

A MIXED INVARIANT OF NONORIENTABLE SURFACES IN EQUIVARIANT KHOVANOV HOMOLOGY

ROBERT LIPSHITZ AND SUCHARIT SARKAR

ABSTRACT. We construct a mixed invariant of nonorientable surfaces from the Lee and Bar-Natan deformations of Khovanov homology and use it to distinguish pairs of surfaces bounded by the same knot, including some exotic examples.

CONTENTS

| | |
|---|----|
| 1. Introduction | 1 |
| Acknowledgments | 4 |
| 2. Background on the Lee and Bar-Natan deformations | 4 |
| 3. Behavior under (possibly nonorientable) cobordisms | 7 |
| 4. Admissible cuts | 15 |
| 5. The mixed invariant | 22 |
| 6. Properties | 24 |
| 6.1. First observations | 24 |
| 6.2. Stabilizations | 28 |
| 6.3. Closed surfaces | 33 |
| 7. Computations, applications, and questions | 35 |
| 7.1. A first direct computation | 35 |
| 7.2. A more interesting example | 36 |
| 7.3. An exotic pair of surfaces | 38 |
| 7.4. Some questions | 40 |
| References | 41 |

1. INTRODUCTION

Ozsváth-Szabó's Heegaard Floer homology associates $\mathbb{Z}[U]$ -modules $HF^-(Y, \mathfrak{s})$ and $HF^+(Y, \mathfrak{s})$ to a closed, connected, oriented 3-manifold Y and Spin^C -structure \mathfrak{s} , and $\mathbb{Z}[U]$ -module homomorphisms $F^\pm(W, \mathfrak{t}): HF^\pm(Y_1, \mathfrak{s}_1) \rightarrow HF^\pm(Y_2, \mathfrak{s}_2)$ to a Spin^C -cobordism $(W, \mathfrak{t}): (Y_1, \mathfrak{s}_1) \rightarrow (Y_2, \mathfrak{s}_2)$ [OSz04, OSz06]. For a closed 4-manifold W with $b_2^+ > 0$, viewed as a cobordism from S^3 to itself by deleting two balls, the maps $HF^\pm(W, \mathfrak{t})$ vanish. Using the proof of vanishing, Ozsváth-Szabó define a Heegaard Floer-theoretic analogue of the Seiberg-Witten invariant, the *Heegaard Floer*

Date: May 5, 2022.

RL was supported by NSF Grant DMS-1810893.

SS was supported by NSF Grant DMS-1905717.

mixed invariant, which is a map $HF^-(Y_1, \mathfrak{s}_1) \rightarrow HF^+(Y_2, \mathfrak{s}_2)$ associated to a Spin^C -cobordism with b_2^+ sufficiently large.

The goal of this paper is to give an analogous construction in Khovanov homology, for smoothly embedded surfaces in $[0, 1] \times S^3$ with crosscap number at least 3.

Following Finashin-Kreck-Viro [FKV87], call a pair of smoothly embedded surfaces $F, F' \subset [0, 1] \times S^3$ *exotic* if there is a self-homeomorphism of $[0, 1] \times S^3$ which is the identity on $\{0, 1\} \times S^3$ and takes F to F' , but there is no such self-diffeomorphism of $[0, 1] \times S^3$. (See also Lemma 4.7 and Section 7.3.) At the time of writing, we believe no pairs of exotic closed, orientable surfaces in $[0, 1] \times S^3$ are known. There are, however, exotic nonorientable surfaces in $[0, 1] \times S^3$, as first shown by Finashin-Kreck-Viro using results from Donaldson theory [FKV88]. Recently, examples of exotic orientable cobordisms in $[0, 1] \times S^3$ have also appeared, in work of Juhász-Miller-Zemke [JMZ21], Hayden [Hay], and Hayden-Kjuchukova-Krishna-Miller-Powell-Sunukjian [HKK⁺], which use Heegaard Floer homology to distinguish them, and Hayden-Sundberg [HS], which uses Khovanov homology.

Many question about exotic pairs of surfaces remain open. For example, Baykur-Sunukjian introduced stabilization operations for surfaces, and showed that all known examples of exotic pairs of closed surfaces become diffeomorphic after a single stabilization [BS16]; it is not known if this holds in general. Building on these ideas, one can study the *total stabilization distance* between two surfaces F, F' , the minimum number of stabilizations or destabilizations needed to turn one into the other [Miy86, MP19], or the *max stabilization distance*, the minimum over sequences $F = F_0, F_1, \dots, F_n = F'$, where F_i and F_{i+1} are related by a stabilization, destabilization, or taking the connected sum with a knotted 2-sphere, of $\max\{|g(F_1) - g(F)|, \dots, |g(F_n) - g(F)|\}$ (where g denotes the genus) [JZb]. (See also [Mel77, p. 6].) Another notion is the *generalized total stabilization distance*, which is defined the same way as the total stabilization distance except that if two surfaces differ by taking the connected sum with a 2-sphere then they are declared to be at distance 0, so the generalized total stabilization distance, like the max stabilization distance, focuses on global, rather than local, knotting [MP19]. By using the Alexander module, Miyazaki shows there are pairs of embedded spheres in S^4 with arbitrarily high total stabilization distance [Miy86], and Miller-Powell show there are pairs of embedded disks with arbitrarily high generalized total stabilization distance [MP19]. Juhász-Zemke use Heegaard Floer homology to give pairs of disks with max stabilization distance at least 3 [JZb]. The analogous questions for exotic pairs are open.

A key strategy for distinguishing knotted closed surfaces in S^4 has been to apply gauge theory, like the Heegaard Floer mixed invariant, to their branched double covers. Ozsváth-Szabó showed that the Heegaard Floer homology of the branched double cover of a knot K is closely related to the Khovanov homology of K [OSz05]. So, it seems natural to look for an analogue of the Heegaard Floer mixed invariant in Khovanov homology

In this paper, we give one such analogue. Khovanov homology admits a family of deformations [Kho06b]; we will focus on two particular ones, the Lee deformation [Lee05] and the Bar-Natan deformation [Bar05]. Rasmussen showed that the map of Lee homologies associated to a nonorientable cobordism vanishes [Ras10]. Viewing these deformations as modules over polynomial algebras, analogous to the Heegaard Floer invariant HF^- , Rasmussen's result says that the map on the analogue of HF^∞ associated to a nonorientable cobordism vanishes. Using this, and a notion of admissible cuts analogous to Heegaard Floer theory, we formulate a Khovanov mixed invariant

Φ_F of a surface F with crosscap number ≥ 3 in the Lee and Bar-Natan deformations of Khovanov homology. Note that F having crosscap number ≥ 3 has no implication for b_2^+ , so the Khovanov mixed invariant is defined for some surfaces F where the Heegaard Floer mixed invariant of $\Sigma(F)$ is not.

Verifying that the mixed invariant is well-defined (up to sign) has two steps: observing that the map on (deformed) Khovanov homology associated to a nonorientable cobordism is well-defined (up to sign), and verifying independence of the choice of admissible cut. The proof of the first statement is a straightforward extension of the literature [Jac04, Kho06a, Bar05, MWW]; see Section 3. (Unlike the orientable case, the sign ambiguity here is essential; see Remark 3.8.) To prove independence of the admissible cut, we use arguments involving the one-sided curve complex of a nonorientable surface; see Section 4.

It turns out that, unlike the Heegaard Floer mixed invariant, this Khovanov mixed invariant does not distinguish closed, connected surfaces (Section 6.3), though the proof of this fact is somewhat intricate. (We do not know if the mixed invariant distinguishes some closed, disconnected surfaces, and in particular have not generalized Gujral-Levine's results [GL] to this setting.) On the other hand, both the mixed invariant and the map on Khovanov homology associated to a nonorientable cobordism do distinguish pairs of nonorientable surfaces with common boundary. Indeed, this was essentially already shown in computations of Sundberg-Swann [SS]: combined with the functoriality result mentioned above, their computations show the following.

Theorem 1.1. *There is a pair of connected surfaces F, F' with boundary on $3_1 \# m(3_1)$ with crosscap number 3 and normal Euler number -6 which are not isotopic, and do not become isotopic after taking the connected sum with any knotted 2-sphere. Further, F is not obtained from a connected surface F'' by attaching a 1-handle, or by taking the connected sum with a standard \mathbb{RP}^2 or $\overline{\mathbb{RP}}^2$.*

Theorem 1.1 is proved in Section 7.2. The second half of the theorem uses the behavior of Φ_F under various local modifications to the surface, which are summarized in Theorem 6.18.

Hayden-Sundberg's examples of exotic pairs of slice disks distinguished by Khovanov homology can be enhanced to give exotic pairs of nonorientable surfaces distinguished by Khovanov homology. In particular, we have:

Theorem 1.2. *There is an exotic pair of surfaces with boundary 12_{309}^n , crosscap number 3, and normal Euler number -6 .*

Theorem 1.2 is proved in Section 7.3. As far as we know, this is the first gauge theory-free proof that there are pairs of exotic nonorientable surfaces.

This paper is organized as follows. We review the Lee and Bar-Natan deformations of Khovanov homology in Section 2, in an algebraic framework parallel to Heegaard Floer homology. Section 3 shows that these deformed Khovanov complexes are functorial with respect to nonorientable cobordisms. For convenience later, we also allow our cobordisms to be decorated with stars (following the notation of [KR22]). Section 4 formulates the notion of admissible cuts, and shows that for surfaces with crosscap number ≥ 3 all admissible cuts are equivalent in a suitable sense. Section 5 defines the mixed invariant and proves it is well-defined. Section 6 gives basic properties of the maps associated to nonorientable cobordisms and the mixed invariant, and Section 7 gives some

computations and applications of these invariants, including proving Theorems 1.1 and 1.2, and concludes with some questions.

Acknowledgments. We thank Ian Zemke for helpful comments on the first draft of this paper, and the referee for further comments and corrections.

2. BACKGROUND ON THE LEE AND BAR-NATAN DEFORMATIONS

Khovanov homology has two well-known deformations, the Lee deformation [Lee05] and the Bar-Natan deformation [Bar05]. The two theories are similar, although they have some essential differences as well. Most of the constructions and results of this paper work for either of the two theories, so we will use the same notations for both the theories. When the two theories diverge, we will explicitly mention that in the text.

Fix a commutative ring R with unit. All chain complexes and modules will be defined over R , though we will often suppress R from the notation. If we are using the Lee theory, we assume 2 is a unit in R .

Fix an oriented link diagram L with N crossings, N_+ of which are positive and N_- of which are negative. Consider the Kauffman cube of resolutions of L . A *Khovanov generator* y is a choice of vertex v and a labeling $y(Z) \in \{1, X\}$ of each circle Z in the v -resolution. Denoting the homological, quantum bigrading by $(\text{gr}_h, \text{gr}_q)$, a Khovanov generator y lying over a vertex $v \in \{0, 1\}^N$ has bigrading

$$\begin{aligned} \text{gr}_h(y) &= -N_- + |v| \\ \text{gr}_q(y) &= N_+ - 2N_- + |v| + \#\{Z \mid y(Z) = 1\} - \#\{Z \mid y(Z) = X\}. \end{aligned}$$

The deformed Khovanov complex $\mathcal{C}^-(L)$ is freely generated by these generators over a polynomial algebra over R , and is obtained by feeding the cube of resolutions into a Frobenius algebra over that polynomial algebra.

- (1) In the Lee theory, the polynomial algebra is $R[T]$ with T in bigrading $(0, -4)$, and the Frobenius algebra is $R[T, X]/(X^2 = T)$, with co-multiplication given by

$$\Delta(1) = 1 \otimes X + X \otimes 1 \qquad \Delta(X) = X \otimes X + T1 \otimes 1$$

and counit $\epsilon: R[T, X]/(X^2 = T) \rightarrow R[T]$ given by $\epsilon(1) = 0$, $\epsilon(X) = 1$.

- (2) In the Bar-Natan theory, the polynomial algebra is $R[H]$ with H in bigrading $(0, -2)$, and the Frobenius algebra is $R[H, X]/(X^2 = XH)$, with co-multiplication given by

$$\Delta(1) = 1 \otimes X + X \otimes 1 - H1 \otimes 1 \qquad \Delta(X) = X \otimes X$$

and counit $\epsilon: R[H, X]/(X^2 = XH) \rightarrow R[H]$ given by $\epsilon(1) = 0$, $\epsilon(X) = 1$.

(In Khovanov's paper [Kho06b], these are denoted \mathcal{F}_7 and \mathcal{F}_3 , respectively.) In either theory, the differential increases the bigrading by $(1, 0)$. To keep the notation the same, let $R[U]$ denote the polynomial algebra for either theory. That is, when discussing the Lee theory we take $U = T$ (in bigrading $(0, -4)$), and when discussing the Bar-Natan theory we take $U = H$ (in bigrading $(0, -2)$). In either case, the original non-deformed Khovanov complex is obtained by setting $U = 0$; in analogy

with Heegaard Floer homology, we will denote the non-deformed complex $\widehat{\mathcal{C}}(L) = \mathcal{C}^-(L)/\{U = 0\}$. The homology $\widehat{\mathcal{H}}(L)$ of $\widehat{\mathcal{C}}(L)$ is ordinary Khovanov homology, often denoted $Kh(L)$.

Continuing the analogy with Heegaard Floer homology, let

$$\begin{aligned}\mathcal{C}^\infty(L) &= U^{-1}\mathcal{C}^-(L) \\ \mathcal{C}^+(L) &= \mathcal{C}^\infty(L)/\mathcal{C}^-(L),\end{aligned}$$

where the notation U^{-1} denotes localization or, equivalently, tensoring over $\mathbb{R}[U]$ with $\mathbb{R}[U^{-1}, U]$. (These conventions are not exactly parallel to Heegaard Floer homology [OSz04, Section 4.1].) Let $\mathcal{H}^-(L)$, $\mathcal{H}^\infty(L)$, and $\mathcal{H}^+(L)$ be the homologies of $\mathcal{C}^-(L)$, $\mathcal{C}^\infty(L)$, and $\mathcal{C}^+(L)$. See Figure 7.3 for an example of $\mathcal{C}^\infty(L)$ and its subcomplex $\mathcal{C}^-(L)$ and quotient complex $\mathcal{C}^+(L)$ (up to quasi-isomorphism), and a comparison with the usual formulation of the Lee deformation.

There are short exact sequences

$$\begin{aligned}(2.1) \quad & 0 \rightarrow \mathcal{C}^-(L) \xrightarrow{\iota} \mathcal{C}^\infty(L) \xrightarrow{\pi} \mathcal{C}^+(L) \rightarrow 0 \\ & 0 \rightarrow \mathcal{C}^-(L) \xrightarrow{U} \mathcal{C}^-(L) \xrightarrow{\pi} \widehat{\mathcal{C}}(L) \rightarrow 0 \\ & 0 \rightarrow \widehat{\mathcal{C}}(L) \xrightarrow{\iota} \mathcal{C}^+(L) \xrightarrow{U} \mathcal{C}^+(L) \rightarrow 0.\end{aligned}$$

and corresponding long exact sequences

$$\begin{aligned}(2.2) \quad & \cdots \rightarrow \mathcal{H}^-(L) \xrightarrow{\iota_*} \mathcal{H}^\infty(L) \xrightarrow{\pi_*} \mathcal{H}^+(L) \xrightarrow{\partial} \mathcal{H}^-(L) \rightarrow \cdots, \\ & \cdots \rightarrow \mathcal{H}^-(L) \xrightarrow{U} \mathcal{H}^-(L) \xrightarrow{\pi_*} \widehat{\mathcal{H}}(L) \xrightarrow{\partial} \mathcal{H}^-(L) \rightarrow \cdots, \\ & \cdots \rightarrow \widehat{\mathcal{H}}(L) \xrightarrow{\iota_*} \mathcal{H}^+(L) \xrightarrow{U} \mathcal{H}^+(L) \xrightarrow{\partial} \widehat{\mathcal{H}}(L) \rightarrow \cdots.\end{aligned}$$

The homomorphisms U decrease the bigrading by $(0, 2)$ for the Bar-Natan deformation and $(0, 4)$ for the Lee deformation, the homomorphism $\widehat{\mathcal{H}}(L) \rightarrow \mathcal{H}^+(L)$ increases bigrading by $(0, 2)$ for the Bar-Natan deformation and $(0, 4)$ for the Lee deformation, the connecting homomorphisms $\mathcal{H}^+(L) \rightarrow \mathcal{H}^-(L)$ and $\mathcal{H}^+(L) \rightarrow \widehat{\mathcal{H}}(L)$ increase the bigrading by $(1, 0)$, the connecting homomorphism $\widehat{\mathcal{H}}(L) \rightarrow \mathcal{H}^-(L)$ increases bigrading by $(1, 2)$ for the Bar-Natan deformation and $(1, 4)$ for the Lee deformation, and the other maps preserve the bigrading.

Commutativity of the diagrams

$$\begin{array}{ccc} 0 \rightarrow \mathcal{C}^-(L) \xrightarrow{\iota} \mathcal{C}^\infty(L) \xrightarrow{\pi} \mathcal{C}^+(L) \rightarrow 0 & & 0 \rightarrow \mathcal{C}^-(L) \xrightarrow{U} \mathcal{C}^-(L) \xrightarrow{\pi} \widehat{\mathcal{C}}(L) \rightarrow 0 \\ \downarrow \pi & \downarrow \pi \circ U^{-1} & \downarrow \text{Id} \\ 0 \rightarrow \widehat{\mathcal{C}}(L) \xrightarrow{\iota} \mathcal{C}^+(L) \xrightarrow{U} \mathcal{C}^+(L) \rightarrow 0 & \text{and} & 0 \rightarrow \mathcal{C}^-(L) \xrightarrow{\iota} \mathcal{C}^\infty(L) \xrightarrow{\pi} \mathcal{C}^+(L) \rightarrow 0 \\ & & \downarrow \text{Id} \quad \downarrow U^{-1} \circ \iota \quad \downarrow \iota \end{array}$$

of short exact sequences and naturality of the snake lemma imply that the following diagrams commute:

$$(2.3) \quad \begin{array}{ccc} & \mathcal{H}^-(L) & \\ \partial \nearrow & & \searrow \pi_* \\ \mathcal{H}^+(L) & \xrightarrow{\partial} & \widehat{\mathcal{H}}(L) \end{array} \quad \text{and} \quad \begin{array}{ccc} & \mathcal{H}^+(L) & \\ \iota_* \nearrow & & \searrow \partial \\ \widehat{\mathcal{H}}(L) & \xrightarrow{\partial} & \mathcal{H}^-(L). \end{array}$$

For the empty link, $\mathcal{H}^-(\emptyset) \cong \mathbb{R}[U]$, $\mathcal{H}^\infty(\emptyset) \cong \mathbb{R}[U^{-1}, U]$, $\mathcal{H}^+(\emptyset) \cong \mathbb{R}[U^{-1}, U]/\mathbb{R}[U]$, and $\widehat{\mathcal{H}}(\emptyset) = \mathbb{R}$. More generally, for \mathcal{H}^∞ we have the following well-known result. (Recall that for the Lee deformation we assume 2 is invertible in \mathbb{R} .)

Proposition 2.1. *In the Bar-Natan theory, there is a canonical isomorphism*

$$(2.4) \quad \mathcal{H}^\infty(L) \cong \bigoplus_{o \in o(L)} \mathbb{R}[H^{-1}, H]$$

where $o(L)$ is the set of orientations of L . In the Lee theory, after adding a formal square root of T , there is a canonical isomorphism

$$(2.5) \quad \mathcal{H}^\infty(L) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}] \cong \bigoplus_{o \in o(L)} \mathbb{R}[T^{-\frac{1}{2}}, T^{\frac{1}{2}}].$$

In each case, the summand corresponding to an orientation o is supported in homological grading $2\text{lk}(L_o, L \setminus L_o)$, where L_o is the sublink of L consisting of components whose original orientations agree with o and lk is the linking number.

Proof. The proof is well-known (see [Lee05, BNM06, Tur20]), so we merely sketch it. The Bar-Natan Frobenius algebra $\mathbb{R}[H^{-1}, H][X]/(X^2 = XH)$ has a basis $\{A := X, B := H - X\}$ over $\mathbb{R}[H^{-1}, H]$, which diagonalizes it:

$$\begin{aligned} A^2 &= HA, & B^2 &= HB, & AB &= 0 \\ \Delta(A) &= A \otimes A, & \Delta(B) &= -B \otimes B. \end{aligned}$$

For the Lee case, after adding a formal square root of T , the Frobenius algebra $\mathbb{R}[T^{-\frac{1}{2}}, T^{\frac{1}{2}}][X]/(X^2 = T)$ has a basis $\{A := \sqrt{T} + X, B := \sqrt{T} - X\}$, which diagonalizes it:

$$\begin{aligned} A^2 &= 2\sqrt{T}A, & B^2 &= 2\sqrt{T}B, & AB &= 0 \\ \Delta(A) &= A \otimes A, & \Delta(B) &= -B \otimes B. \end{aligned}$$

In the Lee case, note that the homology of $\mathcal{C}^\infty(L) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}]$ is isomorphic to $\mathcal{H}^\infty(L) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}]$, since $\mathbb{R}[\sqrt{T}]$ is free over $\mathbb{R}[T]$.

For any vertex $v \in \{0, 1\}^N$ in the cube of resolutions, let L_v be the corresponding complete resolution of the link diagram L . With respect to the above basis, the chain group $\mathcal{C}^\infty(L)$ is freely generated (over $\mathbb{R}[H^{-1}, H]$ in the Bar-Natan case or $\mathbb{R}[T^{-\frac{1}{2}}, T^{\frac{1}{2}}]$ in the Lee case) by all possible labelings of the circles of L_v by $\{A, B\}$, for all v . Call two such generators *equivalent* if one can be obtained from the other by changing the resolutions at some crossings (0 to 1 or 1 to 0) so that the circles have consistent labelings before and after the change. That is, given resolutions L_v and L_w , there is a cobordism $\Sigma_{v,w}$ from L_v to L_w consisting of saddles at the crossings where v and w differ; a generator over v and a generator over w are equivalent if for each component Σ of $\Sigma_{v,w}$, all circles in the boundary of Σ have the same label.

Since the basis $\{A, B\}$ diagonalizes the Frobenius algebra, the complex $\mathcal{C}^\infty(L)$ decomposes as a direct sum along equivalence classes. Moreover, in each equivalence class, the complex is isomorphic to the tensor product of some number of copies of the two-step complex $\mathbb{R}[H^{-1}, H] \xrightarrow{\text{Id}} \mathbb{R}[H^{-1}, H]$

in the Bar-Natan case or $R[T^{-\frac{1}{2}}, T^{\frac{1}{2}}] \xrightarrow{\text{Id}} R[T^{-\frac{1}{2}}, T^{\frac{1}{2}}]$ in the Lee case. These complexes are acyclic, unless the tensor product is over zero copies, that is, the equivalence class contains just a single element. Therefore, the homology $\mathcal{H}^\infty(L)$ is generated by equivalence classes containing a single element, which are generators where every crossing connects two circles in the resolution with different labels.

Such generators, in turn, are in canonical correspondence with orientations of L , as follows. For any resolution L_v , consider the checkerboard coloring of $\mathbb{R}^2 \setminus L_v$ where the unbounded region is colored white. For any generator over v , orient each circle in L_v as the boundary of the black (respectively, white) region if it is labeled A (respectively, B). This orientation of L_v induces an orientation of L precisely for the above type of generators. The statement about the homological gradings is straightforward from the description of the generators above. \square

3. BEHAVIOR UNDER (POSSIBLY NONORIENTABLE) COBORDISMS

We will study the maps induced on these deformed Khovanov complexes and their homologies by a (possibly nonorientable) cobordism $F \subset [0, 1] \times S^3$ from an oriented link $L_0 \subset \{0\} \times S^3$ to an oriented link $L_1 \subset \{1\} \times S^3$. Throughout the paper, all link cobordisms will be assumed to be products near the boundary.

Recall the definition of the *normal Euler number*. Pick Seifert surfaces for the L_i and take a transverse pushoff F' of F so that the pushoff of L_i is in the direction of its Seifert surface. (It follows from the Mayer-Vietoris theorem applied to the decomposition $S^3 = \text{nbd}(L_i) \cup (S^3 \setminus L_i)$ that these pushoffs are independent of the choice of Seifert surfaces.) Then, the normal Euler number e of F is the signed count of intersection points between F and F' , where the signs come from picking a local orientation of F near each intersection point and using the induced local orientation of F' . This number is independent of the choice of pushoff. The normal Euler number is zero for oriented cobordisms from L_0 to L_1 and is some even number in general.

We will consider compact link cobordisms decorated with finitely many marked points which, to be consistent with Khovanov-Robert [KR22], we will call *stars*. So, an *elementary cobordism* between link diagrams is one of the following moves:

- (EC-1) A planar isotopy of the diagram.
- (EC-2) A Reidemeister move.
- (EC-3) A birth or death of an unknot disjoint from the rest of the diagram.
- (EC-4) No change to the link diagram but a choice of a distinguished point (star) in the interior of an arc of the diagram; we interpret this as the identity cobordism with a single star in its interior, lying over this distinguished point.
- (EC-5) A planar saddle.
- (EC-6) The identity cobordism from a link L to the same link with a different orientation on some components.

Associated to each elementary cobordism is a map of the Khovanov complexes. For Reidemeister moves, these are the maps from the proof of invariance of these theories. Specifically, Bar-Natan associates particular picture-world maps to each Reidemeister move [Bar05], and feeding these pictures into the Lee or Bar-Natan Frobenius algebra gives the map of deformed Khovanov complexes. The map associated to a birth is the unit 1 and associated to a death is the counit ϵ . The map

associated to a saddle is obtained by applying the corresponding multiplication m or comultiplication Δ to each resolution. The map associated to the identity cobordism with a star on some arc A multiplies the label of A , in each resolution, by $2X$ for the Lee deformation and $2X - H$ for the Bar-Natan deformation (compare [KR22, Formula (16)]). This map depends only on the arc containing the star, not the location of the star on that arc. The map associated to the identity cobordism with inconsistent orientations is the identity map.

Suppose F is a (possibly nonorientable) cobordism from L_0 to L_1 , with a finite number of marked stars in the interior of F . If F is represented by a movie of elementary cobordisms, then there is an induced map $\mathcal{C}^-(F): \mathcal{C}^-(L_0) \rightarrow \mathcal{C}^-(L_1)$, obtained by composing the maps associated to elementary cobordisms. This induces maps on all the four versions \mathcal{C}^\bullet of the Khovanov complexes, $\bullet \in \{+, -, \infty, \wedge\}$, as well as their homologies \mathcal{H}^\bullet .

Lemma 3.1. *The maps $\mathcal{C}^\bullet(F): \mathcal{C}^\bullet(L_0) \rightarrow \mathcal{C}^\bullet(L_1)$, $\bullet \in \{+, -, \infty, \wedge\}$, induce maps $\mathcal{H}^\bullet: \mathcal{H}(L_0) \rightarrow \mathcal{H}(L_1)$, and the long exact sequences from Formula (2.2) are natural with respect to these maps.*

Proof. This is immediate from the definitions. \square

Assuming the link diagrams are oriented coherently before and after the move, for planar isotopies and Reidemeister moves, the maps preserve the bigrading, and for births and deaths, the maps preserve gr_h and increase gr_q by 1. The map associated to a star preserves gr_h and decreases gr_q by 2. The behavior of saddles is more complicated.

Lemma 3.2. *Let F be a planar saddle from an oriented link diagram L_0 to an oriented link diagram L_1 , which is not necessarily orientable coherently with the orientations of L_0 and L_1 . Let e be its normal Euler number. Then, the map $\mathcal{C}^-(F): \mathcal{C}^-(L_0) \rightarrow \mathcal{C}^-(L_1)$ decreases gr_h by $e/2$ and decreases gr_q by $1 + 3e/2$.*

Proof. Ozsváth-Stipsicz-Szabó show that the normal Euler number e of the planar saddle F is $w(L_0) - w(L_1)$, where $w(L_i) = N_+(L_i) - N_-(L_i)$ is the writhe of the link diagram L_i [OSS17, Proof of Lemma 4.3]. They write their proof only for knots but, as we sketch in the next paragraph, it works equally well for links.

Fix any normal direction to the plane of projection of the link diagrams and consider a small pushoff of L_i in this normal direction; call this the *blackboard pushoff*. Since the total linking number of L_i with its blackboard pushoff is the writhe $w(L_i)$ while the total linking number of L_i with its Seifert pushoff is zero, the identity cobordism from L_i to L_i has a pushoff which intersects itself $w(L_i)$ times, so that the pushoff restricts to the Seifert pushoff at one end and the blackboard pushoff at the other. The planar saddle also has a (similarly defined) blackboard pushoff without any self-intersection, and it restricts to the blackboard pushoffs of L_0 and L_1 at the two ends. Putting these pieces together, we get a pushoff with $w(L_0) - w(L_1)$ self-intersections connecting the Seifert pushoffs of L_0 and L_1 .

Let N be the total number of crossings in either L_0 or L_1 . Recall that the complex $\mathcal{C}^-(L_i)$ is obtained from the total complex of a cube-shaped diagram—call it $\mathcal{C}'(L_i)$ —by increasing gr_h by $-N_-(L_i) = (w(L_i) - N)/2$ and gr_q by $N_+(L_i) - 2N_-(L_i) = (3w(L_i) - N)/2$. The saddle F induces a map $\mathcal{C}'(L_0) \rightarrow \mathcal{C}'(L_1)$ that preserves the homological grading and decreases the quantum grading by 1. Therefore, after the grading shifts, the map $\mathcal{C}^-(F): \mathcal{C}^-(L_0) \rightarrow \mathcal{C}^-(L_1)$ decreases gr_h by $(w(L_0) - w(L_1))/2 = e/2$ and decreases gr_q by $1 + 3(w(L_0) - w(L_1))/2 = 1 + 3e/2$. \square

Corollary 3.3. (*Compare [Bal, Proposition 4.7]*) *The map associated to a cobordism with Euler characteristic χ , normal Euler number e , and s stars decreases gr_h by $e/2$ and increases gr_q by $\chi - 3e/2 - 2s$.*

Proof. Consider a movie decomposition into elementary cobordisms (EC-1)–(EC-6). Choose orientations of all the link diagrams that appear in the movie, so that link diagrams before and after each of the moves (EC-1)–(EC-4) are oriented coherently. (We may choose the orientations inductively, starting with the given orientation of L_0 , and by using a move (EC-6) if necessary, we may ensure that the chosen orientation of L_1 agrees with the given orientation of L_1 .)

We now check that the statement holds for each of the elementary cobordisms. The elementary cobordisms (EC-1)–(EC-3) have $e = s = 0$, and the associated maps increase the bigrading by $(0, \chi) = (-e/2, \chi - 3e/2 - 2s)$. The elementary cobordism (EC-4) has $e = \chi = 0$ and $s = 1$, and the associated map increases the bigrading by $(0, -2) = (-e/2, \chi - 3e/2 - 2s)$. The elementary cobordism (EC-5) has $\chi = -1$ and $s = 0$, and by Lemma 3.2, the associated map increases the bigrading by $(-e/2, -1 - 3e/2) = (-e/2, \chi - 3e/2 - 2s)$. Finally, for cobordisms of type (EC-6), we have $\chi = s = 0$, and the associated identity map increases the bigrading by $(-e/2, -3e/2) = (-e/2, \chi - 3e/2 - 2s)$ as well. (The proof is similar to, but easier than, the proof of Lemma 3.2.) Since e , χ , and s are additive, the composition of these maps also increases the bigrading by $(-e/2, \chi - 3e/2 - 2s)$. \square

Remark 3.4. The corollary suggests that another natural grading is $\text{gr}_\gamma = \text{gr}_q - 3\text{gr}_h$: the map associated to any cobordism (possibly nonorientable) increases gr_γ by the Euler characteristic of the cobordism minus twice the number of stars.

We also recall a well-known result of Rasmussen’s [Ras10]. Given a link L , let $o(L)$ be the set of orientations of L . Similarly, given a cobordism F from L_0 to L_1 , let $o(F)$ be the set of orientations of F . There are restriction maps $o(L_0) \leftarrow o(F) \rightarrow o(L_1)$; that is, $o(F)$ is a *correspondence* from $o(L_0)$ to $o(L_1)$. (Here, we choose orientation conventions so that if $F = [0, 1] \times L$ then $o(F)$ is the identity correspondence of $o(L)$.)

Proposition 3.5. *Given a cobordism F from L_0 to L_1 , for the Bar-Natan and Lee theories, respectively, we have commutative diagrams*

$$\begin{array}{ccc}
\mathcal{H}^\infty(L_0) & \xrightarrow{\mathcal{H}^\infty(F)} & \mathcal{H}^\infty(L_1) \\
\downarrow \cong & & \cong \downarrow \\
\bigoplus_{o \in o(L_0)} \mathbb{R}[H^{-1}, H] & \xrightarrow{F_*} & \bigoplus_{o \in o(L_1)} \mathbb{R}[H^{-1}, H] \\
\\
\mathcal{H}^\infty(L_0) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}] & \xrightarrow{\mathcal{H}^\infty(F) \otimes \text{Id}} & \mathcal{H}^\infty(L_1) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}] \\
\downarrow \cong & & \cong \downarrow \\
\bigoplus_{o \in o(L_0)} \mathbb{R}[T^{-\frac{1}{2}}, T^{\frac{1}{2}}] & \xrightarrow{F_*} & \bigoplus_{o \in o(L_1)} \mathbb{R}[T^{-\frac{1}{2}}, T^{\frac{1}{2}}]
\end{array}$$

where the vertical arrows are from Proposition 2.1, and the bottom arrow is some map F_* that refines the correspondence $o(L_0) \leftarrow o(F) \rightarrow o(L_1)$. That is, for $i \in \{0, 1\}$ and for any orientation $o_i \in o(L_i)$ and any generator g_i of the $\mathbb{R}[H^{-1}, H]$ (respectively, $\mathbb{R}[T^{-\frac{1}{2}}, T^{\frac{1}{2}}]$) summand corresponding to o_i , the coefficient of g_1 in $F_*(g_0)$ is a sum $\sum_{o \in o(F), o|_{L_i}=o_i} e_o$, where each e_o is a unit in $\mathbb{R}[H^{-1}, H]$ (respectively, $\mathbb{R}[T^{-\frac{1}{2}}, T^{\frac{1}{2}}]$). In particular, if F is nonorientable, then in either theory, the map $\mathcal{H}^\infty(F): \mathcal{H}^\infty(L_0) \rightarrow \mathcal{H}^\infty(L_1)$ is zero.

Proof. For the first part, Rasmussen proved [Ras10] the result for Lee's deformation with $\mathbb{R} = \mathbb{Q}$ using a change of basis that diagonalized Lee's Frobenius algebra (after adding a square root of T). The change of basis from the proof of Proposition 2.1 diagonalizes the Frobenius algebra, both for the Bar-Natan theory (over any ring \mathbb{R}) and the Lee theory (over any ring \mathbb{R} with 2 invertible, and after adding a square root of T). Using this diagonalized basis, Rasmussen's proof goes through without any essential changes. (For an elementary star cobordism, the map is multiplication by $A - B$, which sends each orientation to itself with coefficient ± 1 , and hence fits into Rasmussen's framework.)

The last assertion is automatic for the Bar-Natan theory. For the Lee theory, note that if the map $\mathcal{H}^\infty(F) \otimes \text{Id}: \mathcal{H}^\infty(L_0) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}] \rightarrow \mathcal{H}^\infty(L_1) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}]$ is zero, then the map $\mathcal{H}^\infty(F): \mathcal{H}^\infty(L_0) \rightarrow \mathcal{H}^\infty(L_1)$ must be zero as well. This follows from commutativity of the diagram

$$\begin{array}{ccc} \mathcal{H}^\infty(L_0) & \xrightarrow{\mathcal{H}^\infty(F)} & \mathcal{H}^\infty(L_1) \\ \downarrow & & \downarrow \\ \mathcal{H}^\infty(L_0) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}] & \xrightarrow{\mathcal{H}^\infty(F) \otimes \text{Id}} & \mathcal{H}^\infty(L_1) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}] \end{array}$$

and noting that the rightmost vertical map

$$\mathcal{H}^\infty(L_1) \rightarrow \mathcal{H}^\infty(L_1) \otimes_{\mathbb{R}[T]} \mathbb{R}[\sqrt{T}] \cong \mathcal{H}^\infty(L_1) \otimes_{\mathbb{R}[T]} (\mathbb{R}[T] \oplus \sqrt{T}\mathbb{R}[T]) \cong \mathcal{H}^\infty(L_1) \oplus \sqrt{T}\mathcal{H}^\infty(L_1)$$

is the inclusion as the first factor, and therefore is injective. \square

Finally, we confirm well-definedness of the cobordism maps. Before stating the main result, we note some relations involving elementary star cobordisms (cobordisms of type (EC-4)).

Lemma 3.6. *Up to sign, the map on \mathcal{C}^- associated to an elementary star cobordism commutes with the map associated to any elementary cobordism disjoint from the star, and commutes with planar isotopies in general in the obvious sense. If p and q are points on opposite sides of a crossing then the map associated to the elementary star cobordism at p is chain homotopic to -1 times the map associated to the elementary star cobordism at q ; in particular, these two maps also agree up to homotopy and sign.*

Proof. The first statement is straightforward from the definitions. The second is immediate from a lemma of Hedden-Ni's [HN13, Lemma 2.3] in the Lee case and a lemma of Alishahi's [Ali19, Lemma 2.2] in the Bar-Natan case. \square

Well-definedness of the cobordism maps is the following:

Proposition 3.7. *Let $F \subset [0, 1] \times S^3$ be a (possibly nonorientable) cobordism from L_0 to L_1 . For $\bullet \in \{+, -, \infty, \hat{\cdot}\}$, the induced map $\mathcal{C}^\bullet(F): \mathcal{C}^\bullet(L_0) \rightarrow \mathcal{C}^\bullet(L_1)$ in the homotopy category of complexes over $\mathbb{R}[U]$ is well-defined up to sign, and invariant under isotopy of F in $[0, 1] \times S^3$ rel. boundary. In fact, if $\Phi: [0, 1] \times S^3 \rightarrow [0, 1] \times S^3$ is a diffeomorphism which is the identity near the boundary, then $\mathcal{C}^\bullet(F)$ and $\mathcal{C}^\bullet(\Phi(F))$ are chain homotopic.*

Proof. Since the maps on \mathcal{C}^+ , \mathcal{C}^∞ , and $\hat{\mathcal{C}}$ are induced by the map on \mathcal{C}^- , it suffices to prove the result for \mathcal{C}^- (compare Lemma 3.1). For isotopies of oriented cobordisms in $[0, 1] \times \mathbb{R}^3$, this follows from Bar-Natan's result [Bar05, Theorem 4] since both the Lee perturbation and the Bar-Natan perturbation can be obtained functorially from Bar-Natan's diagrammatic invariants. His proof works equally well for nonorientable cobordisms: any two movies representing isotopic nonorientable cobordisms also differ by a sequence of Carter-Saito's movie moves [CS93], every local movie move (between sequences of tangles) can be given a consistent orientation, and the map induced by a movie is independent of the choice of orientations. (In particular, we can suppress cobordisms of type (EC-6) in these movies.) Finally, functoriality for starred cobordisms follows easily from Lemma 3.6. (See, for instance, [Sar20, Lemma 2.1 and 4.1] for more details.)

To verify invariance under isotopies in $[0, 1] \times S^3$ we must also check invariance under Morrison-Walker-Wedrich's *sweep-around move* [MWW, Formula (1.1)]. Their proof works mutatis mutandis for the Lee and Bar-Natan deformations [MWW, Remark 2.2]. Nevertheless, for the sake of completeness, we present their proof adapted to our setting below. We will mostly use their notation, but a slightly different language.

Consider Morrison-Walker-Wedrich's picture [MWW, Formula (3.1)]. The picture shows two ways of moving a strand from top to bottom to get from a link diagram L to a link diagram L' : in the first method, this strand moves in front of the rest of the link, while in the second, it moves behind. Let L_+^i (respectively, L_-^i) denote the link diagram at the i^{th} stage in the first (respectively, second) method. The two sequences of link diagram L_\pm^i produce two chain maps $\mathcal{C}^-(L) \rightarrow \mathcal{C}^-(L')$ by composing maps associated to Reidemeister moves. By choosing the Reidemeister maps carefully, we will show that the two maps agree on the nose.

Recall that for any link diagram, the homological grading of any resolution is given by the number of crossings that have been resolved as the 1-resolution minus the total number of negative crossings in the diagram. For the link diagrams that appear above, call a crossing *external* (respectively, *internal*) if it involves (respectively, does not involve) the moving horizontal strand, and define the *external grading* (respectively, *internal grading*) of any resolution to be the number of external (respectively, internal) crossings that have been resolved as the 1 resolution minus the total number of negative external (respectively, internal) crossings. The differential preserves or increases the external grading.

For any of the above link diagrams, let \mathcal{C}_0^- denote the subgroup of the chain group that lives in external grading 0. Note that $\mathcal{C}_0^-(L) = \mathcal{C}^-(L)$ and $\mathcal{C}_0^-(L') = \mathcal{C}^-(L')$ since these diagrams have no external crossings. Also note that there is a natural isomorphism of groups $\mathcal{C}_0^-(L_+^i) \cong \mathcal{C}_0^-(L_-^i)$ since any resolution of L_+^i in external grading 0, when viewed as a resolution of L_-^i , is also in external grading 0. (This uses the fact that Morrison-Walker-Wedrich start with a braid closure.)

The Reidemeister maps will be chosen in such a way that they will preserve or decrease the external grading. Since the composition is a map $\mathcal{C}_0^-(L) \rightarrow \mathcal{C}_0^-(L')$, it is therefore enough to consider

the portion of the maps that preserve the external grading, namely, the maps

$$\mathcal{C}^-(L) \rightarrow \mathcal{C}_0^-(L_\pm^0) \rightarrow \cdots \rightarrow \mathcal{C}_0^-(L_\pm^i) \rightarrow \cdots \rightarrow \mathcal{C}_0^-(L_\pm^\ell) \rightarrow \mathcal{C}^-(L').$$

(These individual maps are typically not chain maps; however, perhaps surprisingly, that is irrelevant to the proof.) Furthermore, these components of the maps will commute with the isomorphisms $\mathcal{C}_0^-(L_+^i) \cong \mathcal{C}_0^-(L_-^i)$, that is, the following diagram will commute:

$$\begin{array}{ccccccc} & & \mathcal{C}_0^-(L_+^0) & \rightarrow \cdots & \mathcal{C}_0^-(L_+^i) & \rightarrow \cdots & \mathcal{C}_0^-(L_+^\ell) & & \\ & \nearrow & \downarrow \cong & & \downarrow \cong & & \downarrow \cong & \searrow & \\ \mathcal{C}^-(L) & & \mathcal{C}_0^-(L_-^0) & \rightarrow \cdots & \mathcal{C}_0^-(L_-^i) & \rightarrow \cdots & \mathcal{C}_0^-(L_-^\ell) & & \mathcal{C}^-(L') \end{array}$$

That will establish that the two compositions agree on the nose.

For the Reidemeister I and II moves at the beginning and end of the sequence, the above check is almost automatic. The maps preserve the homological grading. Since all the crossings involved are external, the maps also preserve the internal grading, and therefore preserve the external grading as well.

The Reidemeister III move requires more care. The proof is illustrated in Figure 3.1. Assume L_\pm^{i+1} is obtained from L_\pm^i by moving the horizontal strand past an internal crossing, as shown in the figure. The other Reidemeister III move is obtained by mirroring all the link diagrams, and the proof in that case follows formally from the following proof by reversing all arrows.

The 3-dimensional cubes of resolution for the four link diagrams L_\pm^* , $* \in \{i, i+1\}$, are shown. The two external crossings are numbered 1 and 2—left to right for L_+^* and right to left for L_-^* —and the internal crossing is numbered 3. The eight vertices in each cube of resolutions decompose according to the local external grading, which is the sum of the first two coordinates of the vertices (up to a shift); this is shown by boxing them with a **dashed**, **solid**, or **dotted** line. The differentials are shown in light gray.

The Reidemeister maps go from the cube of resolution of L_\pm^i to the cube of resolution of L_\pm^{i+1} . The maps either preserve or decrease the external grading. We are only interested in the maps which preserve the external grading, so we have drawn them in black, and the other maps in light gray. The maps (and the differentials) are decorated with the cobordisms that induce them, with s , b , d being shorthand for saddle, birth, and death, respectively.

The top row (corresponding to L_+^*) is essentially a copy of Bar-Natan's picture [Bar05, Figure 9]; we have merely rotated Bar-Natan's tangle so that his vertical over-strand has become our horizontal moving strand, and we have reordered the crossings as well, so our signs differ from Bar-Natan's. For example, the surface highlighted in Bar-Natan's picture corresponds to our map from the 010 vertex of L_+^i to the 100 vertex of L_-^{i+1} ; it is decorated $-dssb$, so it is the negative of a death, followed by two saddles (which are easy to figure out from the diagrams), followed by a birth.

The bottom row (corresponding to L_-^*) is also obtained from Bar-Natan's picture [Bar05, Figure 9]; this time we have rotated Bar-Natan's tangle so that his northwest-to-southeast under-strand has become our horizontal moving strand, and once again, we have reordered the crossings. The diagram thus obtained is not quite the bottom row of our figure: it does not have the map from

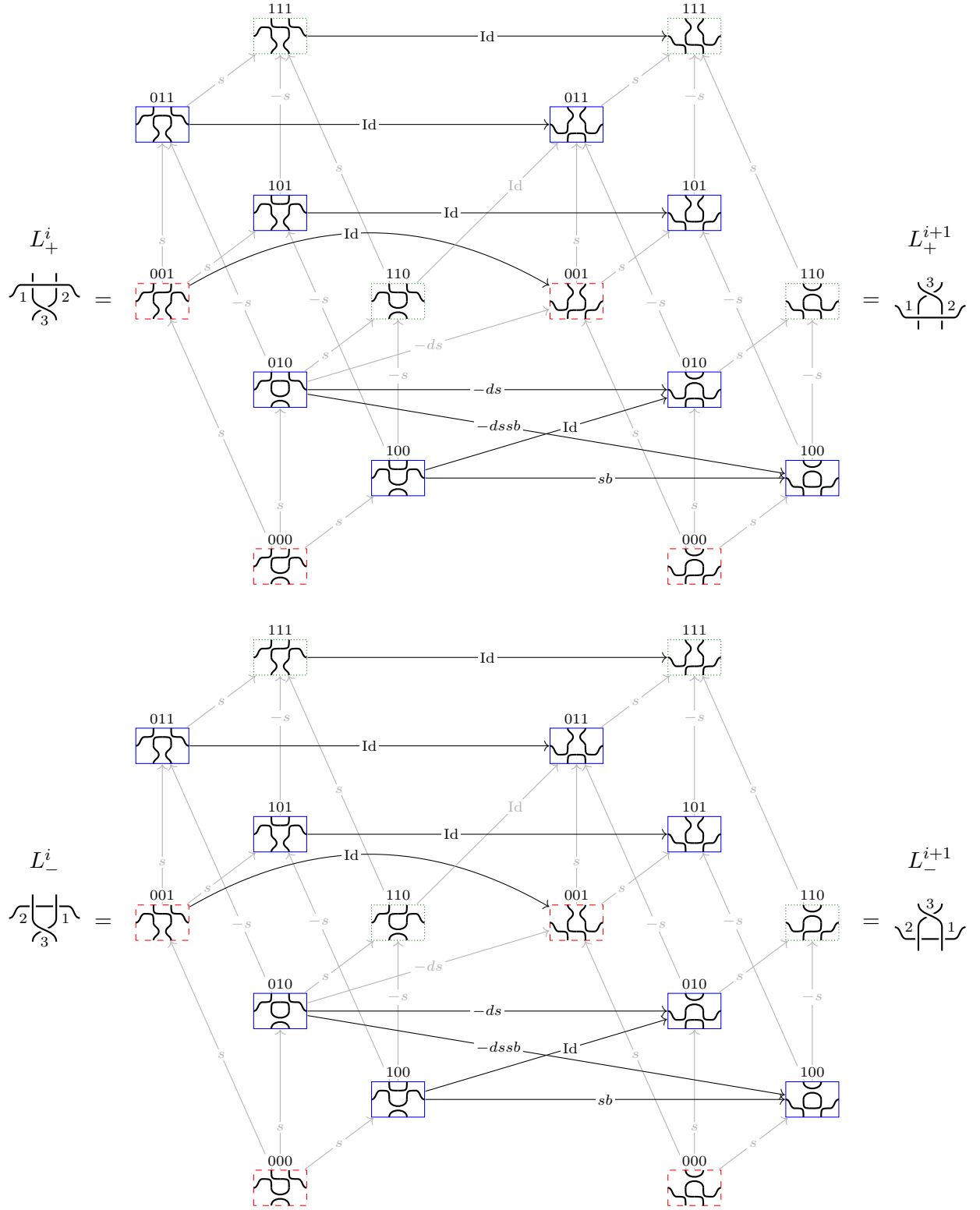


FIGURE 3.1. The Reidemeister III move during the proof of invariance under the sweep-around move.

the 100 vertex of L_-^i to the 100 vertex of L_-^{i+1} , nor the map from the 101 vertex of L_-^i to the 101 vertex of L_-^{i+1} , but instead has maps from the 001 vertex of L_-^i to the 100 vertex of L_-^{i+1} and from the 101 vertex of L_-^i to the 110 vertex of L_-^{i+1} . The latter maps increase the external gradings, so we modify the chain map by a null-homotopy $\partial f + f\partial$ to get to our diagram, where f is the map from the 101 vertex of L_-^i to the 100 vertex of L_-^{i+1} corresponding to a birth.

The natural isomorphism $\mathcal{C}_0^-(L_+^*) \xrightarrow{\cong} \mathcal{C}_0^-(L_-^*)$ sends the **dotted** vertices to the **dashed** vertices and vice-versa, and sends the **solid** vertices to the corresponding **solid** vertices. These isomorphisms commute with the Reidemeister maps that preserve external gradings (which are the black arrows in the figure). This gives invariance under the sweep-around move.

Finally, the proof that the map is invariant under diffeomorphisms follows from another argument of Morrison-Walker-Wedrich [MWW, Section 4.2]; we refer the reader there, though a key point in the proof is quoted as Lemma 4.7, below. \square

Remark 3.8. The Khovanov Frobenius algebra—the $U = 0$ specialization of the Lee and Bar-Natan algebras—corresponds to a (1+1)-dimensional TQFT. It is natural to ask if this TQFT extends to non-orientable cobordisms. TQFTs for oriented 1-manifolds but allowing non-orientable cobordisms are called *Klein TQFTs* [AN06] or *unoriented TQFTs* [TT06]. Unoriented (1+1)-dimensional TQFTs correspond to Frobenius algebras V with extra structure: an element $\theta \in V$ corresponding to a Möbius band and an involution ϕ corresponding to the mapping cylinder of an orientation-reversing involution of S^1 , satisfying the conditions that $\phi(m(\theta, v)) = m(\theta, v)$ for all $v \in V$ and $(m \circ (\phi \otimes \text{Id}) \circ \Delta)(1) = m(\theta, \theta)$ [TT06, Proposition 2.9]. For the Khovanov TQFT, the fact that ϕ respects the unit and counit implies that $\phi = \text{Id}$, so the second identity implies that $m(\theta, \theta) = 2X$ which is impossible (cf. [TT06, Section 4.2]). It is possible to extend V to a projective unoriented TQFT, by defining $\theta = 0$, $\phi(1) = 1$, and $\phi(X) = -X$; the map ϕ only respects the counit up to sign.

Imitating part of this argument, we can see that it is impossible to remedy the sign ambiguity in for non-orientable surfaces (without equipping the surfaces with some extra data). There is a movie which starts with a 0-crossing unknot, performs a Reidemeister I move on half of it, introducing one crossing, then performs a Reidemeister I move on the other half eliminating the crossing. The induced map $V \rightarrow V$ is either $(1 \mapsto 1, X \mapsto -X)$ or $(1 \mapsto -1, X \mapsto X)$. One can compute this directly, but it is also forced by the fact that the invariant of a once-punctured Klein bottle is zero (see the proof of Corollary 6.13): the map associated to a once-punctured Klein bottle factors as a birth, then a split, then applying the map just described to one of the two circles, and then a merge. (This is an embedded version of the proof of the relation $(m \circ (\phi \otimes \text{Id}) \circ \Delta)(1) = m(\theta, \theta)$.) However, following the first option by a death (counit) gives a cobordism isotopic to a death, but sending $X \mapsto -1$ instead of $X \mapsto 1$; and preceding the second option by a birth gives a cobordism isotopic to a birth, but sending $1 \mapsto -1$.

(Note that, in the construction of the Khovanov cube, all the surfaces that arise are orientable; in fact, a checkerboard coloring of the knot projection induces an orientation of them. So, to construct the Khovanov cube one does not need the extension of the TQFT to non-orientable surfaces. See also [TT06] for further discussion.)

Mikhail Khovanov informs us that Greg Kuperberg mentioned to him around 2003 that Khovanov homology is functorial with respect to nonorientable cobordisms in $[0, 1] \times \mathbb{R}^3$, up to a sign, which is part of Proposition 3.7, above.

4. ADMISSIBLE CUTS

Any compact, connected, nonorientable surface F is diffeomorphic to $(\#^g \mathbb{RP}^2) \# (\#^k \mathbb{D}^2)$, where $k = |\pi_0(\partial F)|$ is the number of boundary components. The number $g = 2 - \chi(F) - k$ is called the *crosscap number* of F . For any surface F (not necessarily connected), define its crosscap number to be the sum of the crosscap numbers of its nonorientable components.

Definition 4.1. Fix a small $\epsilon > 0$. Let $F \subset [0, 1] \times S^3$ be a nonorientable cobordism from L_0 to L_1 , which is a product near the boundary. An *admissible cut* for F consists of the data (S, V, ϕ) , where:

- $S \subset (0, 1) \times S^3$ is a smoothly embedded 3-manifold;
- $V \subset (0, 1) \times S^3$ is a tubular neighborhood of S ; and
- $\phi: V \rightarrow (\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon) \times S^3$ is a diffeomorphism,

satisfying:

- (AC-1) ϕ takes S to $\{\frac{1}{2}\} \times S^3$ and $F \cap V$ to a product cobordism;
- (AC-2) the intersection of F with each of the 2 components of $([0, 1] \times S^3) \setminus S$ is nonorientable; and
- (AC-3) there exists a diffeomorphism $\Phi: ([0, 1] \times S^3, V) \xrightarrow{\cong} ([0, 1] \times S^3, (\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon) \times S^3)$, which is the identity near the boundary and agrees with ϕ on V .

Call a pair of admissible cuts (S, V, ϕ) and (S', V', ϕ') for F *elementary equivalent* if $V \cap V' = \emptyset$ and there is a diffeomorphism

$$(4.1) \quad \begin{aligned} ([0, 1] \times S^3, V, V') &\cong ([0, 1] \times S^3, (\tfrac{1}{3} - \epsilon, \tfrac{1}{3} + \epsilon) \times S^3, (\tfrac{2}{3} - \epsilon, \tfrac{2}{3} + \epsilon) \times S^3) \quad \text{or} \\ ([0, 1] \times S^3, V', V) &\cong ([0, 1] \times S^3, (\tfrac{1}{3} - \epsilon, \tfrac{1}{3} + \epsilon) \times S^3, (\tfrac{2}{3} - \epsilon, \tfrac{2}{3} + \epsilon) \times S^3), \end{aligned}$$

which is the identity near the boundary and which agrees with ϕ and ϕ' on V and V' , respectively, after post-composition by a translation in the first factor. Call admissible cuts (S, V, ϕ) and (S', V', ϕ') for a pair of surfaces F and F' *diffeomorphic* if there is a diffeomorphism $\Psi: ([0, 1] \times S^3, F, V) \xrightarrow{\cong} ([0, 1] \times S^3, F', V')$ which is the identity near the boundary and satisfies $\phi' \circ \Psi = \phi$ on V . Call admissible cuts (S, V, ϕ) and (S', V', ϕ') for a pair of surfaces F and F' *equivalent* if they differ by a sequence of elementary equivalences and diffeomorphisms.

Proposition 4.2. *Suppose F is a cobordism (not necessarily connected) with crosscap number ≥ 2 . Then, F has an admissible cut. Further, if F has crosscap number ≥ 3 , and if F' is obtained from F by a self-diffeomorphism of $[0, 1] \times S^3$ which is the identity near the boundary, then any admissible cut for F is equivalent to any admissible cut for F' .*

We recall some results about the curve complex of nonorientable surfaces before proving Proposition 4.2.

Let F be a compact, nonorientable surface. Consider the long exact sequence for the pair $(F, \partial F)$,

$$\tilde{H}^0(\partial F; \mathbb{F}_2) \rightarrow H^1(F, \partial F; \mathbb{F}_2) \rightarrow H^1(F; \mathbb{F}_2) \rightarrow H^1(\partial F; \mathbb{F}_2).$$

The first Stiefel-Whitney class $w_1(TF) \in H^1(F; \mathbb{F}_2)$ maps to zero in $H^1(\partial F; \mathbb{F}_2)$ (since $TF|_{\partial F}$ is orientable), hence is in the image of $H^1(F, \partial F; \mathbb{F}_2)$. Call a closed curve α in the interior of F *complement-orientable* if $\text{PD}([\alpha]) \in H^1(F, \partial F; \mathbb{F}_2)$ maps to $w_1(TF)$; assuming α is embedded, this is equivalent to the condition that $F \setminus \alpha$ is orientable, since for any other curve β , $\langle w_1(TF), \beta \rangle = \alpha \cdot \beta \pmod{2}$. Call the other curves *complement-nonorientable*. Call a closed curve $\alpha \subset F$ *one-sided* if $\langle w_1(TF), [\alpha] \rangle = 1$; assuming α is embedded, this is equivalent to $TF|_{\alpha}$ being a Möbius band.

Lemma 4.3. *Let $\alpha \subset F$ be a complement-orientable, embedded circle. Then, F has a single nonorientable component F_0 . Moreover, α is one-sided if and only if the crosscap number of F_0 (equivalently, F) is odd.*

Proof. By hypothesis, $[\alpha] = \text{PD}(w_1(TF))$. If F has multiple nonorientable components, then $[\alpha]$ cannot be represented by a single curve. For the second part, we have to calculate $\langle w_1(TF), [\alpha] \rangle = \langle w_1(TF_0) \cup w_1(TF_0), [F_0] \rangle = \alpha \cdot \alpha \pmod{2}$. By classification of surfaces and a direct computation, this number equals the parity of the crosscap number of F_0 . \square

The *one-sided curve complex* of F is the graph with vertices isotopy classes of embedded, one-sided curves α in the interior of F and an edge from α to β if and only if there are disjoint representatives of α and β . The *restricted one-sided curve complex* is the full sub-graph spanned by the complement-nonorientable one-sided curves α . (By Lemma 4.3, if F has multiple nonorientable components or if the crosscap number of F is even, then the restricted one-sided curve complex is the same as the one-sided curve complex.)

Proposition 4.4. *Let F be a compact, nonorientable surface (with boundary) of crosscap number ≥ 2 . Then, the restricted one-sided curve complex of F is connected.*

Proof. This is essentially due to Pieloch [Pie16, Proposition 2.7], and we follow his argument.

If F has more than one nonorientable connected component then the statement is obvious. So, it suffices to prove the result when F is a connected surface with crosscap number ≥ 2 .

Let θ be the one-sided curve shown in Figure 4.1 and let λ be any other complement-nonorientable one-sided curve. From the classification of surfaces, there is a homeomorphism from F to itself sending θ to λ . So, it suffices to show that the mapping class group of F takes the path component of θ in the restricted one-sided curve complex to itself. For that, it suffices to show that a set of generators for the mapping class group take this path component to itself, i.e., take θ to curves which can be connected to θ in the restricted one-sided curve complex.

Here, we will not require homeomorphisms to be the identity on ∂F or to take boundary components to themselves. In fact, since deleting ∂F has no effect on the restricted one-sided curve complex, we can view F as a punctured surface and the homeomorphism as an element of the mapping class group of the punctured surface F . This mapping class group was studied by Korkmaz [Kor02], who denoted it $\mathcal{M}_{g,n}$, where g is the crosscap number and n is the number of punctures z_1, \dots, z_n .

In particular, Korkmaz gave a set of generators for this mapping class group [Kor02, Section 4]. There are three cases, depending on the crosscap number: crosscap number 2, $2k+1$ for $k \geq 1$, or $2k$ for $k > 1$. In each case, the mapping class group is generated by a finite set of Dehn twists, braid generators in the z_i , one crosscap slide (see [Kor02, Section 2] and the references he gives for

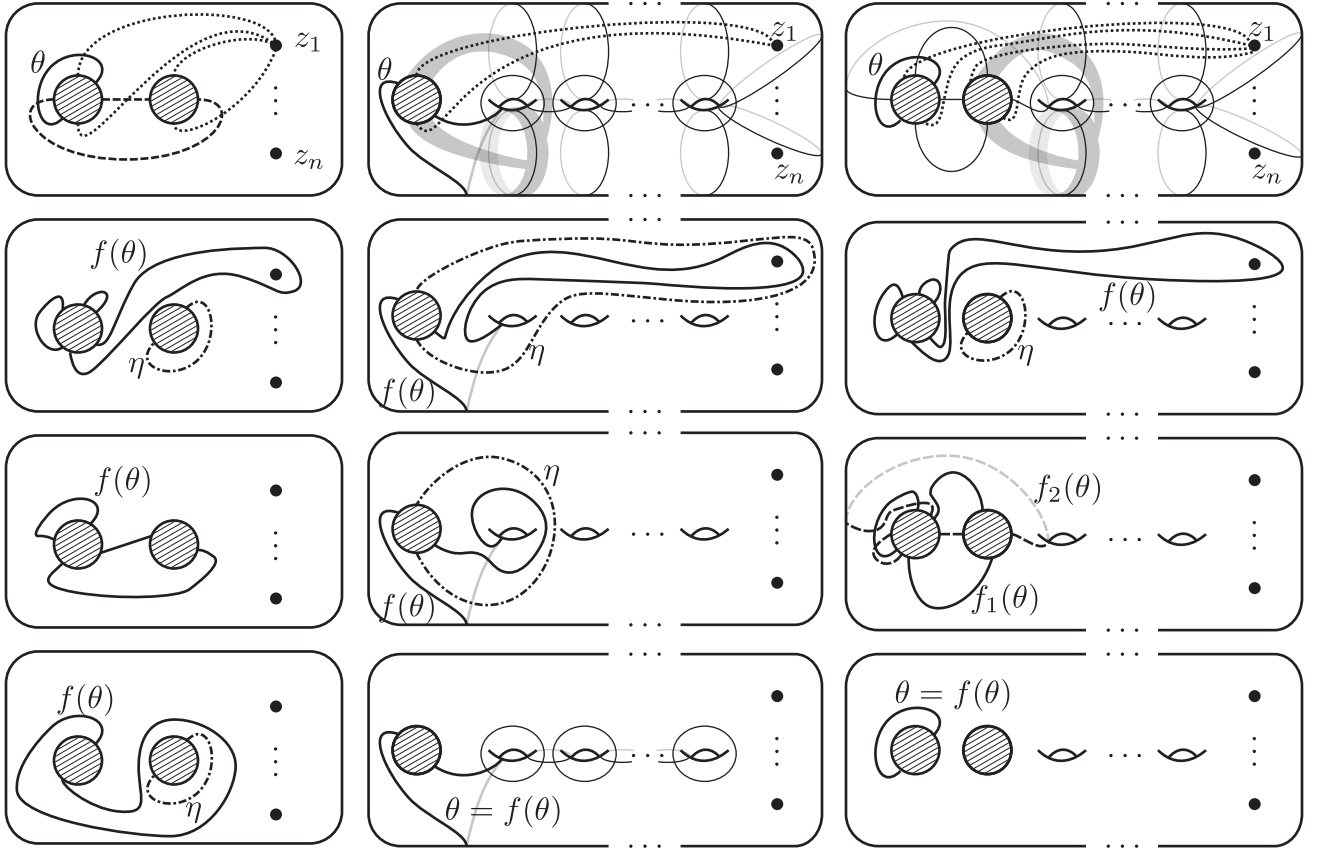


FIGURE 4.1. **Generators for the mapping class group of a punctured surface.** The left column is the crosscap number 2 case, center is the crosscap number $2k + 1$ ($k \geq 1$) case, and right is the crosscap number $2k$ ($k > 1$) case. Crosscaps are shaded, and hidden lines are gray. Top: the mapping class group is generated by elementary braids in the z_i , Dehn twists around the thin curves and the dashed curve, boundary slides (push maps) along the dotted curves, and a crosscap slide. In the left picture, the crosscap slide pulls one crosscap along the dashed curve. In the other pictures, the crosscap slide occurs in the shaded region (a punctured Klein bottle). Second row: the image of θ under the boundary slide that moves it. Third row: the images of θ under the Dehn twist(s) that move it. Bottom row: the image of θ under the crosscap slide. When $f(\theta)$ is neither disjoint from nor equal to θ , a third curve η disjoint from both is shown (dash-dotted).

the definition), and one or two boundary slides (again, see [Kor02, Section 2]); the generators are shown in Figure 4.1. In each case, most of the generators fix θ . The remaining ones either take θ to a curve disjoint from θ (up to isotopy) or to a curve $f(\theta)$ so that there is a third curve η disjoint from both θ and $f(\theta)$. See Figure 4.1. So, in all cases, $f(\theta)$ lies in the same path component as θ , as desired. \square

We note an easier lemma:

Lemma 4.5. *If F is a nonorientable surface of crosscap number > 1 (i.e., is not a punctured \mathbb{RP}^2 union an orientable surface) then the restricted curve complex has at least two points. If F has crosscap number > 2 (i.e., is also neither a punctured Klein bottle nor a punctured $\mathbb{RP}^2 \amalg \mathbb{RP}^2$, union an orientable surface) then the restricted curve complex contains a 3-cycle.*

Proof. This is straightforward from the classification of surfaces, and is left to the reader. \square

Proposition 4.4 and Lemma 4.5 together imply the following.

Lemma 4.6. *Let F be a compact, nonorientable surface with crosscap number > 2 . For any complement-nonorientable one-sided embedded curves α, β in F , there exists an even-length sequence $\alpha = \gamma_0, \gamma_1, \dots, \gamma_{2n} = \beta$ of complement-nonorientable one-sided embedded curves connecting α to β so that every pair of consecutive curves γ_i, γ_{i+1} are disjoint.*

Proof. Using Proposition 4.4, we may choose a walk in the restricted curve complex connecting α to β . Using the 3-cycle from the second part of Lemma 4.5 if needed, we may ensure that the walk has even length. Choose embedded curves representing the vertices of the walk to get a sequence $\gamma_0, \gamma_1, \dots, \gamma_{2m}$.

We may choose $\gamma_0 = \alpha$ and we may choose γ_i inductively to ensure that it is disjoint from γ_{i-1} . The final curve γ_{2m} will be isotopic to β , but need not equal β . Let ϕ_t for $t \in [0, 1]$ be an ambient isotopy taking γ_{2m} to β .

Using the first part of Lemma 4.5, choose a complement-nonorientable one-sided embedded curve δ which is disjoint from γ_{2m} . To finish the proof, we will construct a sequence $\gamma_{2m}, \gamma_{2m+1}, \dots, \gamma_{2n} = \beta$ of complement-nonorientable one-sided embedded curves with every consecutive pair disjoint, as follows:

$$\gamma_{2i} = \phi_{\frac{i-m}{n-m}}(\gamma_{2m}), \quad m \leq i \leq n \qquad \gamma_{2i+1} = \phi_{\frac{i-m}{n-m}}(\delta), \quad m \leq i < n.$$

Clearly, γ_{2i+1} is disjoint from γ_{2i} , and by compactness, for n large enough, it will be disjoint from γ_{2i+2} as well. For instance, fix a metric on F so that length of $\frac{\partial}{\partial t}\phi_t$ is bounded above by 1; let $D = \min_{t \in [0,1]} \text{dist}(\phi_t(\gamma_{2m}), \phi_t(\delta))$; then $n > m + (1/D)$ suffices. \square

Finally we need the following analogue of a result of Morrison-Walker-Wedrich's [MW], Lemma 4.7]:

Lemma 4.7. *Let $\Sigma \subset [0, 1] \times S^3$ be a link cobordism and $f: [0, 1] \times S^3 \rightarrow [0, 1] \times S^3$ a diffeomorphism which is the identity near the boundary. Then, Σ is isotopic to $f(\Sigma)$, and the isotopy may also be assumed to be the identity near the boundary.*

Proof. This is proved by replacing \mathbb{R}^3 by S^3 throughout Morrison-Walker-Wedrich's proof [MWW, Lemma 4.7], but for completeness, we sketch the proof below.

Let U be a collar neighborhood of $\{0, 1\} \times S^3$ on which f is the identity. Fix a point $p \in S^3$ so that $[0, 1] \times \{p\}$ is disjoint from Σ . By postcomposing f by an isotopy that takes $f([0, 1] \times \{p\})$ to $[0, 1] \times \{p\}$ (and is the identity on U), we may assume f is the identity on $[0, 1] \times \{p\}$. Let $B \subset S^3$ be a ball around p with $[0, 1] \times B$ disjoint from Σ . Let ϕ be a diffeomorphism of $[0, 1] \times S^3$ which is the identity on $U \cup ([0, 1] \times (\{p\} \cup (S^3 \setminus B)))$, so that on the normal bundle of $[0, 1] \times \{p\}$ inside $[0, 1] \times S^3$ (which is a trivial \mathbb{R}^3 bundle in an obvious way) $d\phi$ induces the non-trivial element of $\pi_1(\text{SO}(3))$. By post-composing f by ϕ if necessary, and a further small isotopy, we may assume f is the identity on $[0, 1] \times B$. Now let g_t be an isotopy of $[0, 1] \times S^3$ which is the identity near the boundary, so that $g_0 = \text{Id}$ and $g_1(\Sigma) \subset U \cup ([0, 1] \times B)$. Then, the isotopy g_t takes Σ to $g_1(\Sigma)$ and the isotopy $f \circ g_{1-t}$ takes $g_1(\Sigma) = f(g_1(\Sigma))$ to $f(\Sigma)$. \square

Proof of Proposition 4.2. Let $\pi: [0, 1] \times S^3 \rightarrow S^3$ be projection. Given a curve $\gamma \subset F$, let

$$C_{\leq \gamma} = \{(t, p) \in [0, 1] \times S^3 \mid \exists t' \in [0, 1] \text{ so that } (t', p) \in \gamma \text{ and } t \leq t'\}$$

$$C_{\geq \gamma} = \{(t, p) \in [0, 1] \times S^3 \mid \exists t' \in [0, 1] \text{ so that } (t', p) \in \gamma \text{ and } t \geq t'\},$$

so $\pi^{-1}(\pi(\gamma)) = C_{\leq \gamma} \cup C_{\geq \gamma}$ and if $\pi|_{\gamma}$ is injective then $C_{\leq \gamma} \cap C_{\geq \gamma} = \gamma$.

For the first statement, that F has an admissible cut, choose disjoint one-sided embedded curves $\gamma, \eta \subset F$; this is possible by Lemma 4.5. Perturbing F slightly, we may assume:

- (G-1) $\pi|_{\gamma}$ is injective,
- (G-2) $d\pi$ restricted to $TF|_{\gamma}$ has rank 2 everywhere,
- (G-3) F intersects $C_{\leq \gamma}$ transversely away from γ ,
- (G-4) $\pi|_{\eta}$ is injective,
- (G-5) $d\pi$ restricted to $TF|_{\eta}$ has rank 2 everywhere,
- (G-6) F intersects $C_{\geq \eta}$ transversely away from η , and
- (G-7) $\pi(\gamma) \cap \pi(\eta) = \emptyset$.

Let $U_{\leq \gamma}, U_{\geq \eta}$ be tubular neighborhoods of $C_{\leq \gamma}$ and $C_{\geq \eta}$ with disjoint closures. Choose $U_{\leq \gamma}$ small enough that $U_{\leq \gamma} \cap F$ consists of a Möbius band around γ and disks around the other points in $C_{\leq \gamma} \cap F$. Choose $U_{\geq \eta}$ similarly. Choose $\epsilon > 0$ small enough that

$$([0, \epsilon] \times S^3) \cup \overline{U_{\leq \gamma}} \cap ([1 - \epsilon, 1] \times S^3) \cup \overline{U_{\geq \eta}} = \emptyset.$$

Let

$$S_{\leq \gamma} = \partial([0, \epsilon] \times S^3) \cup \overline{U_{\leq \gamma}} \setminus (\{0\} \times S^3), \quad S_{\geq \eta} = \partial([1 - \epsilon, 1] \times S^3) \cup \overline{U_{\geq \eta}} \setminus (\{1\} \times S^3).$$

Then, after smoothing corners, both $S_{\leq \gamma}$ and $S_{\geq \eta}$ are admissible cuts for F , since either side of either cut is nonorientable (containing one of the one-sided curves γ or η). The diffeomorphisms ϕ from Definition 4.1 for $S_{\leq \gamma}$ and $S_{\geq \eta}$ are induced from the natural isotopies that take γ to $\{\epsilon\} \times \pi(\gamma)$ along $C_{\leq \gamma}$ and η to $\{1 - \epsilon\} \times \pi(\eta)$ along $C_{\geq \eta}$.

For the second statement, fix admissible cuts S and S' for F and F' and choose diffeomorphisms $\Phi: ([0, 1] \times S^3, S) \xrightarrow{\cong} ([0, 1] \times S^3, \{\frac{1}{2}\} \times S^3)$ and $\Phi': ([0, 1] \times S^3, S') \xrightarrow{\cong} ([0, 1] \times S^3, \{\frac{1}{2}\} \times S^3)$ as in Definition 4.1. It is enough to show that the admissible cuts $\{\frac{1}{2}\} \times S^3$ for $\Phi(F)$ and $\Phi'(F')$ are equivalent.

Choose one-sided curves $\gamma, \eta \subset \Phi(F)$ on the left and right side of $\{\frac{1}{2}\} \times S^3$ and choose a one-sided curve $\alpha \subset \Phi'(F')$ on the left side of $\{\frac{1}{2}\} \times S^3$. Consider the admissible cuts $S_{\leq \gamma}$ and $S_{\geq \eta}$ for $\Phi(F)$ as defined above. They are both elementary equivalent to the cut $\{\frac{1}{2}\} \times S^3$ (as well as to each other). Similarly the admissible cut $S_{\leq \alpha}$ for $\Phi'(F')$ is elementary equivalent to the cut $\{\frac{1}{2}\} \times S^3$. In particular, it is enough to show that the cut $S_{\leq \gamma}$ for $\Phi(F)$ is equivalent to the cut $S_{\leq \alpha}$ for $\Phi'(F')$.

Using Lemma 4.7, choose an ambient isotopy $\psi_t: [0, 1] \times S^3 \rightarrow [0, 1] \times S^3$ which is the identity near the boundary, with $\psi_0 = \text{Id}$, and $\psi_1(\Phi(F)) = \Phi'(F')$. We may perturb the isotopy to ensure that the genericity assumptions (G-1)–(G-7) hold for the curves $\psi_t(\gamma), \psi_t(\eta) \subset \psi_t(\Phi(F))$, except for finitely many $t \in (0, 1)$ when exactly one of them fails. Choose non-exceptional points $0 = t_0 < t_1 < \dots < t_{m-1} < t_m = 1$ such that each interval $[t_i, t_{i+1}]$ contains at most one such exceptional point; call such an interval γ -good (respectively, η -good) if the genericity conditions (G-1)–(G-3) (respectively, (G-4)–(G-6)) hold on the interval. Since at most one genericity condition fails at each exceptional point, each of these intervals $[t_i, t_{i+1}]$ is either γ -good or η -good (or both, if (G-7) fails).

At non-exceptional t (such as t_0, \dots, t_m), both the cuts $S_{\leq \psi_t(\gamma)}$ and $S_{\geq \psi_t(\eta)}$ for $\psi_t(\Phi(F))$ are admissible, and they are elementary equivalent to each other. We need the following lemma.

Lemma 4.8. *For any γ -good (respectively, η -good) interval $[t_i, t_{i+1}]$, the cuts $S_{\leq \psi_{t_i}(\gamma)}$ (respectively, $S_{\geq \psi_{t_i}(\eta)}$) for $\psi_{t_i}(\Phi(F))$ and $S_{\leq \psi_{t_{i+1}}(\gamma)}$ (respectively, $S_{\geq \psi_{t_{i+1}}(\eta)}$) for $\psi_{t_{i+1}}(\Phi(F))$ are diffeomorphic.*

Proof. We will explain the γ -good case; the η -good case is similar. For notational convenience, let $a = t_i$, $b = t_{i+1}$, $F_a = \psi_a(\Phi(F))$, $\gamma_a = \psi_a(\gamma)$, $F_b = \psi_b(\Phi(F))$, and $\gamma_b = \psi_b(\gamma)$. Note also that, in the definition of $S_{\leq \gamma_a}$ and $S_{\leq \gamma_b}$, up to diffeomorphism, we can choose the collar neighborhoods of the boundary and the tubular neighborhoods of $C_{\leq \gamma_a}$ and $C_{\leq \gamma_b}$ to be as small as we like.

To construct the diffeomorphism

$$([0, 1] \times S^3, F_a, S_{\leq \gamma_a}) \xrightarrow{\cong} ([0, 1] \times S^3, F_b, S_{\leq \gamma_b})$$

we first construct an ambient isotopy θ_t , $t \in [a, b]$, of $[0, 1] \times S^3$ which carries $F_a \cup C_{\leq \gamma_a}$ to $F_b \cup C_{\leq \gamma_b}$.

The map $\psi_t \circ \psi_a^{-1}$ restricts to an isotopy θ_t^F from F_a to F_b , and the isotopy $(\psi_t \circ \psi_a^{-1})|_{\gamma_a}$ induces an isotopy θ_t^γ from $C_{\leq \gamma_a}$ to $C_{\leq \gamma_b}$. (This uses conditions (G-1) and (G-2).) The isotopy θ_t^γ will not be the identity on the boundary, but we can choose it so that, for all small s , it sends $C_{\leq \gamma_a} \cap (\{s\} \times S^3)$ to $C_{\leq \gamma_b} \cap (\{s\} \times S^3)$ for each t . The maps θ_t^F and θ_t^γ typically will not agree on $F_a \cap C_{\leq \gamma_a}$, but by Condition (G-3), for any $t \in [a, b]$, $\theta_t^F(F_a)$ and the interior of $\theta_t^\gamma(C_{\leq \gamma_a})$ intersect transversally in finitely many points, say P_t . So, we get one-parameter families of points $(\theta_t^F)^{-1}(P_t)$ on F_a and $(\theta_t^\gamma)^{-1}(P_t)$ on $C_{\leq \gamma_a}$. Let $\tilde{\theta}_t^F$ be an isotopy of F_a which preserves it setwise, is the identity near the boundary and near γ_a , and satisfies $\tilde{\theta}_t^F(P_a) = (\theta_t^F)^{-1}(P_t)$. Similarly, let $\tilde{\theta}_t^\gamma$ be an isotopy of $C_{\leq \gamma_a}$ which preserves it setwise, is the identity near the boundary, and satisfies $\tilde{\theta}_t^\gamma(P_a) = (\theta_t^\gamma)^{-1}(P_t)$. On F_a , set $\theta_t = \theta_t^F \circ \tilde{\theta}_t^F$, and on $C_{\leq \gamma_a}$, set $\theta_t = \theta_t^\gamma \circ \tilde{\theta}_t^\gamma$. These two isotopies do agree on $F_a \cap C_{\leq \gamma_a}$, so let $\theta_t: F_a \cup C_{\leq \gamma_a} \rightarrow [0, 1] \times S^3$, $t \in [a, b]$, be their union. By the isotopy extension lemma, we can extend θ_t to a smooth ambient isotopy θ_t of $[0, 1] \times S^3$. (This again uses Conditions (G-2) and (G-3).) We can ensure that this extension preserves the slices $\{s\} \times S^3$ for $s \in [0, 2\epsilon] \cup [1 - 2\epsilon, 1]$, for some sufficiently small ϵ ; shrinking ϵ if necessary, we may assume that θ_t is the identity on $F_a \cap ([0, 2\epsilon] \cup [1 - 2\epsilon, 1]) \times S^3$.

The ambient isotopy θ_t is not the identity on the whole boundary, but $\theta_t|_{\{0,1\} \times S^3}$ is, of course, isotopic to the identity. So, we can modify θ_t to a new isotopy θ'_t so that:

- θ'_t is the identity on a small collar neighborhood $S^3 \times ([0, \epsilon] \cup [1 - \epsilon, 1])$ of the boundary.
- θ'_t preserves $\{s\} \times S^3$ for $s \in [0, 2\epsilon] \cup [1 - 2\epsilon, 1]$.
- θ'_t agrees with θ_t on $[2\epsilon, 1 - 2\epsilon] \times S^3$.
- θ'_t is the identity on $F_a \cap ([0, 2\epsilon] \cup [1 - 2\epsilon, 1]) \times S^3$. So, in particular, $\theta'_b(F_a) = F_b$.

Choose the collar neighborhoods of the boundary, in the definition of $S_{\leq \gamma_a}$ and $S_{\leq \gamma_b}$, to be $([0, 2\epsilon] \cup [1 - 2\epsilon, 1]) \times S^3$. Then, for appropriate choices of tubular neighborhoods of $C_{\leq \gamma_a}$ and $C_{\leq \gamma_b}$, the map θ'_b sends F_a to F_b and $S_{\leq \gamma_a}$ to $S_{\leq \gamma_b}$, as desired. \square

We can now conclude that for any interval $[t_i, t_{i+1}]$, the cuts $S_{\leq \psi_{t_i}(\gamma)}$ for $\psi_{t_i}(\Phi(F))$ and $S_{\leq \psi_{t_{i+1}}(\gamma)}$ for $\psi_{t_{i+1}}(\Phi(F))$ are equivalent. If the interval is γ -good, then this follows from Lemma 4.8. On the other hand, if the interval is η -good, then the cuts $S_{\geq \psi_{t_i}(\eta)}$ for $\psi_{t_i}(\Phi(F))$ and $S_{\geq \psi_{t_{i+1}}(\eta)}$ for $\psi_{t_{i+1}}(\Phi(F))$ are diffeomorphic, again by Lemma 4.8; however, the former is elementary equivalent to $S_{\leq \psi_{t_i}(\gamma)}$, while the latter is elementary equivalent to $S_{\leq \psi_{t_{i+1}}(\gamma)}$. Therefore, we get that the cut $S_{\leq \psi_0(\gamma)} = S_{\leq \gamma}$ for $\psi_0(\Phi(F)) = \Phi(F)$ and the cut $S_{\leq \psi_1(\gamma)}$ for $\psi_1(\Phi(F)) = \Phi'(F')$ are equivalent. So all that remains is to show that the two cuts $S_{\leq \psi_1(\gamma)}$ and $S_{\leq \alpha}$ for $\Phi'(F')$ are equivalent.

Using Lemma 4.6, choose an even-length sequence $\alpha = \delta_0, \delta_1, \dots, \delta_{2n} = \psi_1(\gamma)$ of complement-nonorientable one-sided embedded curves on $\Phi'(F')$ so that every pair of consecutive curves are disjoint. Perturbing slightly, we may assume that the genericity conditions (G-1)–(G-3) hold for each of these curves, and the genericity condition (G-7) holds for each pair of consecutive curves. Then, $S_{\leq \delta_0}, S_{\geq \delta_1}, S_{\leq \delta_2}, \dots, S_{\leq \delta_{2n}}$ is a sequence of admissible cuts connecting $S_{\leq \alpha}$ and $S_{\leq \psi_1(\gamma)}$ where every consecutive pair is elementary equivalent. \square

Remark 4.9. The proof that all admissible cuts are equivalent is inspired by the $b_2^+ \geq 3$ case of Ozsváth-Szabó's argument [OSz06, Theorem 8.5]. Their argument is terse, so for comparison with the proof of Proposition 4.2 we expand the $b_2^+ \geq 3$ case of their argument here.

Given a compact, oriented, connected 4-dimensional cobordism $W: Y_0 \rightarrow Y_1$, Y_i connected, an *admissible cut* for W is a decomposition $W = W_0 \cup_N W_1$ along a closed, connected 3-manifold N so that both $b_2^+(W_0) > 0$ and $b_2^+(W_1) > 0$ (and $Y_i \subset W_i$). Ozsváth-Szabó show that any 4-manifold W with $b_2^+(W) \geq 2$ has an admissible cut, as follows. Choose closed, connected, oriented surfaces $\Sigma_0, \Sigma_1 \subset W$ with $[\Sigma_0]^2, [\Sigma_1]^2 > 0$ and $[\Sigma_0] \cdot [\Sigma_1] = 0$. One can make Σ_0 and Σ_1 disjoint by repeatedly performing embedded surgery on Σ_0 to cancel pairs of points $p, q \in \Sigma_0 \cap \Sigma_1$ of opposite sign, along an arc in Σ_1 from p to q . Let W_0 be a neighborhood of the union of Σ_0 , an arc from Σ_0 to Y_0 , and Y_0 . Choose the arc generically and its neighborhood small enough to be disjoint from Σ_1 , and let W_1 be the complement of the interior of W_0 . Then, $W_0 \cup W_1$ is an admissible cut for W .

Next, call admissible cuts $W = W_0 \cup_N W_1 = W'_0 \cup_{N'} W'_1$ *elementary equivalent* if $N \cap N' = \emptyset$, and *equivalent* if they differ by a sequence of elementary equivalences. If $b_2^+(W) \geq 3$ then Ozsváth-Szabó argue that any two admissible cuts for W are equivalent. Fix admissible cuts $W = W_0 \cup_N W_1 = W'_0 \cup_{N'} W'_1$. For convenience, assume that $b_2^+(W'_1) \geq 2$; the other case is symmetric. Choose connected surfaces $\Sigma_0 \subset W_0$ with $[\Sigma_0]^2 > 0$ and $\Sigma'_1 \subset W'_1$ with $[\Sigma'_1]^2 > 0$, and so that $[\Sigma_0] \cdot [\Sigma'_1] = 0$ (this uses the assumption on $b_2^+(W'_1)$). (This part of the argument is

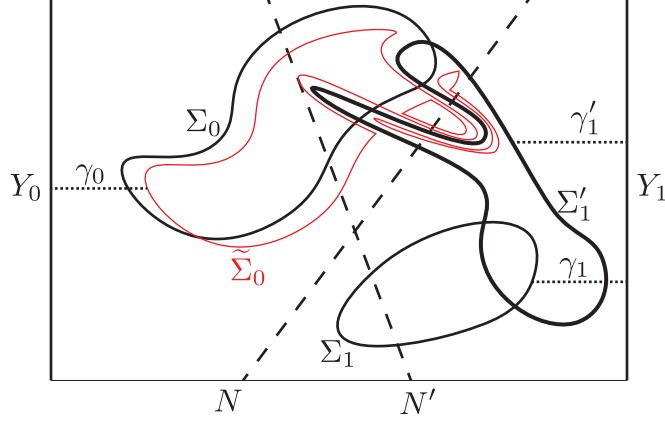


FIGURE 4.2. **Equivalence of admissible cuts in Ozsváth-Szabó's setting.** This is a schematic illustration of the surfaces and paths in Remark 4.9. The cuts N and N' are dashed, the paths γ are dotted, the surface Σ'_1 is thick, and the surface $\tilde{\Sigma}_0$ is **thin**. A key point is that $\tilde{\Sigma}_0$ is disjoint from Σ_1 and Σ'_1 .

illustrated schematically in Figure 4.2.) Performing surgery on Σ_0 along arcs Γ in Σ'_1 with interiors disjoint from Σ_0 gives a new surface $\tilde{\Sigma}_0$ homologous to Σ_0 and disjoint from Σ'_1 . Let $\Sigma_1 \subset \mathring{W}_1$ be a surface with $[\Sigma_1]^2 > 0$. Since $\Sigma_0 \subset \mathring{W}_0$ and $\Sigma_1 \subset \mathring{W}_1$, $\Sigma_0 \cap \Sigma_1 = \emptyset$. Perturbing Σ_1 slightly, we can assume that Σ_1 is also disjoint from the arcs Γ , and hence from $\tilde{\Sigma}_0$.

Choose an arc $\gamma_0 \subset W_0$ connecting $\tilde{\Sigma}_0$ to Y_0 , disjoint from Σ'_1 ; $\gamma'_1 \subset W'_1$ connecting Σ'_1 to Y_1 , disjoint from $\tilde{\Sigma}_0 \cup \gamma_0$; and $\gamma_1 \subset W_1$ connecting Σ_1 to Y_1 , disjoint from Γ . Let \tilde{N}_0 be the boundary of a neighborhood of $Y_0 \cup \gamma_0 \cup \tilde{\Sigma}_0$, N'_1 the boundary of a neighborhood of $Y_1 \cup \gamma'_1 \cup \Sigma'_1$, and N_1 the boundary of a neighborhood of $Y_1 \cup \gamma_1 \cup \Sigma_1$. Observe that \tilde{N}_0 , N_1 , and N'_1 are all admissible cuts. Further, N is equivalent to N_1 , which is equivalent to \tilde{N}_0 , which is equivalent to N'_1 , which is equivalent to N' , proving the result.

5. THE MIXED INVARIANT

Consider the long exact sequence $\cdots \rightarrow \mathcal{H}^-(L) \xrightarrow{\iota_*} \mathcal{H}^\infty(L) \xrightarrow{\pi_*} \mathcal{H}^+(L) \xrightarrow{\partial} \mathcal{H}^-(L) \rightarrow \cdots$ from Formula (2.2). Define $\mathcal{H}^{\text{red}}(L) = \ker(\iota_*) \cong \text{coker}(\pi_*)$, and give it the grading it inherits as a submodule of $\mathcal{H}^-(L)$; this is $(1, 0)$ higher than its grading as a quotient module of $\mathcal{H}^+(L)$. (This is analogous to the reduced Heegaard Floer invariant HF_{red} , and is not immediately related to reduced Khovanov homology.)

Definition 5.1. Fix an admissible cut (S, V, ϕ) for a cobordism $F \subset [0, 1] \times S^3$, decomposing F into F_0 and F_1 . Let Φ be a diffeomorphism as in Condition (AC-3) of Definition 4.1. Given a movie M_0 representing $\Phi(F_0) \subset [0, 1/2] \times S^3$ (identified with $[0, 1] \times S^3$ in the obvious way) and a movie M_1 representing $\Phi(F_1) \subset [1/2, 1] \times S^3$, we say that the concatenated movie (M_0, M_1) is a *movie compatible with the admissible cut* (S, V, ϕ) .

Lemma 5.2. *Let F be a nonorientable cobordism in $[0, 1] \times S^3$ from L_0 to L_1 . Then, the image of the induced map $\mathcal{H}^-(F): \mathcal{H}^-(L_0) \rightarrow \mathcal{H}^-(L_1)$ lies in $\mathcal{H}^{\text{red}}(L_1) \subset \mathcal{H}^-(L_1)$, and the map $\mathcal{H}^+(F): \mathcal{H}^+(L_0) \rightarrow \mathcal{H}^+(L_1)$ descends to a map $\mathcal{H}^{\text{red}}(L_0) \rightarrow \mathcal{H}^+(L_1)$.*

Proof. By Proposition 3.5, $\mathcal{H}^\infty(F): \mathcal{H}^\infty(L_0) \rightarrow \mathcal{H}^\infty(L_1)$ vanishes. So, both claims follow from the first long exact sequence in Formula (2.2) and Lemma 3.1. \square

Definition 5.3. Let F be a nonorientable cobordism from L_0 to L_1 with crosscap number ≥ 3 . Let S be an admissible cut for F , decomposing F as $F_1 \circ F_0$ and $[0, 1] \times S^3$ as $W_1 \circ W_0$. Choose a movie compatible with the admissible cut, and let $\mathcal{H}^\pm(F_i)$ be the induced maps. By Lemma 5.2, the map $\mathcal{H}^-(F_0)$ induces a map $\mathcal{H}(F_0): \mathcal{H}^-(L_0) \rightarrow \mathcal{H}^{\text{red}}(L)$ and $\mathcal{H}^+(F_1)$ descends to a map $\mathcal{H}(F_1): \mathcal{H}^{\text{red}}(L) \rightarrow \mathcal{H}^+(L_1)$. Define the mixed invariant $\Phi_F: \mathcal{H}^-(L_0) \rightarrow \mathcal{H}^+(L_1)$ to be the composition $\mathcal{H}(F_1) \circ \mathcal{H}(F_0)$. That is, Φ_F is the composition of the two dashed arrows in the following diagram:

$$(5.1) \quad \begin{array}{ccc} & \mathcal{H}^\infty(L) & \xrightarrow{\mathcal{H}^\infty(F_1) = 0} \mathcal{H}^\infty(L_1) \\ & \downarrow & \downarrow \\ & \mathcal{H}^+(L) & \xrightarrow{\mathcal{H}^+(F_1)} \mathcal{H}^+(L_1) \\ & \downarrow & \downarrow \\ & \mathcal{H}^{\text{red}}(L) & \xrightarrow{\mathcal{H}(F_1)} \mathcal{H}^+(L_1) \\ & \downarrow & \\ \mathcal{H}^-(L_0) & \xrightarrow{\mathcal{H}^-(F_0)} \mathcal{H}^-(L) & \\ \downarrow & & \downarrow \\ \mathcal{H}^\infty(L_0) & \xrightarrow{\mathcal{H}^\infty(F_0) = 0} \mathcal{H}^\infty(L) & \end{array}$$

(Note: In the original image, dashed arrows connect $\mathcal{H}^-(L_0) \rightarrow \mathcal{H}^{\text{red}}(L)$ labeled $\mathcal{H}(F_0)$ and $\mathcal{H}^{\text{red}}(L) \rightarrow \mathcal{H}^+(L_1)$ labeled $\mathcal{H}(F_1)$.)

Remark 5.4. Definition 5.3 also makes sense if F has crosscap number 2, but the proof that Φ_F is independent of the choice of admissible cut (Theorem 5.5) requires crosscap number at least 3. That is, in the case $F: L_0 \rightarrow L_1$ has crosscap number 2, there is a map $\Phi_{F,S}: \mathcal{H}^-(L_0) \rightarrow \mathcal{H}^+(L_1)$ which, as far as we know, depends on both the surface F and the admissible cut S .

Theorem 5.5. *Let F be a cobordism from L_0 to L_1 , with crosscap number ≥ 3 . Then, the mixed invariant $\Phi_F: \mathcal{H}^-(L_0) \rightarrow \mathcal{H}^+(L_1)$ is independent of the choices in its construction, up to an overall sign. Further, If F is isotopic to F' or, more generally, if there is a self-diffeomorphism of $[0, 1] \times S^3$ which is the identity on the boundary and sends F to F' then $\Phi_F = \pm \Phi_{F'}: \mathcal{H}^-(L_0) \rightarrow \mathcal{H}^+(L_1)$.*

Proof. We will show (in order):

- (1) Independence of the choice of movies compatible with a fixed admissible cut and, in particular, of isotopies of F_0 and F_1 rel boundary, and of the choice of diffeomorphism Φ in Definition 5.1.
- (2) Invariance under elementary equivalences of admissible cuts.
- (3) Invariance under diffeomorphisms of surfaces and admissible cuts.

By Proposition 4.2, this implies the result.

Throughout the proof, “equal” or “homotopic” will mean equal or homotopic up to an overall sign.

For point (1), by Proposition 3.7, difference choices of movies compatible with the same admissible cut give chain homotopic maps $\mathcal{C}^-(F_0)$ and $\mathcal{C}^-(F_1)$. (For independence of Φ , this uses the last statement in Proposition 3.7.) If $\mathcal{C}^-(F_0) \sim \mathcal{C}^-(F'_0)$, then $\mathcal{H}^-(F_0) = \mathcal{H}^-(F'_0)$; similarly, if $\mathcal{C}^-(F_1) \sim \mathcal{C}^-(F'_1)$, then $\mathcal{C}^+(F_1) \sim \mathcal{C}^+(F'_1)$, and hence $\mathcal{H}^+(F_1) = \mathcal{H}^+(F'_1)$. Notice that the lift $\mathcal{H}(F_0): \mathcal{H}^-(L_0) \rightarrow \mathcal{H}^{\text{red}}(L)$ of $\mathcal{H}^-(F_0)$ is canonical: $\mathcal{H}^{\text{red}}(L)$ is a canonical submodule of $\mathcal{H}^-(L)$. Similarly, the induced map $\mathcal{H}(F_1): \mathcal{H}^{\text{red}}(L) \rightarrow \mathcal{H}^+(L_1)$ is canonical, since $\mathcal{H}^{\text{red}}(L)$ is a canonical quotient module of $\mathcal{H}^+(L)$. So, different choices of movies for F_0 and F_1 give the same mixed invariant.

For point (2), if the admissible cuts (S, V, ϕ) and (S', V', ϕ') are elementary equivalent, let $L = \phi(S \cap F)$ and $L' = \phi'(S' \cap F)$. Assume, without loss of generality, that we are in the first case of Formula (4.1). Choose a movie for F compatible with this decomposition (in a sense analogous to Definition 5.1), so L and L' are frames in the movie. Then, it follows from commutativity of the diagram

$$\begin{array}{ccc} \mathcal{H}^+(L; R) & \longrightarrow & \mathcal{H}^+(L'; R) \\ \downarrow & & \downarrow \\ \mathcal{H}^{\text{red}}(L; R) & \longrightarrow & \mathcal{H}^{\text{red}}(L'; R) \\ \downarrow & & \downarrow \\ \mathcal{H}^-(L; R) & \longrightarrow & \mathcal{H}^-(L'; R) \end{array}$$

(and point (1)) that the mixed invariants with respect to (S, V, ϕ) and (S', V', ϕ') agree.

Finally, for point (3), suppose an admissible cut (S, V, ϕ) for F is diffeomorphic to an admissible cut (S', V', ϕ') for F' , via a diffeomorphism Ψ . Fix a movie for F compatible with (S, V, ϕ) , with respect to a diffeomorphism Φ extending ϕ . Then, the same movie is compatible with (S', V', ϕ') , via the diffeomorphism $\Phi \circ \Psi$. By points (1) and (2), we can compute the mixed invariants of F and F' using these movies, so the mixed invariants agree. \square

6. PROPERTIES

6.1. First observations. We start by noting the mixed invariant's grading:

Lemma 6.1. *The mixed invariant $\Phi_F: \mathcal{H}^-(L_0) \rightarrow \mathcal{H}^+(L_1)$ increases the bigrading by $(-1 - e/2, \chi - 3e/2 - 2s)$, where χ is the Euler characteristic of F , e is the normal Euler number of F , and s is the number of stars on F .*

Proof. This follows immediately from Corollary 3.3. The additional downward grading shift by $(1, 0)$ comes from the identification of $\mathcal{H}^{\text{red}}(L)$ as a quotient module of $\mathcal{H}^+(L)$. \square

There is a simple condition guaranteeing the mixed invariant vanishes:

Lemma 6.2. *If F has an admissible cut S so that the link $L \subset S^3$ corresponding to $S \cap F$ has $\mathcal{H}^{\text{red}}(L) = 0$, then the mixed invariant Φ_F vanishes.*

Proof. This is immediate from Definition 5.3. \square

Remark 6.3. The analogue of Lemma 6.2 for Heegaard Floer homology is factoring through an L -space. Note, however, that L -spaces seem to be much more common than links with vanishing \mathcal{H}^{red} .

The mixed invariant behaves simply with respect to (a particular kind of) mirroring. Let $F: L_0 \rightarrow L_1$ be a cobordism, so F is smoothly embedded in $[0, 1] \times S^3$. Applying the orientation-preserving diffeomorphism $[0, 1] \times S^3 \rightarrow [0, 1] \times S^3$, $(t, x, y, z, w) \mapsto (1 - t, -x, y, z, w)$ gives a new cobordism $m(F): m(L_1) \rightarrow m(L_0)$, where $m(L_i)$ denotes the mirror of L_i .

The statement is a little simpler for the Lee deformation than the Bar-Natan deformation, so we separate the two cases. Given an $R[T]$ -module M , $\text{Hom}_R(M, R)$ inherits the structure of an $R[T]$ -module, as well.

Lemma 6.4. *Let \mathcal{C}^\pm denote the Lee deformation. Given a link L , there is an isomorphism $\mathcal{C}^+(m(L)) \cong \text{Hom}_R(\mathcal{C}^-(L), R)$ of complexes over $R[T]$ so that for any cobordism $F: L_0 \rightarrow L_1$,*

$$\mathcal{C}^+(m(F)): \mathcal{C}^+(m(L_1)) \cong \text{Hom}_R(\mathcal{C}^-(L_1), R) \rightarrow \text{Hom}_R(\mathcal{C}^-(L_0), R) \cong \mathcal{C}^+(m(L_0))$$

is the dual to the map $\mathcal{C}^-(F): \mathcal{C}^-(L_0) \rightarrow \mathcal{C}^-(L_1)$.

Further, if F has crosscap number ≥ 3 and R is a field then the mixed invariant $\Phi_{m(F)}$ is given by the composition

$$\mathcal{H}^-(m(L_1)) \cong \text{Hom}(\mathcal{H}^+(L_1), R) \xrightarrow{\Phi_F^*} \text{Hom}(\mathcal{H}^-(L_0), R) \cong \mathcal{H}^+(m(L_0)).$$

In particular, $\Phi_{m(F)}$ vanishes if and only if Φ_F does.

(Compare [Kho00, Proposition 32], [Kho06b, pp. 184–185], and [CMW09, Proposition 3.1].)

Proof. The isomorphism $\mu: \text{Hom}(\mathcal{C}^-(L), R) \rightarrow \mathcal{C}^+(m(L))$ is defined as follows. Given a generator $T^n(v, y)$ of $\mathcal{C}^-(L)$ (over R), let $[T^n(v, y)]^*$ denote the dual generator of $\text{Hom}(\mathcal{C}^-(L), R)$. The isomorphism is given by $\mu([T^n(v, y)]^*) = T^{-n-1}(\vec{1} - v, y^*)$ where $\vec{1} - v = (1 - v_1, \dots, 1 - v_c)$ and y^* is the result of reversing the label of every circle. (That is, if y labels a circle Z by X then y^* labels the corresponding circle by 1, and vice-versa.) It is straightforward to check that this defines a chain isomorphism.

A movie for F induces a movie for $m(F)$ by mirroring each frame and reversing the order of the frames. So, it suffices to check the second statement for a single elementary cobordism (pair of adjacent frames in a movie). This is a straightforward case check.

For the statement about the mixed invariant, a choice of admissible cut for F , along some link L , and movie compatible with it induce an admissible cut for $m(F)$, along $m(L)$, and a movie compatible with it. The map μ induces an isomorphism of short exact sequences

$$\begin{array}{ccccccc} 0 & \rightarrow & \text{Hom}(\mathcal{C}^+(L), R) & \rightarrow & \text{Hom}(\mathcal{C}^\infty(L), R) & \rightarrow & \text{Hom}(\mathcal{C}^-(L), R) \rightarrow 0 \\ & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \mathcal{C}^-(m(L)) & \longrightarrow & \mathcal{C}^\infty(m(L)) & \longrightarrow & \mathcal{C}^+(m(L)) \longrightarrow 0 \end{array}$$

so naturality of the snake lemma implies that the diagram

$$\begin{array}{ccccc}
 \mathrm{Hom}(\mathcal{H}^-(L), \mathbb{R}) & \xrightarrow{\partial^*} & \mathrm{Hom}(\mathcal{H}^+(L), \mathbb{R}) & & \\
 \downarrow \cong & \searrow & \nearrow & & \downarrow \cong \\
 & \mathrm{Hom}(\mathcal{H}^{\mathrm{red}}(L), \mathbb{R}) & & & \\
 & \downarrow & & & \\
 & \mathcal{H}^{\mathrm{red}}(m(L)) & & & \\
 \nearrow & & \searrow & & \\
 \mathcal{H}^+(m(L)) & \xrightarrow{\partial} & \mathcal{H}^-(m(L)) & &
 \end{array}$$

commutes. Combining this with the definition of the mixed invariant in Diagram (5.1) gives the result. \square

For the analogous results for the Bar-Natan complex, there is an extra sign. There is a ring automorphism $\sigma: \mathbb{R}[H] \rightarrow \mathbb{R}[H]$ induced by $\sigma(H) = -H$. Given a module M over $\mathbb{R}[H]$, let M_σ be the result of restricting (or extending) scalars by σ . That is, M_σ is the same as M except H acts on M_σ the way $-H$ acts on M . Given another $\mathbb{R}[H]$ -module N and a homomorphism $f: M \rightarrow N$, there is an induced homomorphism $f: M_\sigma \rightarrow N_\sigma$ (which, as a map of sets, is the same as $f: M \rightarrow N$).

Then, the following is the analogue of Lemma 6.4:

Lemma 6.5. *Let \mathcal{C}^\pm denote the Bar-Natan deformation. Given a link L , there is an isomorphism $\mathcal{C}^+(m(L))_\sigma \cong \mathrm{Hom}_{\mathbb{R}}(\mathcal{C}^-(L), \mathbb{R})$ of complexes over $\mathbb{R}[H]$ so that for any cobordism $F: L_0 \rightarrow L_1$,*

$$\mathcal{C}^+(m(F)): \mathcal{C}^+(m(L_1))_\sigma \cong \mathrm{Hom}_{\mathbb{R}}(\mathcal{C}^-(L_1), \mathbb{R}) \rightarrow \mathrm{Hom}_{\mathbb{R}}(\mathcal{C}^-(L_0), \mathbb{R}) \cong \mathcal{C}^+(m(L_0))_\sigma$$

is the dual to the map $\mathcal{C}^-(F): \mathcal{C}^-(L_0) \rightarrow \mathcal{C}^-(L_1)$.

Further, if F has crosscap number ≥ 3 and \mathbb{R} is a field then the mixed invariant $\Phi_{m(F)}$, viewed as a map of modules twisted by σ , is given by the composition

$$\mathcal{H}^-(m(L_1))_\sigma \cong \mathrm{Hom}(\mathcal{H}^+(L_1), \mathbb{R}) \xrightarrow{\Phi_F^*} \mathrm{Hom}(\mathcal{H}^-(L_0), \mathbb{R}) \cong \mathcal{H}^+(m(L_0))_\sigma.$$

In particular, $\Phi_{m(F)}$ vanishes if and only if Φ_F does.

Proof. The isomorphism sends $[H^n(v, y)]^*$ to $(-1)^n H^{-n-1}(\vec{1} - v, y^*)$. The rest of the proof is a straightforward adaptation of the Lee case. \square

Remark 6.6. If \mathbb{R} is a field, it follows from the classification of modules over a PID that there is a (perhaps unnatural) isomorphism over $\mathbb{R}[H]$ between $\mathcal{H}^+(m(L))$ and $\mathrm{Hom}(\mathcal{H}^+(L), \mathbb{R})$.

Remark 6.7. There is another mirror one might consider: the map $(t, x, y, z, w) \mapsto (t, -x, y, z, w)$ which mirrors each frame in the movie but does not reverse the order of the frames. Neither $\mathcal{H}^-(F)$ nor the mixed invariant seems to behave simply with respect to this operation, as the example in Section 7.1 shows. (Gauge-theoretic invariants of the branched double cover also do not behave well with respect to this operation.)

The mixed invariant also respects composition, as follows:

Lemma 6.8. *Let $F_0: L_0 \rightarrow L_1$, $F_1: L_1 \rightarrow L_2$, and $F_2: L_2 \rightarrow L_3$ be cobordisms, so that F_1 has crosscap number ≥ 3 . Then,*

$$(6.1) \quad \Phi_{F_2 \circ F_1} = \mathcal{H}^+(F_2) \circ \Phi_{F_1}$$

$$(6.2) \quad \Phi_{F_1 \circ F_0} = \Phi_{F_1} \circ \mathcal{H}^-(F_0).$$

The same result holds if F_1 has crosscap number 2, as long as either $F_1 \circ F_0$ and $F_2 \circ F_1$ have crosscap number ≥ 3 , and we define Φ_{F_1} with respect to any choice of admissible cut for F_1 (compare Remark 5.4).

Proof. This is immediate from the definitions. \square

There is an easy criterion for the mixed invariant to be non-vanishing, in terms of the induced map on ordinary Khovanov homology $\widehat{\mathcal{H}}$:

Lemma 6.9. *Let $F: L_0 \rightarrow L_1$ be a cobordism with crosscap number ≥ 3 . Then, the following diagrams commute:*

$$\begin{array}{ccc} \mathcal{H}^-(L_0) & \xrightarrow{\Phi_F} & \mathcal{H}^+(L_1) \\ \pi_* \downarrow & & \downarrow \partial \\ \widehat{\mathcal{H}}(L_0) & \xrightarrow{\widehat{\mathcal{H}}(F)} & \widehat{\mathcal{H}}(L_1) \end{array} \quad \text{and} \quad \begin{array}{ccc} \widehat{\mathcal{H}}(L_0) & \xrightarrow{\widehat{\mathcal{H}}(F)} & \widehat{\mathcal{H}}(L_1) \\ \partial \downarrow & & \downarrow \iota_* \\ \mathcal{H}^-(L_0) & \xrightarrow{\Phi_F} & \mathcal{H}^+(L_1). \end{array}$$

In particular, if $\widehat{\mathcal{H}}(F) \circ \pi_$ or $\iota_* \circ \widehat{\mathcal{H}}(F)$ is non-zero then the mixed invariant Φ_F is also non-zero.*

Proof. Let L be an admissible cut for F , decomposing F as $F_1 \circ F_0$. Define the map $\partial: \mathcal{H}^{\text{red}}(L) \rightarrow \widehat{\mathcal{H}}(L)$ to be the composition $\mathcal{H}^{\text{red}}(L) \rightarrow \mathcal{H}^-(L) \xrightarrow{\pi_*} \widehat{\mathcal{H}}(L)$. Using the first commutative triangle in Formula (2.3), the map $\partial: \mathcal{H}^+(L) \rightarrow \widehat{\mathcal{H}}(L)$ is the composition $\mathcal{H}^+(L) \rightarrow \mathcal{H}^{\text{red}}(L) \rightarrow \mathcal{H}^-(L) \xrightarrow{\pi_*} \widehat{\mathcal{H}}(L)$, so is also the composition $\mathcal{H}^+(L) \rightarrow \mathcal{H}^{\text{red}}(L) \xrightarrow{\partial} \widehat{\mathcal{H}}(L)$.

To see that $\partial \circ \Phi_F = \widehat{\mathcal{H}}(F) \circ \pi_*$, consider the larger diagram

$$\begin{array}{ccccc} & & \mathcal{H}^{\text{red}}(L) & & \\ & \swarrow \mathcal{H}(F_0) & & \searrow \mathcal{H}(F_1) & \\ \mathcal{H}^-(L_0) & & & & \mathcal{H}^+(L_1) \\ & \searrow \mathcal{H}^-(F_0) & \downarrow \partial & \swarrow \mathcal{H}^+(F_1) & \\ & \mathcal{H}^-(L) & & \mathcal{H}^+(L) & \\ \pi_* \swarrow & & \downarrow \pi_* & & \searrow \partial \\ \widehat{\mathcal{H}}(L_0) & \xrightarrow{\widehat{\mathcal{H}}(F_0)} & \widehat{\mathcal{H}}(L) & \xrightarrow{\widehat{\mathcal{H}}(F_1)} & \widehat{\mathcal{H}}(L_1) \end{array}$$

The middle triangles commute by the discussion above. The outer squares commute by naturality of the long exact sequences (2.2), Lemma 3.1. The triangles at the top commute by the definition of the dashed lifts. Since the map $\mathcal{H}^+(L) \rightarrow \mathcal{H}^{\text{red}}(L)$ is surjective, commutativity of the right square and triangles implies that $\partial \circ \mathcal{H}(F_1) = \widehat{\mathcal{H}}(F_1) \circ \partial: \mathcal{H}^{\text{red}}(L) \rightarrow \widehat{\mathcal{H}}(L_1)$. Commutativity of the square and two triangles on the left then implies the result.

The proof that $\iota_* \circ \widehat{\mathcal{H}}(F) = \Phi_F \circ \partial$ is similar, but instead uses the commutative diagram

$$\begin{array}{ccccc}
 \widehat{\mathcal{H}}(L_0) & \xrightarrow{\widehat{\mathcal{H}}(F_0)} & \widehat{\mathcal{H}}(L) & \xrightarrow{\widehat{\mathcal{H}}(F_1)} & \widehat{\mathcal{H}}(L_1) \\
 \downarrow \partial & & \downarrow \partial & & \downarrow \iota_* \\
 \mathcal{H}^-(L_0) & \xrightarrow{\mathcal{H}^-(F_0)} & \mathcal{H}^-(L) & & \mathcal{H}^+(L_1) \\
 & \searrow \mathcal{H}^-(F_0) & \downarrow & \swarrow \mathcal{H}^+(F_1) & \\
 & & \mathcal{H}^+(L) & & \\
 & \swarrow \mathcal{H}^-(F_0) & \downarrow & \searrow \mathcal{H}^+(F_1) & \\
 & & \mathcal{H}^{\text{red}}(L) & & \\
 & \swarrow \mathcal{H}^-(F_0) & & \searrow \mathcal{H}^+(F_1) & \\
 & & & &
 \end{array}$$

and the fact that the map $\mathcal{H}^{\text{red}}(L) \rightarrow \mathcal{H}^-(L)$ is injective. \square

Remark 6.10. Since $\pi_*: \mathcal{H}^-(\emptyset) \rightarrow \widehat{\mathcal{H}}(\emptyset)$ is surjective, it follows from Lemma 6.9 that if F is a cobordism from \emptyset to L and $\widehat{\mathcal{H}}(F) \neq 0$ then $\Phi_F \neq 0$, as well. Similarly, if F is a cobordism from L to \emptyset and $\widehat{\mathcal{H}}(F) \neq 0$ then $\Phi_F \neq 0$.

6.2. Stabilizations. Next, we turn to the behavior of the mixed invariant under various local changes to the knot. For example, we will study the behavior under Baykur-Sunukjian's *stabilizations* (attaching arbitrary 1-handles to a surface) and *standard stabilizations* (local connect sums with a standard T^2) [BS16], *crosscap stabilizations* (taking local connected sums with a standard \mathbb{RP}^2 or $\overline{\mathbb{RP}}^2$), *local knotting* (taking a local connected sum with a knotted S^2), and more general local connected sums. The main results are summarized in Theorem 6.18, though some technical results along the way (e.g., Corollaries 6.13 and 6.15) may also be of interest.

Most of the results in this section work for all four versions of Khovanov homology, \mathcal{H}^- , \mathcal{H}^+ , \mathcal{H}^∞ , or $\widehat{\mathcal{H}}$, so we will use the symbol \mathcal{H}^\bullet to denote any of these four versions. The key technical property we will use, as usual for these kinds of arguments, is a neck-cutting relation.

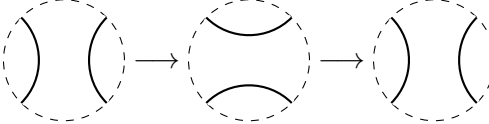
Proposition 6.11. *Let $F: L_0 \rightarrow L_1$ be a cobordism, and let A be an arc in $[0, 1] \times S^3$ with endpoints on a single component F_0 of F and interior disjoint from F . Let F^\cap be the result of attaching a 1-handle to F along A and F^* the result of adding a star to F at one endpoint of A . If F_0 is orientable, assume that the 1-handle is attached in such a way that the resulting component is still orientable. Then, $\mathcal{H}^\bullet(F^\cap) = \mathcal{H}^\bullet(F^*)$.*

Proof. For the Lee deformation, this was essentially shown by Levine-Zemke [LZ19, Proposition 7], so we focus on the case of the Bar-Natan deformation and comment on the Lee deformation at the end.

Let B be an arc in F connecting the endpoints of A , chosen so that the loop $A \cup B$ is two-sided, i.e., so that $TF^\cap|_{A \cup B}$ is trivial. (If F_0 is orientable, any arc B has $TF^\cap|_{A \cup B}$ trivial by the assumption that orientability was preserved; if F_0 is nonorientable, given any arc B connecting the endpoints of A we can modify B by taking the connected sum with a one-sided loop to achieve this property.) Let $H = F^\cap \setminus F$ denote the new 1-handle. Perform an isotopy to F^\cap so that:

- The projection $[0, 1] \times S^3 \rightarrow [0, 1]$ restricts to a Morse function f on F^\cap , and induces a movie decomposition of F^\cap .

- The restriction $f|_H$ has two critical points, both of index 1, corresponding to a pair of saddles in the movie.
- The two saddles corresponding to $f|_H$ in the movie are adjacent, happening at times $t+\epsilon, t+2\epsilon \in (0, 1)$, and there are no other elementary cobordisms between $t-\epsilon$ and $t+3\epsilon$. Further, these frames are obtained by gluing the following local model to the identity cobordism of the rest of the link:

(6.3) 

- The arc B satisfies $f(B) = t$.

(To arrange this, first isotope F^\cap so that H is small, then make H standard with respect to the projection to $[0, 1]$, and then isotope B to lie in the desired level set and use the isotopy extension lemma to push the rest of F^\cap out of the way.)

Given a dotted (not starred) cobordism, decomposed as a movie, there is a corresponding map of Bar-Natan complexes, where the map associated to a dot is multiplication by X . Let $F^{\bullet,0}$ and $F^{\bullet,1}$ be the result of placing a dot on F at each endpoint of A . Then, an easy local computation shows that

$$(6.4) \quad \pm \mathcal{H}^\bullet(F^\cap) = \mathcal{H}^\bullet(F^{\bullet,0}) + \mathcal{H}^\bullet(F^{\bullet,1}) - H\mathcal{H}^\bullet(F).$$

The surface $F^{\bullet,0}$ can be transformed into $F^{\bullet,1}$ by moving the dot along the arc B . Since B is contained in a single level, this corresponds to moving the dot along an arc in a single link diagram in the movie.

Let $F^{\bullet,(i)}$ be the surface after we have moved the dot through i crossings. By Alishahi's lemma [Ali19, Lemma 2.2],

$$\mathcal{H}^\bullet(F^{\bullet,(i)}) = H\mathcal{H}^\bullet(F) - \mathcal{H}^\bullet(F^{\bullet,(i+1)})$$

So, it suffices to show that the arc B has an even number of crossings on it: then $\mathcal{H}^\bullet(F^{\bullet,0}) = \mathcal{H}^\bullet(F^{\bullet,1})$ so the right side of Formula (6.4) is equal to $\mathcal{H}^\bullet(F^\cap)$.

This is where the assumption that $A \cup B$ is two-sided is used. Orient the arc B . This induces an orientation of the arcs in the first frame of (6.3). The fact that $TF^\cap|_{A \cup B}$ is trivial implies that this orientation is compatible with the saddle cobordisms in (6.3); that is, B connects the bottom-left endpoint to the bottom-right one, or the top-left endpoint to the top-right one. Without loss of generality, assume B runs from the top-left endpoint to the top-right one. Choose a checkerboard coloring of the left link diagram so that the region between the arcs shown is black. Then, B starts with a black region to its right, and ends with a black region to its right. Each time B passes through a crossing, the black region switches between the left and right side of B . So, B passes through an even number of crossings, as claimed.

The proof for the Lee case is the same, except that the analogue of Equation (6.4) does not have the $H\mathcal{H}(F)$ term, and Alishahi's lemma is replaced by Hedden-Ni's [HN13, Lemma 2.3]. \square

The standard \mathbb{RP}^2 inside the 4-ball is the surface represented by the following movie:

$$(6.5) \quad \emptyset \xrightarrow{b} \bigcirc \xrightarrow{RI} \bigcirc \xrightarrow{s} \bigcirc \xrightarrow{RI} \bigcirc \xrightarrow{d} \emptyset$$

The standard \mathbb{RP}^2 has $e = -2$. The standard $\overline{\mathbb{RP}}^2$ is the mirror of the above, and has $e = 2$. (Our conventions are chosen to agree with [FKV88].) Define a *crosscap stabilization* to be the result of taking the connected sum with a standard \mathbb{RP}^2 or $\overline{\mathbb{RP}}^2$.

Lemma 6.12. *If F^\otimes is obtained from F by a crosscap stabilization then $\mathcal{H}^\bullet(F^\otimes)$ vanishes.*

Proof. A movie for F^\otimes is obtained from a movie for F by taking the disjoint union with the movie in Formula (6.5) or its mirror, but replacing d with a saddle map between the unknot shown and F . It is straightforward to see that the map on \mathcal{H}^\bullet induced by $s \circ RI \circ b$ vanishes (as does the map associated to the mirror of this movie), so the map associated to the whole movie vanishes, as well. \square

Corollary 6.13. *Let $F: L_0 \rightarrow L_1$ be a cobordism, and suppose that some nonorientable component F_0 of F has a star on it. Then, $\mathcal{H}^\bullet(F) = 0$. Similarly, if F is obtained from another surface by attaching a 1-handle to a nonorientable component then $\mathcal{H}^\bullet(F) = 0$.*

Proof. For the first statement, let F^\wedge be the result of taking the connected sum of F with a local Klein bottle (with normal Euler number 0) at the star on F_0 , and forgetting the star. That is, F^\wedge is obtained from F by attaching a local 1-handle with both feet near the star, in a locally-nonorientable way (and forgetting the star). By Proposition 6.11, $\mathcal{H}^\bullet(F) = \mathcal{H}^\bullet(F^\wedge)$. On the other hand, F^\wedge is also obtained from F by taking the connect sum with $\mathbb{RP}^2 \# \mathbb{RP}^2$ (and forgetting the star). So, by Lemma 6.12, $\mathcal{H}^\bullet(F^\wedge)$ vanishes.

The second statement follows from the first and Proposition 6.11. \square

Adding an even number of stars has a predictable effect on the cobordism maps and the mixed invariant:

Lemma 6.14. *Let $F: L_0 \rightarrow L_1$ be a cobordism, and let F^{**} be the result of adding two stars to the same component of F . Then,*

$$\mathcal{H}^\bullet(F^{**}) = \begin{cases} 4T\mathcal{H}^\bullet(F) & \text{for the Lee deformation} \\ H^2\mathcal{H}^\bullet(F) & \text{for the Bar-Natan deformation.} \end{cases}$$

Further, if the crosscap number of F is at least 3 then

$$\Phi_{F^{**}} = \begin{cases} 4T\Phi_F & \text{for the Lee deformation} \\ H^2\Phi_F & \text{for the Bar-Natan deformation.} \end{cases}$$

Proof. Since the maps are invariant under isotopy of the cobordisms, we can arrange that the two elementary star cobordisms are adjacent. Then, the result is immediate from the definition of the map associated to an elementary star cobordism. \square

Corollary 6.15. *If F is nonorientable then for the Lee deformation, $4T\mathcal{H}^\bullet(F) = 0$, and for the Bar-Natan deformation, $H^2\mathcal{H}^\bullet(F) = 0$. If F has crosscap number at least 3 then additionally, $4T\Phi_F = 0$ and $H^2\Phi_F = 0$ for the Lee and Bar-Natan deformations, respectively.*

Proof. This is immediate from Corollary 6.13 and Lemma 6.14. \square

We note a very mild extension of a result of Rasmussen [Ras] and Tanaka [Tan06]:

Lemma 6.16. *If $F \subset [0, 1] \times S^3$ is a closed, connected, orientable surface of genus g with s stars then $\mathcal{H}^\bullet(F)$ is*

- 0 if $g + s$ is even,
- $2H^{g+s-1}$ if $g + s$ is odd and we are considering the Bar-Natan deformation, and
- $2^{g+s}T^{\frac{g+s-1}{2}}$ if $g + s$ is odd and we are considering the Lee deformation.

If $F \subset [0, 1] \times S^3$ is a closed, connected, nonorientable surface (possibly with stars) then $\mathcal{H}^\bullet(F) = 0$.

(In both cases, the surface may be knotted.)

Proof. We start with the orientable case. Any such surface becomes isotopic to a standardly-embedded one after attaching some number of 1-handles (see, e.g., [BS16, Theorem 1]). By adding an extra one if necessary, we may assume the number of 1-handles added is even, say $2k$. By Proposition 6.11, adding these 1-handles has the same effect as adding the same number of stars which, by Lemma 6.14, multiplies $\mathcal{H}^\bullet(F)$ by $(4T)^k$ or H^{2k} . Considering first $\mathcal{H}^-(F)$, multiplication by $(4T)^k$ or H^{2k} is an injective map $R[H] \rightarrow R[H]$ or $R[T] \rightarrow R[T]$, so the element $\mathcal{H}^-(F)$ depends only on g and s , not the embedding of F . Thus, the result for $\mathcal{H}^-(F)$ follows from an easy model computation for a standardly-embedded surface (which can be made even easier by applying Proposition 6.11 to trade the genus for stars and then applying Lemma 6.14). The results for the other versions— $\mathcal{H}^\infty(F)$, $\mathcal{H}^+(F)$, and $\widehat{\mathcal{H}}(F)$ —follow formally from the case of $\mathcal{H}^-(F)$ and the natural long exact sequences (2.2).

The proof for the nonorientable case is essentially the same. By a result of Baykur-Sunukjian [BS16, Theorem 6], after a finite number of stabilizations, F becomes isotopic to a connected sum of copies of the standard \mathbb{RP}^2 and $\overline{\mathbb{RP}}^2$. A straightforward model computation, or Lemma 6.12, shows the map associated to a connected sum of copies of the standard \mathbb{RP}^2 or $\overline{\mathbb{RP}}^2$ vanishes, so by Proposition 6.11 the map associated to F vanishes, as well. \square

Given a link cobordism $F: L_0 \rightarrow L_1$, a closed surface $E \subset S^4$, and points $p \in E$ and $q \in F$, there is a *standard connected sum* of F and E at q and p ; this is the connected sum of pairs $([0, 1] \times S^3, F) \# (S^4, E)$.

Proposition 6.17. *Let $F: L_0 \rightarrow L_1$ be a cobordism and $E \subset S^4$ a closed, connected surface with no stars on it. Let $F^\# := F \# E$ be a standard connected sum of F and E , and let F^\star be the result of adding a star to F , on the component where the connect sum is occurring. Then,*

(1) *If E is an orientable surface of genus $g > 0$ then for the Lee deformation,*

$$(6.6) \quad \mathcal{H}^\bullet(F^\#) = \begin{cases} (4T)^{\frac{g}{2}} \mathcal{H}^\bullet(F) & g \text{ even} \\ (4T)^{\frac{g-1}{2}} \mathcal{H}^\bullet(F^\star) & g \text{ odd} \end{cases}$$

while for the Bar-Natan deformation

$$(6.7) \quad \mathcal{H}^\bullet(F^\#) = \begin{cases} H^g \mathcal{H}^\bullet(F) & g \text{ even} \\ H^{g-1} \mathcal{H}^\bullet(F^\star) & g \text{ odd.} \end{cases}$$

(2) If E is a nonorientable surface then $\mathcal{H}^\bullet(F^\#) = 0$.

Proof. Let D be a small disk on E , so $E \setminus D$ is a cobordism from the empty set to the unknot U . We will show that $\mathcal{H}^\bullet(E \setminus D)$ is the same as the invariant of a disk with g stars in the orientable case, and vanishes in the nonorientable case. The result then follows from Lemma 6.14 and functoriality, since $F^\#$ is obtained from F by replacing a small disk by $E \setminus D$.

Consider first the version \mathcal{H}^- for the Lee deformation, for the case that E is orientable. Let E^\star be the result of adding a star to E . Write

$$\mathcal{H}^-(E \setminus D) = p(T)1 + q(T)X \in \mathcal{H}^-(U) = \mathbb{R}[T]\langle 1, X \rangle.$$

By Lemma 6.16 applied to E , $q(T) = 0$ if g is even and $q(T) = 2^g T^{\frac{g-1}{2}}$ if g is odd. Also,

$$\mathcal{H}^-(E^\star \setminus D) = 2p(T)X + 2q(T),$$

so by Lemma 6.16 applied to E^\star , $p(T) = 0$ if g is odd and $p(T) = 2^g T^{\frac{g}{2}}$ if g is even. So, $\mathcal{H}^-(E \setminus D)$ is $2^g T^{\frac{g}{2}}$ times the invariant of a disk if g is even, and $2^{g-1} T^{\frac{g-1}{2}}$ times the invariant of a disk with a star if g is odd. The results for the other versions— \mathcal{H}^∞ , \mathcal{H}^+ , and $\widehat{\mathcal{H}}$ —follow formally from the case of \mathcal{H}^- since $\mathcal{H}^-(U)$ is free over $\mathcal{H}^-(\emptyset)$.

The proof for the Bar-Natan deformation in the orientable case is similar.

Now, suppose E is nonorientable and again, for definiteness, consider the Lee deformation. We can again write $\mathcal{H}^-(E \setminus D) = p(T)1 + q(T)X$, but now $\mathcal{H}^-(E) = \mathcal{H}^-(E^\star) = 0$. Consequently, $p(T) = q(T) = 0$. \square

To summarize, both the map $\mathcal{H}^\bullet(F)$ and the mixed invariant Φ_F obstruct surfaces being stabilizations and crosscap stabilizations, and are independent of local knotting.

Theorem 6.18. *Let $F: L_0 \rightarrow L_1$ be a cobordism.*

- (1) *If F has at least one star on some nonorientable component, then $\mathcal{H}^\bullet(F) = 0$; and if in addition F has crosscap number ≥ 3 then $\Phi_F = 0$, as well.*
- (2) *If F is a (possibly nonstandard) stabilization, obtained from another cobordism F' by attaching a handle to some nonorientable component of F' , then $\mathcal{H}^\bullet(F) = 0$.*
- (3) *If F is obtained from another cobordism F' by taking a standard connected sum with a closed, nonorientable surface then $\mathcal{H}^\bullet(F) = 0$; if F' has crosscap number ≥ 2 then $\Phi_F = 0$, as well. In particular, this applies if F is a crosscap stabilization of a surface (with crosscap number ≥ 2 in the case of $\Phi_F = 0$).*
- (4) *If F is obtained from another cobordism F' by taking a standard connected sum of some nonorientable component of F' with a closed, orientable surface of genus $g > 0$, then $\mathcal{H}^\bullet(F) = 0$; if F' has crosscap number ≥ 2 then $\Phi_F = 0$, as well. In particular, this applies if F is a standard stabilization of a surface (with crosscap number ≥ 2 in the case of Φ_F) at some nonorientable component.*

- (5) If F is obtained from another cobordism F' by taking a standard connected sum with a knotted 2-sphere then $\mathcal{H}^\bullet(F) = \mathcal{H}^\bullet(F')$; if in addition F has crosscap number ≥ 3 then $\Phi_F = \Phi_{F'}$.

Proof. For $\mathcal{H}^\bullet(F)$, Points (1) and (2) are Corollary 6.13, Points (3) and (5) are Proposition 6.17, while Point (4) is Proposition 6.17 together with Corollaries 6.13 and 6.15.

For Φ_F , Points (3), (4), and (5) follow by the same methods, after isotoping the surface so the connect sum happens entirely on one side of the admissible cut. For Point (1), choose disjoint one-sided embedded curves $\gamma, \eta \subset F$ so that η contains a star, and then consider the admissible cut $S_{\leq \gamma}$, as in the proof of Proposition 4.2; then one side of the admissible cut contains a nonorientable component with a star, and so Φ_F vanishes by Corollary 6.13. \square

We conclude the section by singling out one consequence of Point (5). Miller-Powell introduced the notion of the *generalized stabilization distance* [MP19] (see also [Miy86, JZb]). In particular, surfaces F and F' have generalized stabilization distance 0 if and only if they are related by taking the connected sums with embedded 2-spheres. (See also [SS].) While they work in the topological category, we will continue to assume all surfaces are smoothly embedded.

Corollary 6.19. *If F and F' are cobordisms with $\mathcal{H}^\bullet(F) \neq \mathcal{H}^\bullet(F')$ or $\Phi_F \neq \Phi_{F'}$ (and the cobordisms have crosscap number ≥ 3) then F and F' have generalized stabilization distance > 0 .*

Remark 6.20. There is also an obstruction to destabilizing orientable surfaces from Heegaard Floer homology [JZa, Proposition 5.5].

6.3. Closed surfaces. The following shows that the mixed invariant is often zero for closed surfaces (and is always zero for connected closed surfaces). By contrast, in Section 7 we will see that for surfaces with boundary the mixed invariant does contain interesting information.

Theorem 6.21. *Let F be a closed surface with crosscap number ≥ 3 , normal Euler number $e(F)$, Euler characteristic $\chi(F)$, $s_o(F)$ stars on orientable components, and $s_n(F)$ stars on nonorientable components. If its mixed invariant Φ_F is non-zero then $e(F) = -2$, $s_n(F) = 0$, and $\chi(F) = 1 + 2s_o(F)$.*

Corollary 6.22. *If F is a closed, connected surface with crosscap number ≥ 3 then its mixed invariant Φ_F vanishes.*

Proof of Theorem 6.21. By Theorem 6.18, $s_n(F) = 0$.

Since the mixed invariant Φ_F is an $\mathbf{R}[U]$ -module homomorphism $\mathbf{R}[U] = \mathcal{H}^-(\emptyset) \rightarrow \mathcal{H}^+(\emptyset) = \mathbf{R}[U^{-1}, U]/\mathbf{R}[U]$, Φ_F may be viewed as an element of $\mathbf{R}[U^{-1}, U]/\mathbf{R}[U]$ (the image of 1). By Lemma 6.1, Φ_F is in bigrading $(-1 - e/2, \chi - 3e/2 - 2s_o)$. Since $\mathcal{H}^+(\emptyset)$ is supported in homological grading 0, this forces $e(F) = -2$.

In the Bar-Natan theory, by Corollary 6.15, $\Phi_F(1) \in \ker(H^2) \subset \mathbf{R}[H^{-1}, H]/\mathbf{R}[H]$, which is $\mathbf{R}\langle H^{-1}, H^{-2} \rangle$, supported in bigradings $(0, 2)$ and $(0, 4)$, which forces $(\chi - 2s_o) \in \{-1, 1\}$. In the Lee theory, again by Corollary 6.15, $\Phi_F(1) \in \ker(4T) \subset \mathbf{R}[T^{-1}, T]/\mathbf{R}[T]$, which is $\mathbf{R}\langle T^{-1} \rangle$ (recall that 2 is invertible in \mathbf{R}), supported in bigrading $(0, 4)$, forcing $\chi - 2s_o = 1$.

Thus, the only case remaining to exclude is $\chi - 2s_o = -1$. So, for the rest of the proof assume $e(F) = -2$, $\chi(F) - 2s_o(F) = -1$, and $s_n(F) = 0$. We need to show $\Phi_F = 0$. To settle this case, we

will need to study the mixed invariants over various Frobenius algebras, so we will use superscripts to denote the various Frobenius algebras that we are working over. We have already observed that the mixed invariant in the Lee theory, $\Phi_F^{\mathbb{R}[T,X]/(X^2=T)}$, vanishes for grading reasons over any ring \mathbb{R} (with 2 invertible).

Consider the Frobenius algebra $\mathbb{R}[\sqrt{T}, X]/(X^2 = T)$ obtained from the Lee Frobenius algebra by adjoining a formal square root of T , as in Proposition 2.1. Since $\mathbb{R}[\sqrt{T}]$ is free over $\mathbb{R}[T]$, all versions of the Khovanov chain complexes and homologies over the Frobenius algebra $\mathbb{R}[\sqrt{T}, X]/(X^2 = T)$ can be obtained from the corresponding versions over the Frobenius algebra $\mathbb{R}[T, X]/(X^2 = T)$ by tensoring with $\mathbb{R}[\sqrt{T}]$ over $\mathbb{R}[T]$; similarly, the maps over the Frobenius algebra $\mathbb{R}[\sqrt{T}, X]/(X^2 = T)$ can be obtained from the maps over the Frobenius algebra $\mathbb{R}[T, X]/(X^2 = T)$ by tensoring with $\mathbb{R}[\sqrt{T}]$ over $\mathbb{R}[T]$. Therefore, the mixed invariant over this new Frobenius algebra, $\Phi_F^{\mathbb{R}[\sqrt{T}, X]/(X^2=T)}$, vanishes over any ring \mathbb{R} (with 2 invertible). In particular, with $\mathbb{R} = \mathbb{Q}$, we get $\Phi_F^{\mathbb{Q}[\sqrt{T}, X]/(X^2=T)} = 0$.

Now consider the Bar-Natan Frobenius algebra over the rationals, $\mathbb{Q}[H, X]/(X^2 = HX)$. This is *twist-equivalent* to the above Frobenius algebra $\mathbb{Q}[\sqrt{T}, X]/(X^2 = T)$, in the sense of Khovanov [Kho06b]. Specifically, we have an isomorphism

$$\begin{aligned} \phi: \mathbb{Q}[\sqrt{T}, X]/(X^2 = T) &\rightarrow \mathbb{Q}[H, X]/(X^2 = HX) \\ \phi(1) &= 1, \quad \phi(X) = 2X - H, \quad \phi(\sqrt{T}) = H \end{aligned}$$

which preserves the algebra structure, and twists the counit η and comultiplication Δ by the invertible element $2 \in \mathbb{Q}$:

$$\eta(\phi(a)) = 2\phi(\eta(a)) \quad \Delta(\phi(a)) = \frac{1}{2}\phi(\Delta(a)) \quad \forall a \in \mathbb{Q}[\sqrt{T}, X]/(X^2 = T).$$

Khovanov's proof of invariance under twist equivalence [Kho06b, Proposition 3] works also for \mathcal{H}^- (respectively, \mathcal{H}^+), and shows that the \mathcal{H}^- (respectively, \mathcal{H}^+) Khovanov homologies over $\mathbb{Q}[\sqrt{T}, X]/(X^2 = T)$ and $\mathbb{Q}[H, X]/(X^2 = HX)$ are isomorphic. Moreover, the proof can be modified to see that for both versions, the maps induced by cobordisms agree over $\mathbb{Q}[\sqrt{T}, X]/(X^2 = T)$ and $\mathbb{Q}[H, X]/(X^2 = HX)$ up to multiplication by (possibly negative) powers of 2. Therefore, $\Phi_F^{\mathbb{Q}[H,X]/(X^2=HX)} = 2^k \Phi_F^{\mathbb{Q}[\sqrt{T}, X]/(X^2=T)}$ for some integer k , and hence is zero.

Finally, consider the Bar-Natan mixed invariant over the integers, $\Phi_F^{\mathbb{Z}[H,X]/(X^2=HX)}$. Recall its definition at the chain level. We choose an admissible cut and decompose the surface F into two cobordisms $F_0: \emptyset \rightarrow L$ and $F_1: L \rightarrow \emptyset$, and choose movies M_0 and M_1 describing F_0 and F_1 . Consider the generator $1 \in \mathcal{C}^-(\emptyset) = \mathbb{Z}[H]$, and its image $\mathcal{C}^-(F_0)(1) \in \mathcal{C}^-(L)$ induced by the movie M_0 . This is a boundary when viewed as a cycle in $\mathcal{C}^\infty(L)$; choose a chain $c \in \mathcal{C}^\infty(L)$ with $\partial c = \mathcal{C}^-(F_0)(1)$. Let \bar{c} be the image of c in $\mathcal{C}^+(L)$ obtained by removing all the terms with non-negative powers of H , and consider its image $\mathcal{C}^+(F_1)(\bar{c}) \in \mathcal{C}^+(\emptyset) = \mathbb{Z}[H^{-1}, H]/\mathbb{Z}[H]$ induced by the movie M_1 . Since we assumed $e = -2$, $\chi - 2s_o = -1$, and $s_n = 0$, $\mathcal{C}^+(F_1)(\bar{c})$ lies in bigrading $(0, 2)$, and so gives an element of \mathbb{Z} (which is the $\text{gr}_q = 2$ part of $\mathbb{Z}[H^{-1}, H]/\mathbb{Z}[H]$); this is the mixed invariant $\Phi_F^{\mathbb{Z}[H,X]/(X^2=HX)}(1)$.

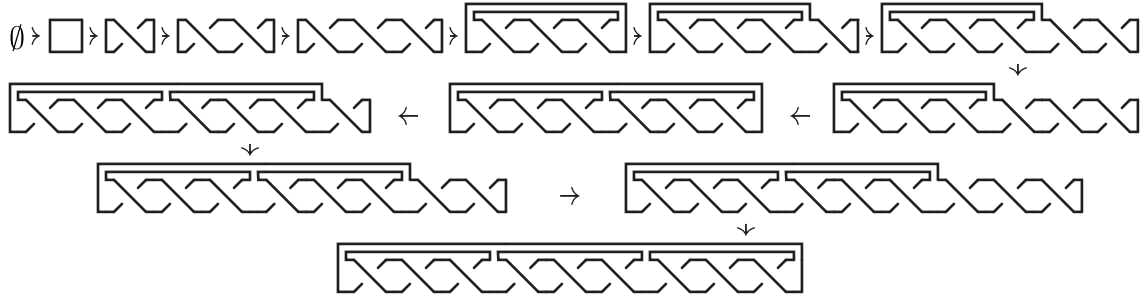
For any ring \mathbb{R} , if we tensor each step in the above chain-level description of the definition of $\Phi_F^{\mathbb{Z}[H,X]/(X^2=HX)}(1)$ with \mathbb{R} , we get a chain-level description of the definition of $\Phi_F^{\mathbb{R}[H,X]/(X^2=HX)}(1)$.

So, the Bar-Natan mixed invariant $\Phi_F^{\mathbb{R}[H,X]/(X^2=HX)}(1)$, viewed as an element of \mathbb{R} (which is the $\text{gr}_q = 2$ part of $\mathcal{C}^+(\emptyset) = \mathbb{R}[H^{-1}, H]/\mathbb{R}[H]$), can be obtained from the above element by tensoring with \mathbb{R} over \mathbb{Z} . Since $\Phi_F^{\mathbb{Q}[H,X]/(X^2=HX)}(1) = 0 \in \mathbb{Q}$, we get $\Phi_F^{\mathbb{Z}[H,X]/(X^2=HX)}(1) = 0 \in \mathbb{Z}$, and therefore, $\Phi_F^{\mathbb{R}[H,X]/(X^2=HX)}(1) = 0 \in \mathbb{R}$ for all rings \mathbb{R} . \square

7. COMPUTATIONS, APPLICATIONS, AND QUESTIONS

Theorems 6.18 and 6.21 give many examples where the mixed invariant vanishes. In this section, we use Lemma 6.9 to give some examples where it does not vanish, and note some corollaries.

7.1. A first direct computation. Let M denote the obvious Möbius band in S^3 with boundary the (right-handed) trefoil 3_1 . View the boundary sum $M \natural M \natural M$ as a cobordism from the empty link to $3_1 \# 3_1 \# 3_1$; explicitly, $M \natural M \natural M$ is given by the following movie:



This movie corresponds to a cobordism with crosscap number 3 and normal Euler number -18 (see the proof of Lemma 3.2, and perform one saddle at a time). We compute directly that the mixed invariant, with respect to the Bar-Natan deformation, is non-vanishing, and then observe that this also follows from Lemma 6.9. The frame 3_1 in the movie above is an admissible cut, decomposing the cobordism as $F_1 \circ F_0$. The normal Euler number of F_1 is -6 , so the map $\mathcal{H}^-(F_0): \mathbb{Z}[H] = \mathcal{H}^-(\emptyset) \rightarrow \mathcal{H}^-(3_1)$ shifts the $(\text{gr}_h, \text{gr}_q)$ -bigrading by $(3, 9)$. The image of $\mathcal{H}^-(\emptyset)$ at each stage of the movie F_0 lies in the all-1 resolution, and a generator labels each circle 1:

$$(7.1) \quad \begin{array}{c} \emptyset \rightarrow \square \rightarrow \text{[diagram]} \rightarrow \text{[diagram]} \rightarrow \text{[diagram]} \rightarrow \text{[diagram]} \rightarrow \text{[diagram]} \\ \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\ 1 \rightarrow \boxed{1} \rightarrow \boxed{1} \rightarrow \boxed{1} \rightarrow \boxed{1} \rightarrow \boxed{1} \rightarrow \boxed{1} \rightarrow \boxed{1} \rightarrow \boxed{1} \rightarrow \boxed{1} \rightarrow \boxed{1} \end{array}$$

The element $\mathcal{H}^-(F_0)(1) \in \mathcal{H}^-(3_1)$ is non-zero, but its image in \mathcal{H}^∞ is zero: the element $\mathcal{C}^-(F_0)(1)$ is the boundary of the following element:

$$\begin{aligned}
& -H^{-1} \left[\begin{array}{|c|c|c|} \hline \hline 1 & 1 & \\ \hline \hline \end{array} \right] + H^{-2} \left[\begin{array}{|c|c|c|} \hline \hline X & 1 & \\ \hline \hline \end{array} \right] - H^{-2} \left[\begin{array}{|c|c|c|} \hline \hline 1 & X & \\ \hline \hline \end{array} \right] \\
& 2H^{-2} \left[\begin{array}{|c|c|c|} \hline \hline X & 1 & \\ \hline \hline \end{array} \right] - H^{-2} \left[\begin{array}{|c|c|c|} \hline \hline 1 & X & \\ \hline \hline \end{array} \right] + H^{-2} \left[\begin{array}{|c|c|c|} \hline \hline X & 1 & \\ \hline \hline \end{array} \right].
\end{aligned}$$

(For computing the signs, we have ordered the crossings from left to right.) In particular, $\mathcal{H}^-(F_0)(1)$ is an element of $\mathcal{H}^{\text{red}}(3_1)$, as expected. The element of $\mathcal{C}^\infty(L)$ shown with boundary $\mathcal{C}^-(F_0)(1)$ lies in $\mathcal{C}^+(L)$, so to compute the mixed invariant, we apply $\mathcal{H}^+(F_1)$ to this element. The result is:

$$\begin{aligned}
& -H^{-1} \left[\begin{array}{|c|c|c|c|c|c|c|c|} \hline \hline 1 & 1 & & 1 & 1 & & 1 & 1 \\ \hline \hline \end{array} \right] + H^{-2} \left[\begin{array}{|c|c|c|c|c|c|c|c|} \hline \hline 1 & X & 1 & 1 & & 1 & 1 & \\ \hline \hline \end{array} \right] \\
& -H^{-2} \left[\begin{array}{|c|c|c|c|c|c|c|c|} \hline \hline 1 & X & & 1 & 1 & & 1 & 1 \\ \hline \hline \end{array} \right] + 2H^{-2} \left[\begin{array}{|c|c|c|c|c|c|c|c|} \hline \hline X & 1 & & 1 & 1 & & 1 & 1 \\ \hline \hline \end{array} \right] \\
& -H^{-2} \left[\begin{array}{|c|c|c|c|c|c|c|c|} \hline \hline 1 & X & & 1 & 1 & & 1 & 1 \\ \hline \hline \end{array} \right] + H^{-2} \left[\begin{array}{|c|c|c|c|c|c|c|c|} \hline \hline 1 & & X & 1 & 1 & & 1 & 1 \\ \hline \hline \end{array} \right]
\end{aligned}$$

We could compute directly that this is a nontrivial element of $\mathcal{H}^+(L_1)$, but it is slightly easier to apply the connecting homomorphism to $\hat{\mathcal{H}}(L_1)$. The image under the connecting homomorphism has a term

$$(7.2) \quad - \left[\begin{array}{|c|c|c|c|c|c|c|c|} \hline \hline 1 & 1 & 1 & & 1 & 1 & & 1 & 1 \\ \hline \hline \end{array} \right]$$

which does not appear in the boundary of any element of $\hat{\mathcal{C}}(L_1)$. So, the mixed invariant, in $\mathcal{H}^+(L_1)$, is nontrivial.

We can obtain the same result slightly more easily using Lemma 6.9. By that lemma, it suffices to show that the image of the class $1 \in \hat{\mathcal{H}}(L_0)$, under $\hat{\mathcal{H}}(F)$, is non-zero. A similar computation to Formula (7.1) shows that $\hat{\mathcal{H}}(F)(1)$ is the element in the all-1 resolution where every circle is labeled 1, i.e., the element shown in Formula (7.2). Since all maps into this resolution are split maps, this is a nontrivial element of $\hat{\mathcal{H}}(L_1)$.

In particular, by Theorem 6.18, this cobordism is not obtained by taking the connected sum of a crosscap-number 2 cobordism with \mathbb{RP}^2 or $\overline{\mathbb{RP}}^2$. The fact that the cobordism does not split off a copy of $\overline{\mathbb{RP}}^2$ also follows from the Gordon-Litherland formula [GL78]: if $F = F' \# \overline{\mathbb{RP}}^2$ then F' is a surface with $b_1(F') = 2$, $e(F') = -20$, and boundary $3_1 \# 3_1 \# 3_1$, so $\sigma(K) - e(F')/2 = 4$ but such a surface violates the inequality $|\sigma(K) - e(F')/2| \leq b_1(F')$. This inequality seems not to obstruct splitting off a copy of \mathbb{RP}^2 . This computation also provides a little more evidence that the 4-dimensional crosscap number of $3_1 \# 3_1 \# 3_1$ is 3, a conjecture which appears to be open.

By contrast, a similar direct computation to the above shows that the mixed invariant associated to the mirror of this cobordism, from \emptyset to $m(3_1) \# m(3_1) \# m(3_1)$, vanishes. (Here, we mean a different mirror from Section 6.1: the map $(t, x, y, z, w) \mapsto (t, -x, y, z, w)$.)

7.2. A more interesting example. Sundberg-Swann showed that the map on Khovanov homology distinguishes two slice disks for the knot 9_{46} . As we will see, their proof actually gives

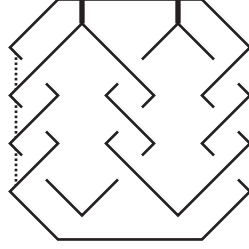


FIGURE 7.1. **The knot 9_{46} .** The slice disks Σ_L and Σ_R are obtained by attaching a saddle at the left and right thick lines, respectively, to obtain a 2-component unlink. The cobordism to $3_1 \# m(3_1)$ comes from attaching saddles at the three dotted lines.

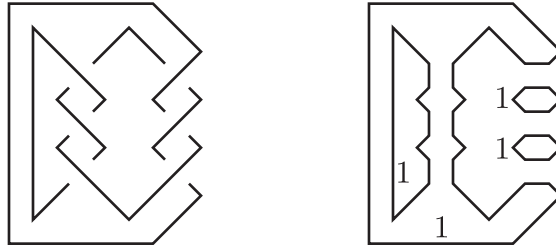


FIGURE 7.2. **The image of $\widehat{\mathcal{H}}(C \circ \Sigma_R)$.** Left: the diagram for $3_1 \# m(3_1)$ obtained by performing three saddle moves to 9_{46} . Right: the element $\widehat{\mathcal{H}}(C \circ \Sigma_R)$ lies in the all-1 resolution, and labels every circle by 1.

somewhat more: two distinct punctured $\mathbb{RP}^2 \# \mathbb{RP}^2 \# \mathbb{RP}^2$ s with boundary $3_1 \# m(3_1)$ and normal Euler number -6 .

Recall that the knot 9_{46} has two slice disks, corresponding to attaching saddles at two of the handles shown in Figure 7.1; we will refer to these as the left and right slice disks Σ_L and Σ_R , respectively. We will view Σ_L and Σ_R as cobordisms from \emptyset to 9_{46} . There is also a cobordism C from 9_{46} to $3_1 \# m(3_1)$ with crosscap number 3 and normal Euler number -6 , obtained by attaching three saddles to 9_{46} ; again, see Figure 7.1. (Attaching just one of these saddles gives 8_{20} , and attaching two gives 6_1 .)

Sundberg-Swann call each of these three saddles a *trim cobordism*. They show, by direct computation, that the map $\widehat{\mathcal{H}}(C \circ \Sigma_L) = 0$ while $\widehat{\mathcal{H}}(C \circ \Sigma_R)$ sends the generator $1 \in \widehat{\mathcal{H}}(\emptyset)$ to the class in $\widehat{\mathcal{H}}(3_1 \# m(3_1))$ shown in Figure 7.2 [SS, Proof of Theorem 6.3]. In particular, by Proposition 3.7, the surfaces $C \circ \Sigma_L$ and $C \circ \Sigma_R$ are not smoothly isotopic.

By Lemma 6.9, the mixed invariant $\Phi_{C \circ \Sigma_R}$ is non-zero, as is $\partial \circ \Phi_{C \circ \Sigma_R}$. By contrast, $\partial \circ \Phi_{C \circ \Sigma_L} = 0$. The element $\Phi_{C \circ \Sigma_L}$ lies in bigrading $(2, 7)$, and the generator in this bigrading is not in the image of multiplication by U ; see Figure 7.3. So, from the long exact sequence (2.2), ∂ is injective in this bigrading, so $\Phi_{C \circ \Sigma_L} = 0$, as well.

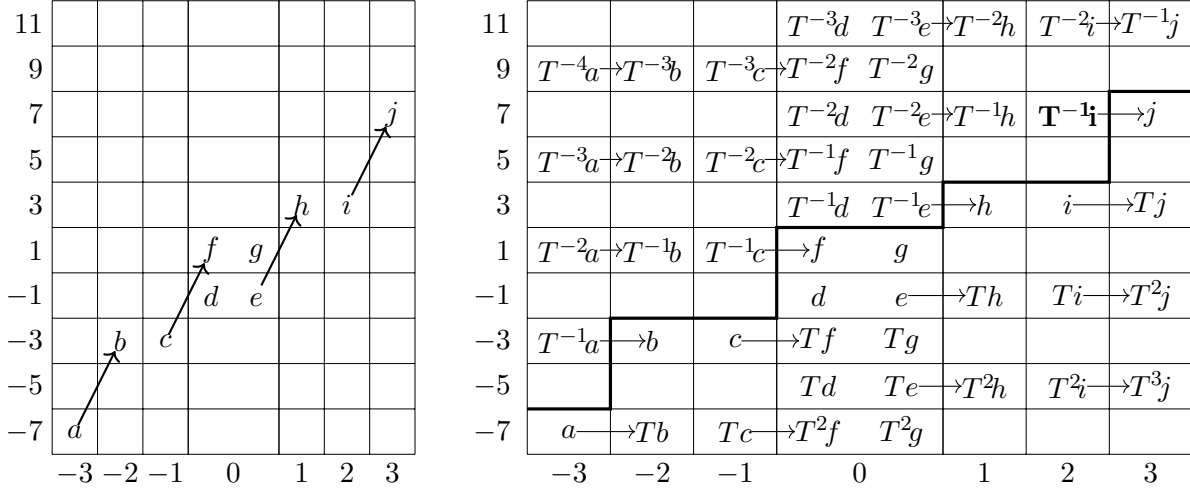


FIGURE 7.3. **Computing $\mathcal{H}^+(3_1 \# m(3_1))$.** Left: $\widehat{\mathcal{H}}$ as computed by knotkit [See], with \mathbb{Q} coefficients. Each letter is a basis element over \mathbb{Q} . (This computation can also be deduced from $\mathcal{H}(3_1)$, with a little work.) The arrows are the differentials in the Lee spectral sequence, which one can deduce from knowing that the Lee homology is $\mathbb{Q} \oplus \mathbb{Q}$ in bidegrees $(0, \pm 1)$, since $s(3_1 \# m(3_1)) = 0$. Right: the differentials on $\mathcal{H}^\infty(3_1 \# m(3_1))$, which one can read off from the top-left computation. The subcomplex \mathcal{C}^- lies below the thick steps, and the quotient complex \mathcal{C}^+ lies above the thick steps. The generator of \mathcal{H}^+ in bigrading $(2, 7)$ is in bold; U^{-1} times this generator is not a cycle in \mathcal{C}^+ . The analogous computation for the Bar-Natan deformation is similar, but the differential on the left has bi-degree $(1, 2)$ instead of $(1, 4)$, and the variable H has bi-degree $(0, -2)$.

In conclusion, both the map $\widehat{\mathcal{H}}$ and the mixed invariant distinguish this pair of surfaces. By Theorem 6.18, this implies that $C \circ \Sigma_L$ and $C \circ \Sigma_R$ do not differ by taking a connected sum with a smoothly embedded 2-sphere. Further, non-vanishing of $\widehat{\mathcal{H}}(C \circ \Sigma_R)$ implies that $C \circ \Sigma_R$ is not obtained from another connected surface by attaching a 1-handle, and is not a crosscap stabilization. Hence, we have proved Theorem 1.1.

Remark 7.1. Using one of the three dotted saddles in Figure 7.1 gives a pair of Möbius bands with boundary 8_{20} distinguished by Khovanov homology, and using two of them gives a pair of punctured Klein bottles with boundary 6_1 distinguished by Khovanov homology.

7.3. An exotic pair of surfaces. Recall that a pair of surfaces $F, F' \subset B^4$ with boundary $K \subset S^3$ is *exotic* if there is a homeomorphism $\phi: B^4 \rightarrow B^4$ so that $\phi|_{S^3} = \text{Id}$ and $\phi(F) = F'$, but no diffeomorphism with these properties. (See also Remark 7.3.) Hayden-Sundberg give a family of exotic pairs of surfaces [HS]. The simplest of their pairs is the pair of slice disks shown in Figure 7.4 for the knot J . The slice disk D (respectively D') is obtained by attaching a saddle along the arc b (respectively b') shown there, and then capping the resulting 2-component unlink with disks. The fact that these surfaces are distinct is witnessed by the map on Khovanov homology: for the

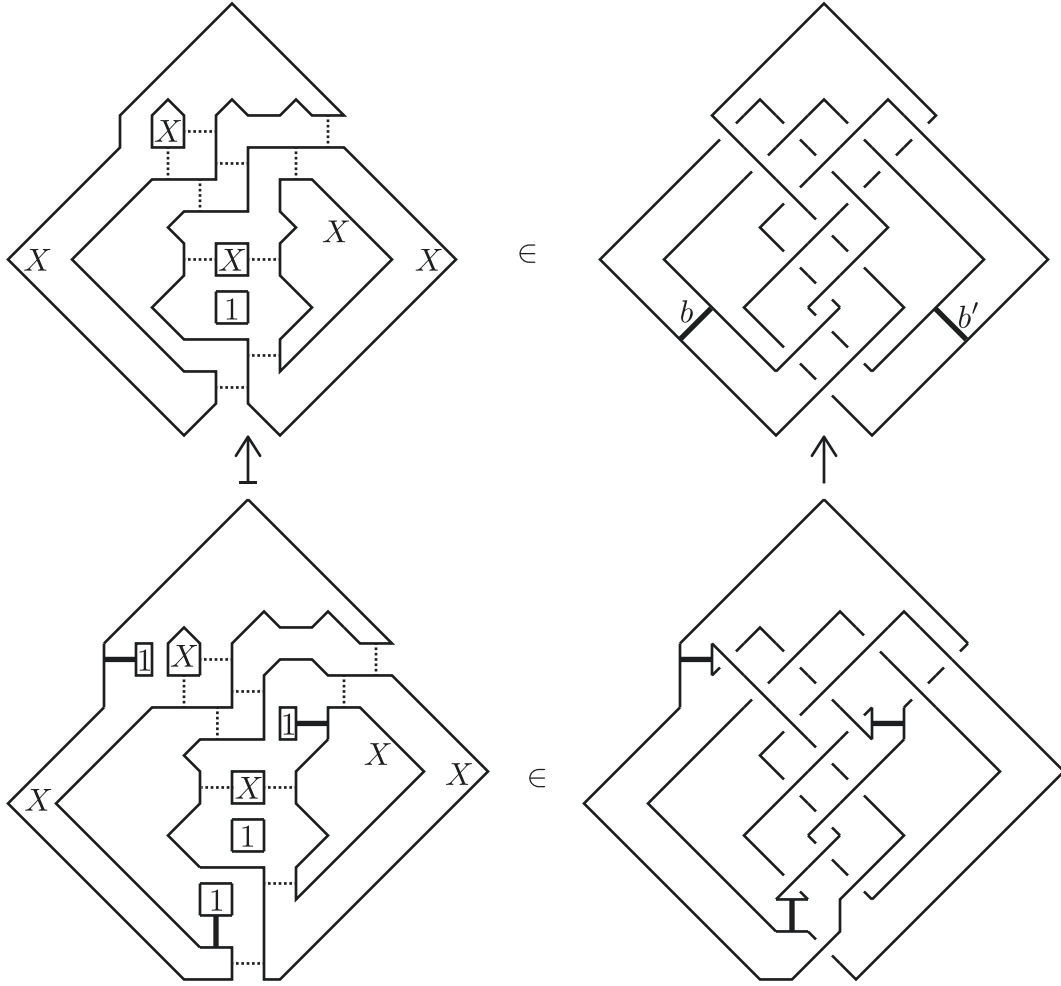


FIGURE 7.4. **An exotic pair.** Top right: Hayden-Sundberg's knot J and the two saddles b and b' giving an exotic pair of distinct slice disks for J , drawn as thick line segments. Top-left: Hayden-Sundberg's cycle ϕ in $\hat{\mathcal{C}}(J)$ witnessing the fact that these slice disks are distinct. Bottom-right: a diagram for 12_{309}^n and three saddles giving a nonorientable cobordism to J . Bottom-left: a cycle ψ in $\hat{\mathcal{C}}(12_{309}^n)$ which maps to Hayden-Sundberg's cycle. As in Hayden-Sundberg's figure, dotted lines indicate which crossings had 0-resolutions.

element ϕ of $\hat{\mathcal{H}}(J)$ shown in Figure 7.4, $\hat{\mathcal{H}}(D')(\phi) = 0$ (obvious) but $\hat{\mathcal{H}}(D)(\phi) = 1 \in \mathbb{Z} = \hat{\mathcal{H}}(\emptyset)$ [HS, Figure 4].

There is a cobordism C with crosscap number 3 and normal Euler number -6 from the knot 12_{309}^n to J , so that $\phi = \widehat{\mathcal{H}}(\psi)$ for an appropriate class $\psi \in \widehat{\mathcal{H}}(12_{309}^n)$: see Figure 7.4. So, $\widehat{\mathcal{H}}(D \circ C)(\psi) = 1$ while $\widehat{\mathcal{H}}(D' \circ C)(\psi) = 0$. Thus, $D \circ C$ is not diffeomorphic to $D' \circ C$ rel boundary. On the other hand, since D is homeomorphic to D' rel boundary, $D \circ C$ is homeomorphic to $D' \circ C$ rel boundary. Thus, we have proved Theorem 1.2.

By the second case of Lemma 6.9, the mixed invariant also distinguishes $D \circ C$ and $D' \circ C$ (compare Remark 6.10). The fact that the pair $D \circ C$ and $D' \circ C$ are not diffeomorphic is, of course, slightly stronger than the statement that D and D' are not diffeomorphic.

Remark 7.2. Hayden-Sundberg's example also immediately gives an exotic pair of crosscap number 3 surfaces with boundary 12_{404}^n , as well as exotic pairs of crosscap number 3 surfaces with boundary on several links, by an easy adaptation of Figure 7.4.

Remark 7.3. Hayden-Sundberg take a slightly different definition of *exotic* than we have: they define a pair of surfaces $F, F' \subset B^4$ to be exotic if there is an ambient isotopy through homeomorphisms taking F to F' but no ambient isotopy through diffeomorphisms (which are, in both cases, the identity on S^3). Since their surfaces are distinguished by the map on Khovanov homology, however, by Proposition 3.7, their computation shows that there is no diffeomorphism from B^4 to itself which is the identity on the boundary and takes F to F' (even one which is not isotopic to the identity). That is, their pairs of surfaces really are exotic in the sense described above.

7.4. Some questions. To put the results above in context, and in particular to acknowledge the cases they do not cover, we conclude with some open questions.

In Corollary 6.22, we showed that the mixed invariant does not distinguish closed, connected surfaces. By Proposition 3.5 for the nonorientable case and work of Gujral-Levine [GL] for the orientable case, the map on \mathcal{H}^\bullet also does not distinguish disconnected surfaces.

Question 1. Is there a pair $F, F' \subset S^4$ of closed, disconnected surfaces with the same topology and componentwise normal Euler numbers so that $\Phi_F \neq \Phi_{F'}$?

If we are only considering surfaces without stars, by Theorem 6.21, these surfaces would have to have (total) normal Euler number -2 and Euler characteristic 1. For example, perhaps F could be a knotted copy of $\mathbb{RP}^2 \amalg (\mathbb{RP}^2 \# \overline{\mathbb{RP}}^2)$ with total normal Euler number -2 and non-vanishing mixed invariant, and F' the standard $\mathbb{RP}^2 \amalg (\mathbb{RP}^2 \# \overline{\mathbb{RP}}^2)$, which has vanishing mixed invariant.

The Seiberg-Witten invariant is not just defined when $b_2^+ \geq 3$, but also when $b_2^+ = 2$. As noted in Remark 5.4, we can define a Khovanov mixed invariant when the crosscap number is 2, but we do not know if it is well-defined.

Question 2. If F, F' are isotopic surfaces (rel boundary) with crosscap number 2 and admissible cuts (S, V, ϕ) and (S', V', ϕ') , respectively, is the mixed invariant of F with respect to (S, V, ϕ) equal to the mixed invariant of F' with respect to (S', V', ϕ') ?

In the examples in Sections 7.2 and 7.3 of surfaces distinguished by their mixed invariants, the surfaces were also distinguished by the induced maps on ordinary Khovanov homology $\widehat{\mathcal{H}}$. By Remark 6.10, for surfaces with connected boundary, the mixed invariant is at least as strong as the map on $\widehat{\mathcal{H}}$.

Question 3. Is there a pair of surfaces with crosscap number ≥ 3 and the same topology and normal Euler number which are distinguished by the mixed invariant but not by the map on $\widehat{\mathcal{H}}$? Is there such a pair not distinguished by the homotopy class of maps on \mathcal{C}^- ? Is there an exotic pair of surfaces with this property?

Lemma 6.9 and its proof give restrictions on what the Khovanov homology of such a pair must look like, as does Corollary 6.15.

On a related point, by Theorem 6.18, the map on \mathcal{H}^\bullet vanishes for stabilizations of surfaces, so never distinguishes them. There is one case in which the mixed invariant could potentially distinguish stabilized surfaces:

Question 4. Is there an exotic pair of Möbius bands $F, F' \subset B^4$ so that the mixed invariant distinguishes their stabilizations? That is, if $F \# T^2$ denotes a standard stabilization of F , is there an exotic pair of Möbius bands F, F' with boundary some knot K so that $\Phi_{F \# T^2} \neq \Phi_{F' \# T^2} \in \mathcal{H}^+(K)$?

The examples of nonorientable surfaces in Sections 7.2 and 7.3 came from pairs of slice disks. Indeed, the nonorientable surfaces were apparent from the slice disks and class in Khovanov homology. Perhaps this phenomenon is general:

Question 5. Is it true that for every exotic pair of slice disks D, D' , for any knot K , there is a nonorientable cobordism F from K to another knot K' so that $F \circ D$ and $F \circ D'$ are also an exotic pair? Can F be chosen to have crosscap number ≥ 3 ?

One can ask the same question, but for exotic pairs detected by Khovanov homology:

Question 6. Is it true that for every pair of slice disks D, D' , for a knot K such that $\widehat{\mathcal{H}}(D) \neq \widehat{\mathcal{H}}(D')$, there is a nonorientable cobordism F from K to another knot K' so that $\widehat{\mathcal{H}}(F \circ D) \neq \widehat{\mathcal{H}}(F \circ D')$?

One could also replace $\widehat{\mathcal{H}}$ by \mathcal{H}^\bullet in the question, or require that F have crosscap number ≥ 3 and ask if $\Phi_{F \circ D} \neq \Phi_{F \circ D'}$.

As noted in the introduction, our mixed invariant is inspired by Ozsváth-Szabó's mixed invariant in Heegaard Floer homology.

Question 7. Is there a precise relationship between the Khovanov mixed invariant of a surface F and the Heegaard Floer mixed invariant of the branched double cover of F ?

Note that having crosscap number ≥ 3 does not give an inequality for b_2^+ (consider the standard $\mathbb{RP}^2 \# \mathbb{RP}^2 \# \mathbb{RP}^2$), nor does having $b_2^+(\Sigma(F)) \geq 2$ imply crosscap number ≥ 3 (consider $\mathbb{RP}^2 \# \mathbb{RP}^2$ or, for that matter, an orientable surface of genus $g \geq 2$), so the two invariants are not defined in exactly the same cases; perhaps this argues against a direct relationship.

REFERENCES

- [Ali19] Akram Alishahi, *Unknotting number and Khovanov homology*, Pacific J. Math. **301** (2019), no. 1, 15–29.
- [AN06] A. Alexeevski and S. Natanzon, *Noncommutative two-dimensional topological field theories and Hurwitz numbers for real algebraic curves*, Selecta Math. (N.S.) **12** (2006), no. 3-4, 307–377.
- [Bal] William Ballinger, *Concordance invariants from the $E(-1)$ spectral sequence on Khovanov homology*, arXiv:2004.10807.
- [Bar05] Dror Bar-Natan, *Khovanov's homology for tangles and cobordisms*, Geom. Topol. **9** (2005), 1443–1499.

- [BNM06] Dror Bar-Natan and Scott Morrison, *The Karoubi envelope and Lee’s degeneration of Khovanov homology*, Algebr. Geom. Topol. **6** (2006), 1459–1469.
- [BS16] R. İnanç Baykur and Nathan Sunukjian, *Knotted surfaces in 4-manifolds and stabilizations*, J. Topol. **9** (2016), no. 1, 215–231.
- [CMW09] David Clark, Scott Morrison, and Kevin Walker, *Fixing the functoriality of Khovanov homology*, Geom. Topol. **13** (2009), no. 3, 1499–1582.
- [CS93] J. Scott Carter and Masahico Saito, *Reidemeister moves for surface isotopies and their interpretation as moves to movies*, J. Knot Theory Ramifications **2** (1993), no. 3, 251–284.
- [FKV87] S. M. Finashin, M. Kreck, and O. Ya. Viro, *Exotic knottings of surfaces in the 4-sphere*, Bull. Amer. Math. Soc. (N.S.) **17** (1987), no. 2, 287–290.
- [FKV88] ———, *Nondiffeomorphic but homeomorphic knottings of surfaces in the 4-sphere*, Topology and geometry—Rohlin Seminar, Lecture Notes in Math., vol. 1346, Springer, Berlin, 1988, pp. 157–198.
- [GL] Onkar Singh Gujral and Adam Simon Levine, *Khovanov homology and cobordisms between split links*, arXiv:2009.03406.
- [GL78] C. McA. Gordon and R. A. Litherland, *On the signature of a link*, Invent. Math. **47** (1978), no. 1, 53–69.
- [Hay] Kyle Hayden, *Exotically knotted disks and complex curves*, arXiv:2003.13681.
- [HKK⁺] Kyle Hayden, Alexandra Kjuchukova, Siddhi Krishna, Maggie Miller, Mark Powell, and Nathan Sunukjian, *Brunnian exotic surface links in the 4-ball*, arXiv:2106.13776.
- [HN13] Matthew Hedden and Yi Ni, *Khovanov module and the detection of unlinks*, Geom. Topol. **17** (2013), no. 5, 3027–3076.
- [HS] Kyle Hayden and Isaac Sundberg, *Khovanov homology and exotic surfaces in the 4-ball*, arXiv:2108.04810.
- [Jac04] Magnus Jacobsson, *An invariant of link cobordisms from Khovanov homology*, Algebr. Geom. Topol. **4** (2004), 1211–1251 (electronic).
- [JMZ21] András Juhász, Maggie Miller, and Ian Zemke, *Transverse invariants and exotic surfaces in the 4-ball*, Geom. Topol. **25** (2021), no. 6, 2963–3012.
- [JZa] András Juhász and Ian Zemke, *New Heegaard Floer slice genus and clasp number bounds*, arXiv:2007.07106.
- [JZb] ———, *Stabilization distance bounds from link Floer homology*, arXiv:1810.09158.
- [Kho00] Mikhail Khovanov, *A categorification of the Jones polynomial*, Duke Math. J. **101** (2000), no. 3, 359–426.
- [Kho06a] ———, *An invariant of tangle cobordisms*, Trans. Amer. Math. Soc. **358** (2006), no. 1, 315–327.
- [Kho06b] ———, *Link homology and Frobenius extensions*, Fund. Math. **190** (2006), 179–190.
- [Kor02] Mustafa Korkmaz, *Mapping class groups of nonorientable surfaces*, Geom. Dedicata **89** (2002), 109–133.
- [KR22] Mikhail Khovanov and Louis-Hadrien Robert, *Link homology and Frobenius extensions II*, Fund. Math. **256** (2022), no. 1, 1–46.
- [Lee05] Eun Soo Lee, *An endomorphism of the Khovanov invariant*, Adv. Math. **197** (2005), no. 2, 554–586.
- [LZ19] Adam Simon Levine and Ian Zemke, *Khovanov homology and ribbon concordances*, Bull. Lond. Math. Soc. **51** (2019), no. 6, 1099–1103.
- [Mel77] Paul Michael Melvin, *Blowing up and down in 4-manifolds*, ProQuest LLC, Ann Arbor, MI, 1977, Thesis (Ph.D.)—University of California, Berkeley.
- [Miy86] Katura Miyazaki, *On the relationship among unknotting number, knotting genus and Alexander invariant for 2-knots*, Kobe J. Math. **3** (1986), no. 1, 77–85.
- [MP19] Allison N. Miller and Mark Powell, *Stabilization distance between surfaces*, Enseign. Math. **65** (2019), no. 3–4, 397–440.
- [MWW] Scott Morrison, Kevin Walker, and Paul Wedrich, *Invariants of 4-manifolds from Khovanov-Rozansky link homology*, arXiv:1907.12194.
- [OSS17] Peter S. Ozsváth, András I. Stipsicz, and Zoltán Szabó, *Unoriented knot Floer homology and the unoriented four-ball genus*, Int. Math. Res. Not. IMRN (2017), no. 17, 5137–5181.
- [OSz04] Peter Ozsváth and Zoltán Szabó, *Holomorphic disks and topological invariants for closed three-manifolds*, Ann. of Math. (2) **159** (2004), no. 3, 1027–1158.
- [OSz05] ———, *On the Heegaard Floer homology of branched double-covers*, Adv. Math. **194** (2005), no. 1, 1–33.

- [OSz06] ———, *Holomorphic triangles and invariants for smooth four-manifolds*, Adv. Math. **202** (2006), no. 2, 326–400.
- [Pie16] Alex Pieloch, *Curve complexes of non-orientable surfaces*, https://lsa.umich.edu/content/dam/math-assets/math-document/reu-documents/Pieloch_REUPaper.pdf, 2016.
- [Ras] Jacob Rasmussen, *Khovanov’s invariant for closed surfaces*, arXiv:math/0502527.
- [Ras10] ———, *Khovanov homology and the slice genus*, Invent. Math. **182** (2010), no. 2, 419–447.
- [Sar20] Sucharit Sarkar, *Ribbon distance and Khovanov homology*, Algebr. Geom. Topol. **20** (2020), no. 2, 1041–1058.
- [See] Cotton Seed, *Knotkit*, <https://github.com/cseed/knotkit>.
- [SS] Isaac Sundberg and Jonah Swann, *Relative Khovanov-Jacobsson classes*, arXiv:2103.01438.
- [Tan06] Kokoro Tanaka, *Khovanov-Jacobsson numbers and invariants of surface-knots derived from Bar-Natan’s theory*, Proc. Amer. Math. Soc. **134** (2006), no. 12, 3685–3689.
- [TT06] Vladimir Turaev and Paul Turner, *Unoriented topological quantum field theory and link homology*, Algebr. Geom. Topol. **6** (2006), 1069–1093.
- [Tur20] Paul Turner, *Khovanov homology and diagonalizable Frobenius algebras*, J. Knot Theory Ramifications **29** (2020), no. 1, 1950095, 10.

Email address: lipshitz@uoregon.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OREGON, EUGENE, OR 97403

Email address: sucharit@math.ucla.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 90095