat{i}^{n-1}_{\max}-\hat{i}^n_{\max}})^{-1}(x).$$

By Lemma 4.9, we know that

$$F_n(y) = F_n(x) \neq 0 \text{ and } y \in \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, j_{\max} + q).$$

This is a contradiction. □

Lemma 4.30. *Suppose $n \in \mathbb{Z}$ satisfies $q_{n-1} - q = q_n \geqslant q + 2g(S)$, and suppose $i, j \in \mathbb{Z}$ satisfying*

$$\hat{i}^n_{min} + 2g(S) \leqslant i, j \leqslant \hat{i}^n_{max} - 2g(S) \text{ and } i - j = q.$$

Then we have

$$\underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, i) \cong \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, j).$$

Proof. The proof is similar to the proof of Lemma 4.20. □

Lemma 4.31. *Let $\mathcal{I}_-(-\widehat{Y}, \widehat{K})$ be defined as in Definition 4.27. Then we have*

$$\dim_{\mathbb{C}} \mathcal{I}_-(-\widehat{Y}, \widehat{K}) = \dim_{\mathbb{C}} \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta).$$

Proof. The proof is similar to the proof of Lemma 4.25. □

Proposition 4.32. *Suppose $n \in \mathbb{Z}$ satisfies $q_{n-1} - q = q_n \geqslant q + 2g(S)$. Let Π_n be the projection*

$$\Pi_n : \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \to \mathcal{I}_-(-\widehat{Y}, \widehat{K}).$$

Then we have an isomorphism

$$\Pi_n \circ G_n : \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta) \xrightarrow{\cong} \mathcal{I}_-(-\widehat{Y}, \widehat{K}).$$

Proof. It suffices to show that $\Pi_n \circ G_n$ is injective. We assume it is not true and derive a contradiction. By assumption, there exists

$$x \neq 0 \in \underline{\mathrm{SHI}}(-\widehat{Y}(1), -\delta) \text{ with } \Pi_n \circ G_n(x) = 0.$$

Write

$$y = G_n(x) = \sum_{j\in\mathbb{Z}} y_j,\ \text{where}\ y_j \in \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n, S^\tau, j).$$

From Lemmas 4.9 and 4.29, we know that G_n is injective, and hence $y \neq 0$. From the assumption, we know that

$$y_j = 0 \text{ for } \hat{i}^n_{\max} - 2g(S) \geqslant j \geqslant \hat{i}^n_{\max} - 2g(S) - q + 1.$$

Also write

$$z = G_{n-1}(x) = \sum_{j\in\mathbb{Z}} z_j,\ \text{where}\ z_j \in \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n-1}, S^\tau, j).$$

From Lemma 4.9, we know that

$$\psi^{n-1}_{-,n}(z) = y = \psi^{n-1}_{+,n}(z).$$

Suppose $j_{\min}$ is the minimal grading j so that

$$j > \hat{i}^n_{\max} - 2g(S) \text{ and } y_j \neq 0.$$

Then we know that

$$y_{j_{\min}-q} = 0 \text{ and } j_{\min} - q \geqslant \hat{i}^n_{\min} + 2g(s).$$

Hence by Lemma 4.28, we know that

$$\begin{aligned} y_{j_{\min}} &= \psi^{n-1, j_{\min}-\hat{i}^n_{\min}+\hat{i}^{n-1}_{\min}}_{-,n}(z_{j_{\min}-\hat{i}^n_{\min}+\hat{i}^{n-1}_{\min}}) \\ &= \psi^{n-1, j_{\min}-\hat{i}^n_{\min}+\hat{i}^{n-1}_{\min}}_{-,n} \circ (\psi^{n-1, j_{\min}-\hat{i}^n_{\min}+\hat{i}^{n-1}_{\min}}_{+,n})^{-1}(y_{j_{\min}-q}) \\ &= 0. \end{aligned}$$

This implies that $y_j = 0$ for all $j \geqslant \hat{i}^n_{\max} - 2g(S) - q + 1$. Similarly we can prove that $y_j = 0$ for all $j < \hat{i}^n_{\max} - 2g(S) - q + 1$, and $y = 0$, which contradicts the injectivity of G_n. $\square$

4.4 | Commutative diagrams for bypass maps

In this subsection, we show there are some commutative diagrams for bypass maps.

Lemma 4.33 [42, Corollary 2.20]. *For any surgery slope q/p, consider the bypass maps $\psi^*_{+,*}$ and $\psi^*_{-,*}$ in Proposition 4.8. For any integer $n \in \mathbb{Z}$, we have the following commutative diagram.*

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \arrow[r, "\psi^n_{-,n+1}"] \arrow[d, "\psi^n_{+,n+1}"'] & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) \arrow[d, "\psi^{n+1}_{+,n+2}"] \\ \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) \arrow[r, "\psi^{n+1}_{-,n+2}"] & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+2}) \end{tikzcd} \tag{4.7}$$

Proof. In Subsection 2.3, we reinterpreted bypass maps by contact gluing maps. So the composition of bypass maps becomes the composition of contact gluing maps. To verify the commutative diagram, it suffices to verify that two contact structures coming from different bypasses are actually the same. Thus, it is free to change the basis of $H_1(T^2)$. It suffices to verify a special case $q/p = 1/0$ and $n = 0$. Then it follows from [22, Lemma 4.14] that the contact structures are the same. $\square$

Lemma 4.34. *For any surgery slope q/p, consider the bypass maps $\psi^*_{+,*}$ and $\psi^*_{-,*}$ in Proposition 4.8. For any $n \in \mathbb{Z}$, we have two commutative diagrams*

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \arrow[rr, "\psi^n_{-,n+1}"] \arrow[dr, "\psi^n_{+,\mu}"'] & & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) \arrow[dl, "\psi^{n+1}_{+,\mu}"] \\ & \underline{\mathrm{SHI}}(-\widehat{Y}(K), -\widehat{\Gamma}_\mu) & \end{tikzcd} \tag{4.8}$$

and

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_n) \arrow[rr, "\psi^n_{-,n+1}"] & & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_{n+1}) \\ & \underline{\mathrm{SHI}}(-\widehat{Y}(K), -\widehat{\Gamma}_\mu) \arrow[ul, "\psi^\mu_{+,n}"] \arrow[ur, "\psi^\mu_{+,n+1}"'] & \end{tikzcd} \tag{4.9}$$

The similar commutative diagrams hold if we switch the roles of $\psi^*_{+,*}$ and $\psi^*_{-,*}$.

Remark 4.35. The bypass maps in Lemma 4.34 are from different bypass exact triangles.

Proof of Lemma 4.34. Similar to the proof of Lemma 4.33, this lemma follows from Honda's classification of tight contact structures on $T^2 \times I$ [22, Lemma 4.14]. $\square$

Corollary 4.36. *For any surgery slope q/p, consider the bypass maps $\psi^*_{+,*}$ and $\psi^*_{-,*}$ in Proposition 4.8. For any $i, j \in \mathbb{Z}$, we have the following commutative diagrams related to $\psi^*_{+,*}$ and $\psi^*_{-,*}$, respectively.*

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu) \arrow[r, "\psi^\mu_{\pm,j}"] \arrow[d, "\psi^\mu_{\pm,i}"'] & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_j) \arrow[d, "\psi^j_{\pm,\mu}"] \\ \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_i) \arrow[r, "\psi^i_{\pm,\mu}"] & \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu) \end{tikzcd} \tag{4.10}$$

Proof. The commutative diagram related to $\psi^*_{+,*}$ follows from (4.8) and (4.9). Explicitly, for $i = j + 1$, both compositions of maps are equal to

$$\psi^{j+1}_{+,\mu} \circ \psi^n_{-,j+1} \circ \psi^\mu_{+,j}.$$

The other commutative diagram follows from Lemma 4.34 similarly. $\square$

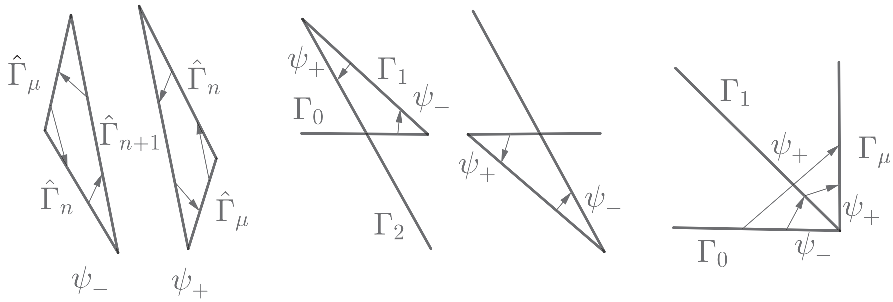

FIGURE 24 Left, bypass maps; middle, illustration of (4.7); right, illustration of (4.8)

Corollary 4.37. *For any surgery slope q/p, consider the bypass maps $\psi^*_{+,*}$ and $\psi^*_{-,*}$ in Proposition 4.8. For any $n \in \mathbb{Z}$, we have*

$$\psi^n_{+,\mu} \circ \psi^\mu_{-,n} = \psi^n_{-,\mu} \circ \psi^\mu_{+,n} = 0$$

and

$$\psi^\mu_{+,n} \circ \psi^n_{+,\mu} = \psi^\mu_{-,n} \circ \psi^n_{-,\mu} = 0$$

Proof. By Lemma 4.34 and the exactness, we have

$$\psi^n_{+,\mu} \circ \psi^\mu_{-,n} = \psi^{n+1}_{+,\mu} \circ \psi^n_{-,n+1} \circ \psi^\mu_{-,n} = 0.$$

Other arguments follow from Lemma 4.34 and the exactness similarly. □

Remark 4.38. The above commutative diagrams can be illustrated by the method described in Remark 4.15. The illustration of the special cases in the proofs is shown in Figure 24. Note that vector spaces are denoted by their sutures (we omit the minus signs), and all maps are bypass maps. They are grading preserving and commute with F_* and G_* by Proposition 4.14 and Lemma 4.9, respectively.

4.5 | The stabilization of integral surgeries

Throughout this subsection, suppose $K \subset Y$ is a null-homologous knot and S is a minimal genus Seifert surface of K. The stabilization S^τ of S is chosen in Definition 4.10. Since we might work with different surgery slopes, we will use the notation $\gamma_{(p,q)}$ in Definition 4.2 to denote the suture on the knot complement. The maximal and minimal gradings of $\underline{\mathrm{SHI}}(-Y(K), -\gamma_{(p,q)}, S^\tau)$ are described explicitly in Lemma 4.12 and we write them as

$$i^q_{\max} = \lceil \frac{q-1}{2} \rceil + g(S) \text{ and } i^q_{\min} = \lceil -\frac{q-1}{2} \rceil - g(S).$$

Note that they are independent of p.

Since we will deal with different surgeries in the current subsection, we will write out the surgery slope explicitly: for the 3-manifold obtained by the q/p-surgery, we write $\widehat{Y}_{q/p}$. For the special class of sutures, we write $\widehat{\Gamma}_n(q/p)$ instead of $\widehat{\Gamma}_n$. For bypass maps, we write $\psi^{n+1}_{\pm,n}(q/p)$. However, if $q/p = 1/0$, we still omit it from the notation, then in this case, we simply write $\widehat{\Gamma}_n$ as Γ_n.

In this subsection, we deal with large integral surgeries. In this case, we know that $\Gamma_n = \gamma_{(-1,n)}$. Hence we have the following by Lemma 4.20 and Proposition 4.26. (Note $(p,q) = (0,1)$.)

Lemma 4.39. *For any $n > 2g(S)$ and $i \in \mathbb{Z}$ so that*

$$\lceil -\frac{n-1}{2} \rceil + g(S) \leqslant i \leqslant \lceil \frac{n-1}{2} \rceil - g(S),$$

we have

$$\underline{\mathrm{SHI}}(-Y(K), -\Gamma_n, S^\tau, i) \cong \underline{\mathrm{SHI}}(-Y(1), -\delta).$$

Remark 4.40. A more direct explanation of Lemma 4.39 is that, apart from the top $2g(S)$ and the bottom $2g(S)$ gradings, the vector spaces in all gradings are isomorphic to $\underline{\mathrm{SHI}}(-Y(1), -\delta)$.

Next, suppose we perform a $(-n)$-surgery. We can take

$$\hat{\lambda} = (0,-1) = -\mu \text{ and } \hat{\mu} = (-1,n) = \lambda - n\mu.$$

Then we compute

$$\widehat{\Gamma}_\mu(-n) = \Gamma_n,\ \widehat{\Gamma}_0(-n) = \Gamma_\mu,\ \widehat{\Gamma}_1(-n) = \gamma_{(-1,n-1)} = \Gamma_{n-1},\ \widehat{\Gamma}_2(-n) = \gamma_{(-2,2n-1)}, \tag{4.11}$$

and also

$$\widehat{\Gamma}_{-1}(-n) = \gamma_{(-1,n+1)} = \Gamma_{n+1},\ \widehat{\Gamma}_{-2}(-n) = \gamma_{(-2,2n+1)}. \tag{4.12}$$

Observe that

$$\widehat{\Gamma}_{-2}(-n) = \widehat{\Gamma}_2(-n-1). \tag{4.13}$$

The following is the first part of 1.15.

Proposition 4.41. *Suppose integer $n \geqslant 2g(S) + 1$, then*

$$\mathcal{I}_+(-\widehat{Y}_{-n}, \widehat{K}, i) \cong \underline{\mathrm{SHI}}(-Y(1), -\delta)$$

for any integer $i \in [0, n - 2g(S) - 1]$ and $i = n - 1$.

Proof. When $0 \leqslant i \leqslant n - 1 - 2g(S)$, we know from equality (4.11), Definition 4.21, Lemmas 4.16, and 4.39 that

$$\begin{aligned}
\mathcal{I}_+(-\widehat{Y}_{-n}, \widehat{K}, i) &= \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_2(-n), S^\tau, i^{2n-1}_{\max} - 2g(S) - i) \\
&\cong \underline{\mathrm{SHI}}(-Y(K), -\widehat{\Gamma}_\mu(-n), S^\tau, i^{n}_{\max} - 2g(S) - i) \\
&= \underline{\mathrm{SHI}}(-Y(K), -\Gamma_n, S^\tau, i^{n}_{\max} - 2g(S) - i) \\
&\cong \underline{\mathrm{SHI}}(-Y(1), -\delta).
\end{aligned}$$

For $i = n-1$, we know similarly that

$$\begin{aligned}\mathcal{I}_+(-\hat{Y}_{-n},\hat{K},i) &= \underline{\mathrm{SHI}}(-Y(K),-\hat{\Gamma}_2(-n),S^\tau,i_{\max}^{2n-1}-2g(S)-n+1)\\ &\cong \underline{\mathrm{SHI}}(-Y(K),-\hat{\Gamma}_\mu(-n),S^\tau,i_{\max}^{n}-2g(S))\\ &= \underline{\mathrm{SHI}}(-Y(K),-\Gamma_n,S^\tau,i_{\max}^{n}-2g(S))\\ &\cong \underline{\mathrm{SHI}}(-Y(1),-\delta).\end{aligned}$$ □

The following is the second part of Proposition 1.15.

Proposition 4.42. *Suppose integer $n \geqslant 2g(S)+1$, then for any integer $i \in [0, n-1]$, we have*

$$\mathcal{I}_+(-\hat{Y}_{-n-1},\hat{K},i+1) \cong \mathcal{I}_+(-\hat{Y}_{-n},\hat{K},i).$$

Proof. For integer $i \in [0, n-1-2g(S)]$ and $i = n-1$, we know from Proposition 4.41 that

$$\mathcal{I}_+(-\hat{Y}_{-n-1},\hat{K},i+1) \cong \underline{\mathrm{SHI}}(-Y(1),-\delta) \cong \mathcal{I}_+(-\hat{Y}_{-n},\hat{K},i)$$

For the rest $(2g(S)-1)$ gradings, we use the following commutative diagram.

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K),-\Gamma_{n-1}) \arrow[r, "\psi^{n-1}_{+,n}"] & \underline{\mathrm{SHI}}(-Y(K),-\Gamma_n) \\ \underline{\mathrm{SHI}}(-Y(K),-\Gamma_n) \arrow[u, "\psi^{\mu}_{-,1}(-n)"] \arrow[r, "\psi^{n}_{-,n+1}"] & \underline{\mathrm{SHI}}(-Y(K),-\Gamma_{n+1}) \arrow[u, "\psi^{\mu}_{-,1}(-n-1)"'] \end{tikzcd} \tag{4.14}$$

This is directly from facts (4.11) and (4.12), and Corollary 4.36 by taking $i = 1$ and $j = -1$. Note that

$$\psi^{n-1}_{+,n} = \psi^{1}_{-,\mu}(-n), \psi^{n}_{-,n+1} = \psi^{\mu}_{-,-1}(-n), \text{ and } \psi^{\mu}_{-,1}(-n-1) = \psi^{-1}_{-,\mu}(-n).$$

From Proposition 4.14 and Definition 4.21, we obtain a graded commutative diagram

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K),-\Gamma_{n-1},S^\tau,i^{n-1}_{\min}+n-2-i) \arrow[r, "\psi^{n-1}_{+,n}"] & \underline{\mathrm{SHI}}(-Y(K),-\Gamma_n,S^\tau,i^{n}_{\min}+n-2-i) \\ \underline{\mathrm{SHI}}(-Y(K),-\Gamma_n,S^\tau,i^{n}_{\max}+n-2g-1-i) \arrow[u, "\psi^{\mu}_{-,1}(-n)"] \arrow[r, "\psi^{n}_{-,n+1}"] & \underline{\mathrm{SHI}}(-Y(K),-\Gamma_{n+1},S^\tau,i^{n+1}_{\max}+n-2g-1-i) \arrow[u, "\psi^{\mu}_{-,1}(-n-1)"'] \end{tikzcd}$$

For any fixed integer $i \in [n-2g(S), n-1]$, we have a graded exact triangle

$$\begin{tikzcd} \underline{\mathrm{SHI}}(-Y(K),-\Gamma_{n-1},S^\tau,i^{n-1}_{\min}+n-2-i) \arrow[rr] & & \mathcal{I}_+(-\hat{Y}_{-n},\hat{K},i) \arrow[dll] \\ \underline{\mathrm{SHI}}(-Y(K),-\Gamma_n,S^\tau,i^{n}_{\max}+n-2g-1-i) \arrow[u, "\psi^{\mu}_{-,1}(-n)"] & & \end{tikzcd}$$

Hence we know that

$$\dim_{\mathbb{C}} \mathcal{I}_+(-\hat{Y}_{-n}, \hat{K}, i) = \dim_{\mathbb{C}} \ker(\psi^{\mu}_{-,1}(-n)) + \dim_{\mathbb{C}} \operatorname{coker}(\psi^{\mu}_{-,1}(-n)).$$

Similarly, we know that

$$\dim_{\mathbb{C}} \mathcal{I}_+(-\hat{Y}_{-n-1}, \hat{K}, i+1) = \dim_{\mathbb{C}} \ker(\psi^{\mu}_{-,1}(-n-1)) + \dim_{\mathbb{C}} \operatorname{coker}(\psi^{\mu}_{-,1}(-n-1))$$

Since the maps $\psi^{n-1}_{+,n}$ and $\psi^{n}_{-,n+1}$ are both isomorphisms in the corresponding gradings by Lemma 4.16, we conclude that

$$\dim_{\mathbb{C}} \mathcal{I}_+(-\hat{Y}_{-n-1}, \hat{K}, i+1) = \dim_{\mathbb{C}} \mathcal{I}_+(-\hat{Y}_{-n}, \hat{K}, i).$$

By five lemma, there is indeed an isomorphism between $\mathcal{I}_+(-\hat{Y}_{-n-1}, \hat{K}, i+1)$ and $\mathcal{I}_+(-\hat{Y}_{-n}, \hat{K}, i)$. ☐

5 | SOME REMARKS AND FUTURE DIRECTIONS

In this section, we state some remarks and further directions.

First, the condition in Theorem 1.12 that the boundary of the Seifert surface S of $\hat{K}$ is connected can be removed by modifying the hypothesis and the statement as follows.

(1) Suppose the order of $[\hat{K}] \in H_1(\hat{Y})$ is a, that is, a is the minimal positive integer so that $a[K] = 0$. Suppose the number of the boundary components of S is b, and λ is a simple closed curve on $\partial \hat{Y}(\hat{K})$ so that $\partial S = b\lambda$.

(2) Choose another simple closed curve μ on $\partial \hat{Y}(\hat{K})$ so that $\mu \cdot \lambda = -1$. Suppose the meridian of $\hat{K}$ has homology class $(q\mu + p\lambda)$. It can be shown that q is independent of the choice of μ. Indeed, we know that $q = a/b$.

(3) Let definitions of i^y_{max} and i^y_{min} in Definition 4.13 be replaced by

$$i^y_{max} = \lceil \frac{1}{2}(yb - \chi(S)) \rceil \text{ and } i^y_{min} = \lceil -\frac{1}{2}(yb - \chi(S)) \rceil.$$

(4) All proofs of Theorem 1.12 and Proposition 1.16 apply without essential change and we will obtain a decomposition associated to $\hat{K}$

$$I^{\sharp}(\hat{Y}) \cong \bigoplus_{i=0}^{a-1} I^{\sharp}(\hat{Y}, i).$$

Note that in the original statement of Theorem 1.12, the integer q is indeed the order of $[\hat{K}]$.

Second, as mentioned in Remark 1.14, it is not known if the decomposition of $I^{\sharp}(\hat{Y})$ is independent of the choice of $\hat{K}$. Explicitly, we have the following conjecture.

Conjecture 5.1. *Suppose $(\hat{Y}, \hat{K})$ satisfies the hypothesis of Theorem 1.12. Suppose further that another knot $\hat{K}' \subset \hat{Y}$ satisfies the similar conditions to those of $\hat{K}$, and*

$$[\hat{K}] = [\hat{K}'] \in H_1(\hat{Y}).$$

Then there exists a grading preserving isomorphism

$$\mathcal{I}_+(\hat{Y}, \hat{K}) \cong \mathcal{I}_+(\hat{Y}, \hat{K}')$$

up to a $\mathbb{Z}_q$ *grading shift, where* $\mathcal{I}_+$ *is defined as in Definition 4.21.*

Last, though we have had a decomposition of $I^\sharp(-\hat{Y})$, we do not know if it works well with the cobordism maps. For example, let $K \subset S^3$ be a knot and let $S^3_{-n}(K)$ be obtained from S^3 by a $(-n)$-surgery along K. Then there is a natural cobordism W from S^3 to $S^3_{-n}(K)$, which induces a cobordism map

$$I^\sharp(W_{-n}) : I^\sharp(-S^3_{-n}(K)) \to I^\sharp(-S^3).$$

Baldwin and Sivek [6, Section 7] proved that the cobordism map decomposes in basic classes:

$$I^\sharp(W_{-n}) = I^\sharp(W_{-n}, t_i),$$

where $t_i : H_2(W_{-n}) \to \mathbb{Z}$ maps $[S']$ to $(2i)$, where $[S']$ is the homology class of the surface obtained by capping off of the Seifert surface of K. We have the following conjecture, basically saying that the decomposition of the cobordism map $I^\sharp(W_{-n})$ is compatible with the decomposition of $I^\sharp(-S^3_{-n}(K))$.

Conjecture 5.2. *There exists an integer N so that for any integer $i \in [0, n-1]$, under the identifications $I^\sharp(-S^3_{-n}(K)) \cong \mathcal{I}_+(-S^3_{-n}(K), \hat{K})$ and $I^\sharp(-S^3) \cong \mathcal{I}_+(-S^3_{1/0}(K), K)$, we have*

$$I^\sharp(W_{-n}, t_i) = I^\sharp(W_{-n})|_{\mathcal{I}_+(-S^3_{-n}(K), \hat{K}, N-i)}$$

and this cobordism map can be recovered by bypass maps.

ACKNOWLEDGEMENTS

The first author is partially supported by his Ph.D. advisor Tomasz Mrowka's NSF Grant 1808794. The second author would like to thank his supervisor Jacob Rasmussen for his patient guidance and helpful comments and thank his parents for support and constant encouragement. The authors would also like to thank Ciprian Manolescu for helpful comments on the draft of the paper and thank John A. Baldwin, Qilong Guo, Joshua Wang, Yi Xie, and Ruide Fu for helpful conversations. The second author is also grateful to Yi Liu for inviting him to BICMR, Peking University, and Zekun Chen, Honghuai Fang, and Longke Tang for the company when he was writing this paper.

JOURNAL INFORMATION

The *Journal of Topology* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES

1. A. Alfieri, J. A. Baldwin, I. Dai, and S. Sivek, *Instanton Floer homology of almost-rational plumbings*, arXiv:2010.03800, v1, 2020.
2. J. A. Baldwin, Z. Li, and F. Ye, *Sutured instanton homology and Heegaard diagrams*, arXiv:2011.09424, v1, 2020.
3. J. A. Baldwin and S. Sivek, *Naturality in sutured monopole and instanton homology*, J. Differ. Geom. **100** (2015), no. 3, 395–480.
4. J. A. Baldwin and S. Sivek, *Instanton Floer homology and contact structures*, Selecta Math. (N.S.) **22** (2016), no. 2, 939–978.
5. J. A. Baldwin and S. Sivek, *Khovanov homology detects the trefoils*, Duke Math. J. **174** (2022), no. 4, 885–956.
6. J. A. Baldwin and S. Sivek, *Instanton and L-space surgeries*, arXiv:1910.13374, v1, 2019.
7. J. A. Baldwin and S. Sivek, *Framed instanton homology and concordance*, J. Topol. **14** (2021), no. 4, 1113–1175.
8. J. Bloom, *A link surgery spectral sequence in monopole Floer homology*, Adv. Math. **226** (2009), no. 4, 3216–3281.
9. V. Colin, P. Ghiggini, and K. Honda, *The equivalence of Heegaard Floer homology and embedded contact homology III: from hat to plus*, arXiv:1208.1526, v2, 2017.
10. V. Colin, P. Ghiggini, K. Honda, and M. Hutchings, *Sutures and contact homology I*, Geom. Topol. **15** (2010), no. 3, 1749–1842.
11. A. Daemi and C. Scaduto, *Equivariant aspects of singular instanton Floer homology*, arXiv:1912.08982, v2, 2020.
12. A. Floer, *An instanton-invariant for* 3*-manifolds*, Comm. Math. Phys. **118** (1988), no. 2, 215–240.
13. A. Floer, *Instanton homology, surgery, and knots*, Geometry of Low-Dimensional Manifolds, 1 (Durham, 1989), London Math. Soc. Lecture Note Ser., vol. 150, Cambridge Univ. Press, Cambridge, 1990, pp. 97–114.
14. D. Gabai, *Foliations and the topology of* 3*-manifolds*, J. Differ. Geom. **18** (1983), no. 3, 445–503.
15. S. Ghosh and Z. Li, *Decomposing sutured monopole and instanton Floer homologies*, arXiv:1910.10842, v2, 2019.
16. S. Ghosh, Z. Li, and C.-M. M. Wong, *Tau invariants in monopole and instanton theories*, arXiv:1910.01758, v3, 2019.
17. H. Goda, H. Matsuda, and T. Morifuji, *Knot Floer homology of (1,1)-knots*, Geom. Dedicata **112** (2005), no. 1, 197–214.
18. J. E. Greene and A. S. Levine, *Strong Heegaard diagrams and strong L–spaces*, Algebr. Geom. Topol. **16** (2016), no. 6, 3167–3208.
19. J. Hanselman, J. Rasmussen, and L. Watson, *Bordered Floer homology for manifolds with torus boundary via immersed curves*, arXiv:1604.03466, v2, 2017.
20. J. Hanselman, J. Rasmussen, and L. Watson, *Heegaard Floer homology for manifolds with torus boundary: properties and examples*, arXiv:1810.10355, v1, 2018.
21. M. Hedden, C. Herald, and P. Kirk, *The pillowcase and perturbations of traceless representations of knot groups*, Geom. Topol. **18** (2014), no. 1, 211–287.
22. K. Honda, *On the classification of tight contact structures I*, Geom. Topol. **4** (2000), 309–368.
23. K. Honda, *Gluing tight contact structures*, Duke Math. J. **115** (2002), no. 3, 435–478.
24. K. Honda, W. H. Kazez, and G. Matić, *Contact structures, sutured Floer homology and TQFT*, arXiv:0807.2431, v1, 2008.
25. M. Hutchings, *Embedded contact homology and its applications*, Proceedings of the International Congress of Mathematicians 2010, ICM 2010, 1022–1041, 2010.
26. A. Juhász, *Holomorphic discs and sutured manifolds*, Algebr. Geom. Topol. **6** (2006), 1429–1457.
27. A. Juhász, *Floer homology and surface decompositions*, Geom. Topol. **12** (2008), no. 1, 299–350.
28. A. Juhász, D. P. Thurston, and I. Zemke, *Naturality and mapping class groups in Heegaard Floer homology*, Mem. Amer. Math. Soc. **273** (2021), no. 1338, v+174 pp.
29. A. Juhász and I. Zemke, *Contact handles, duality, and sutured Floer homology*, Geom. Topol. **24** (2020), no. 1, 179–307.
30. N. Kavi, *Cutting and gluing surfaces*, arXiv:1910.11954, v1, 2019.
31. P. B. Kronheimer and T. S. Mrowka, *Witten's conjecture and property P*, Geom. Topol. **8** (2004), 295–310.
32. P. B. Kronheimer and T. S. Mrowka, *Monopoles and three-manifolds*, New Mathematical Monographs, vol. 10, Cambridge Univ. Press, Cambridge, 2007.

33. P. B. Kronheimer and T. S. Mrowka, *Instanton Floer homology and the Alexander polynomial*, Algebr. Geom. Topol. **10** (2010), no. 3, 1715–1738.
34. P. B. Kronheimer and T. S. Mrowka, *Knots, sutures, and excision*, J. Differ. Geom. **84** (2010), no. 2, 301–364.
35. P. B. Kronheimer and T. S. Mrowka, *Khovanov homology is an unknot-detector*, Publ. Math. Inst. Hautes Études Sci. **113** (2011), 97–208.
36. P. B. Kronheimer and T. S. Mrowka, *Knot homology groups from instantons*, J. Topol. **4** (2011), no. 4, 835–918.
37. P. B. Kronheimer and T. S. Mrowka, *Filtrations on instanton homology*, Quantum Topol. **5** (2014), no. 1, 61–97.
38. C. Kutluhan, Y.-J. Lee, and C. H. Taubes, *HF=HM I: Heegaard Floer homology and Seiberg-Witten Floer homology*, Geom. Topol. **24** (2020), no. 6, 2829–2854.
39. C. Kutluhan, S. Sivek, and C. H. Taubes, *Sutured ECH is a natural invariant*, Mem. Amer. Math. Soc. **275** (2022), no. 1350, v+136 pp.
40. Z. Li, *Contact structures, excisions, and sutured monopole Floer homology*, Algebr. Geom. Topol. **20** (2020), no. 5, 2553–2588.
41. Z. Li, *Gluing maps and cobordism maps for sutured monopole Floer homology*, arXiv:1810.13071, v3, 2018.
42. Z. Li, *Knot Floer homologies in monopole and instanton theory via sutures*, arXiv:1901.06679, v6, 2019.
43. Z. Li and F. Ye, *SU(2) representations and a large surgery formula*, arXiv:2107.11005, v1, 2021.
44. Z. Li and F. Ye, *An enhanced Euler characteristic of sutured instanton homology*, arXiv:2107.10490, v1, 2021.
45. Z. Li and F. Ye, *Instanton Floer homology, sutures, and Euler characteristics*, arXiv:2011.09424, v3, 2021.
46. T. Lidman, J. Pinzón-Caicedo, and C. Scaduto, *Framed instanton homology of surgeries on L-space knots*, arXiv:2003.03329, v1, 2020.
47. Y. Lim, *Instanton homology and the Alexander polynomial*, Proc. Amer. Math. Soc. **138** (2010), no. 10, 3759–3768.
48. A. Lobb and R. Zentner, *On spectral sequences from Khovanov homology*, Algebr. Geom. Topol. **20** (2020), no. 2, 531–564.
49. L. Moser, *Elementary surgery along a torus knot*, Pacific J. Math. **38** (1971), 737–745.
50. Y. Ni, *Knot Floer homology detects fibred knots*, Invent. Math. **170** (2007), no. 3, 577–608.
51. B. Ozbagci, *Contact handle decompositions*, Topol. Appl. **158** (2011), no. 5, 718–727.
52. P. S. Ozsváth and Z. Szabó, *Holomorphic disks and knot invariants*, Adv. Math. **186** (2004), no. 1, 58–116.
53. P. S. Ozsváth and Z. Szabó, *Holomorphic disks and topological invariants for closed three-manifolds*, Ann. of Math. (2) **159** (2004), no. 3, 1027–1158.
54. P. S. Ozsváth and Z. Szabó, *On knot Floer homology and lens space surgeries*, Topology **44** (2005), 1281–1300.
55. J. Rasmussen, *Floer homology and knot complements*, arXiv:math/0306378, v1, 2003.
56. J. Rasmussen, *Knot polynomials and knot homologies*, Geometry and topology of manifolds, Fields Inst. Commun., vol. 47, Amer. Math. Soc., Providence, RI, 2005, pp. 261–280.
57. J. Rasmussen, *Lens space surgeries and L-space homology spheres*, arXiv:0710.2531, v1, 2007.
58. J. Rasmussen and S. D. Rasmussen, *Floer simple manifolds and L-space intervals*, Adv. Math. **322** (2017), 738–805.
59. C. Scaduto, *Instantons and odd Khovanov homology*, J. Topol. **8** (2015), no. 3, 744–810.
60. C. Scaduto and M. Stoffregen, *Two-fold quasi-alternating links, Khovanov homology and instanton homology*, Quantum Topol. **9** (2018), no. 1, 167–205.
61. C. H. Taubes, *Embedded contact homology and Seiberg-Witten Floer cohomology I*, Geom. Topol. **14** (2010), no. 5, 2497–2581.
62. J. Wang, *The cosmetic crossing conjecture for split links*, arXiv:2006.01070, v1, 2020.
63. Y. Xie and B. Zhang, *Instanton Floer homology for sutured manifolds with tangles*, arXiv:1907.00547, v2, 2019.
64. F. Ye, *Constrained knots in lens spaces*, arXiv:2007.04237, v3, 2021.