AGT

Algebraic & Geometric Topology

Volume 23 (2023)

An algorithmic definition of Gabai width

Ricky Lee

msp

Algebraic & Geometric Topology 23:6 (2023) 2415–2448
DOI: 10.2140/agt.2023.23.2415
Published: 7 September 2023

An algorithmic definition of Gabai width

RICKY LEE

We define the Wirtinger width of a knot and prove that this equals its Gabai width. This leads to an efficient technique for establishing upper bounds on Gabai width. We demonstrate an application of this technique by calculating the Gabai width of 54 756 tabulated prime 4–bridge knots. This is done by writing code for a special category of prime 4–bridge tabulated knots to get upper bounds on Gabai width via the Wirtinger width, then comparing with the theoretical lower bound on Gabai width for prime 4–bridge knots. We also provide results showing the advantages our methods have over the obvious method of obtaining upper bounds on Gabai width via planar projections.

57M25, 57M27

1 Introduction

Gabai width is a geometric invariant of knots that was first used by Gabai in his proof of the property R conjecture [6]. Since then, the notion of Gabai width has played central roles in many important results in 3–manifold topology. Some examples are the resolution of the knot complement problem by Gordon and Luecke [8], the recognition problem for S^3 by Thompson [12], and the leveling of unknotting tunnels by Goda, Scharlemann and Thompson [7]. The importance of Gabai width is largely due to its deep connections with the topology of the knot exterior. For example, Gabai width can often be used to find incompressible surfaces; see Thompson [13] and Wu [15].

The bridge number of a knot is a closely related geometric invariant, defined as the minimal number of local maxima needed to construct an embedding of the knot. Roughly speaking, Gabai width depends on the number of critical points of a projection as well as their relative heights. Like most geometric invariants, both bridge number and Gabai width are notoriously difficult to calculate. However, there has been recent progress on finding algorithmically accessible definitions of bridge number. Blair, Kjuchukova, Velazquez and Villanueva [4] defined the Wirtinger number of a link and showed that it is equal to the bridge number. The Wirtinger number is calculated using a combinatorial coloring algorithm applied to a link diagram. Using ideas inspired by the Wirtinger number, we define the Wirtinger width of a knot and show it is equal to the Gabai width of a knot.

We now briefly summarize our procedure. The formal definition of Wirtinger width is given in Section 3. The Wirtinger width is also computed by coloring knot diagrams. Let D be a knot diagram. View D as the image of the knot $K \subset \mathbb{R}^3$ under the standard projection onto the xy–plane. Our goal, given the diagram D, is to obtain a knot K' in the same ambient isotopy class of K, but embedded so that K' realizes the Gabai width. Our coloring procedure allows us to obtain a knot $\widehat{K}$ from D such that $\widehat{K}$ is ambient isotopic to K, and the relative heights of the critical points of $\widehat{K}$ are controlled by combinatorial data attached to our coloring.

The coloring proceeds as follows. Suppose the knot diagram D has J strands. Then there are $J+1$ stages in the procedure. The knot diagram D begins uncolored at stage 0. To transition from one stage to the next, one can either add a new color to an uncolored strand, or extend an existing color to include another uncolored strand. The procedure terminates once all strands of D are colored.

In general, there are many different ways to color a knot diagram. Not all colorings will give data which corresponds to a thin position embedding of the knot. We assign a natural number to each coloring of a knot diagram, then let the Wirtinger width of the diagram D, denoted by $\mathbb{W}(D)$, be the minimum of these numbers over all colorings of D. Finally, for any ambient isotopy class of knots $\mathcal{K}$, we define the Wirtinger width of $\mathcal{K}$, denoted by $\mathbb{W}(\mathcal{K})$, to be the minimum of $\mathbb{W}(D)$ over all diagrams of knots in the ambient isotopy class $\mathcal{K}$. Letting $w(\mathcal{K})$ be the Gabai width of $\mathcal{K}$, we can state our main theorem as follows:

Theorem 1.1 *If $\mathcal{K}$ is an ambient isotopy class of knots, then* $\mathbb{W}(\mathcal{K}) = w(\mathcal{K})$.

The coloring can be viewed as an attempt to discretize the following process. Suppose now $K \subset \mathbb{R}^3$ is a knot in thin position with respect to the standard height function $h(x,y,z) := z$. Let $h^{-1}(r)$ be a level surface above K. The Gabai width of K is calculated by analyzing the intersection set $K \cap h^{-1}(r)$ as $r \to -\infty$ and $h^{-1}(r)$ sweeps across the maxima and minima of K. The addition of a new color to D represents $h^{-1}(r)$ sweeping across a maximum of K. The occurrence of a *multicolored crossing* (crossings where the over-strand is colored and both under-strands are assigned different colors) represents $h^{-1}(r)$ sweeping across a minimum of K. The order in which new colors and multicolored crossings appear in our coloring procedure dictates the ordering of the maxima and minima of $\widehat{K}$ by height.

There is an easy method of obtaining upper bounds on Gabai width. One can take a knot diagram, perform some planar isotopies if necessary, and use the original Gabai definition of width to obtain an upper bound in the obvious way. While our coloring procedure is less straightforward, it is more computationally accessible and enjoys the following advantage over any potential algorithm written to calculate upper bounds on Gabai width utilizing only planar isotopies on a knot diagram. Let $w_p(D)$ denote the planar width of a knot diagram D. A formal definition of planar width will be given in Section 2, but, roughly speaking, $w_p(D)$ is the upper bound on width one would get by applying the original Gabai definition to calculate width on D, after minimizing over all planar isotopies of D. We will prove:

Theorem 1.2 *For any ambient isotopy class $\mathcal{K}$ of knots and any positive integer n, there exist infinitely many diagrams D of knots in $\mathcal{K}$ such that $\mathbb{W}(D) = w(\mathcal{K})$ but $w_p(D) \geq \mathbb{W}(D) + n$.*

Colloquially, Theorem 1.2 states that, if a planar isotopy algorithm were to be implemented, there would still be an infinite number of cases where Wirtinger width performs better.

Since there are many different ways to completely color a knot diagram, the problem of finding a coloring which corresponds to a calculation of Gabai width is subtle. However, one can modify the Wirtinger number algorithm of Villanueva [14] to exhaust all possible colorings of a given diagram. This is possible because the rules for extending a coloring in the Wirtinger width procedure are the same as those for extending a coloring in the Wirtinger number procedure. We illustrate these ideas in Section 8, where we describe an algorithm that we implemented in Python [10] and used to calculate the Gabai width of 54 756 prime 4–bridge knots.

Our algorithm runs fast in practice, but depends on knowing beforehand that the inputted Gauss codes are of prime knots with bridge number 4 and such that the code from [14] can actually detect bridge number 4. The algorithm takes as input such a Gauss code, and outputs upper bounds on Wirtinger width. By Theorem 1.1, this gives upper bounds on Gabai width. It is known, and explained in Section 8, that the Gabai width of a prime 4–bridge knot must be 32 or 28. Of 86 981 knots tested, our code gave an upper bound of 28 on Wirtinger width for 54 756 knots. Since our upper bound equals the theoretical lower bound on Gabai width for such prime 4–bridge knots, this means we got the exact Gabai width in this case.

Structure of the paper In Section 2, we give preliminary definitions. In Section 3, we give the formal definition of Wirtinger width via a coloring procedure similar to the coloring algorithm of Wirtinger number in [4]. Section 4 contains results showing how Wirtinger number is related to Wirtinger width. In Section 5, we describe a specific coloring sequence, which, when performed on a projection of a knot in thin position, shows that $\mathbb{W}(\mathcal{K}) \leq w(\mathcal{K})$. In Section 6, we show how to use our coloring data to obtain Morse embeddings of knots from a colored knot diagram. This is used to show $\mathbb{W}(\mathcal{K}) \leq w(\mathcal{K})$. In Section 7, we use the results of the previous sections to prove Theorems 1.1 and 1.2. Many technical lemmas and results from Sections 4, 5 and 6 do not apply to diagrams of the unknot, so Section 7 handles this special case separately. In Section 8, we explain how we used Wirtinger width to write an algorithm in Python that obtained our numerical results, and present some open questions.

Acknowledgments The author would like to thank Ryan Blair for introducing this topic and for many helpful discussions, especially about Theorem 1.2. We also thank the referee for a close reading of our initial drafts, leading to many corrections and a great improvement to our exposition.

2 Preliminaries

Let $\mathcal{K}$ denote an ambient isotopy class of knots in $\mathbb{R}^3$. As stated in the introduction, let $h\colon \mathbb{R}^3 \to \mathbb{R}$ defined by $h(x, y, z) := z$ be the standard height function. Let $K \subset \mathbb{R}^3$ denote a knot in the ambient isotopy class $\mathcal{K}$. We will always assume that the embedding of K is such that $h|_K$ is a Morse function.

Let $p\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by $p(x, y, z) := (y, z)$ be the projection map onto the yz–plane. We will always assume K is embedded so that $p|_K$ is a regular projection. Then $p(K)$ is a finite four-valent graph in the yz–plane. We say that D is a *knot diagram* of K resulting from the projection p if D is the graph $p(K)$ together with labels at each vertex to indicate which edges are over and which are under. By convention, these labels take the form of deleting parts of the under-arc at every crossing. Thus, we can view D as a disjoint union of closed arcs in the plane. Let $\alpha_1, \dots, \alpha_J$ denote the connected components of D. For each α_i, we let s_i denote the union of all edges in

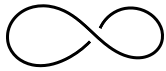

Figure 1: The unique knot diagram containing a strand adjacent to itself.

$p(K)$ whose interiors have nonempty intersection with α_i. We refer to each s_i as a *strand* and let $s(D)$ denote the set of strands of D. We refer to the vertices of $p(K)$ as *crossings* and denote the set of vertices by $v(D)$.

If $s \in s(D)$, then the two endpoints of s will be referred to as the crossings *incident* to s. If s_p and s_q are the under-strands of the same crossing $x \in v(D)$, then we say s_p and s_q are *adjacent at* x, or just *adjacent*. We say the subset $A \subseteq s(D)$ is *connected* if there exists a reordering of the strands $s_{i_1}, s_{i_2}, \dots, s_{i_{|A|}}$ in A such that s_{i_j} is adjacent to $s_{i_{j+1}}$ for all $1 \le j \le |A|$. Note that there is a unique knot diagram up to planar isotopy for which there exists a strand adjacent to itself (see Figure 1). In all cases considered, we assume that adjacent strands are distinct. We say a knot diagram is *trivial* if it is a diagram of the unknot.

For $s \in s(D)$, we define $h(s) := \max_{y \in s} h(y)$ and refer to $h(s)$ as the *height of the strand* s. For a crossing $x \in v(D)$, we refer to $h(x)$ as the *height of the crossing* x.

Note we do not consider the labels of the knot diagram when we calculate the height of a strand. It is therefore possible that a strand and a crossing have equal heights. In fact, if a strand is monotonic with respect to h, then it must have height equal to one of its incident crossings.

By *critical points* of D we will always be referring to images of the critical points of $h|_K$ under the projection p. We say that D is in *general position with respect to* h if all the critical points and crossings of D have distinct heights with respect to h, $h|_K$ is Morse, and $p(K)$ is a regular projection. Observe that, if the knot diagram D is in general position with respect to h, then all the strands must have different heights. See Figure 2.

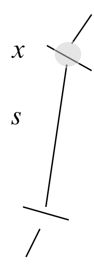

Figure 2: The strand s and the incident crossing x have equal heights ($h(s) = h(x)$).

Now we recall the definition of bridge number. We let $\beta(K)$ denote the number of maxima of $h|_K$. Then the bridge number $\beta(\mathcal{K})$ is defined as $\min_{K' \in \mathcal{K}} \beta(\mathcal{K})$, where the minimum is taken over all Morse embeddings of knots in the equivalence class $\mathcal{K}$.

We now recall the definition of Gabai width. Order the critical values of $h|_K$ by $c_1 > \cdots > c_N$. Let $r_i \in (c_{i+1}, c_i)$ denote arbitrarily chosen regular values of $h|_K$ for $1 \leq i \leq N-1$. For any $y \in \mathbb{R}$, define $w(y) := |K \cap h^{-1}(y)|$. Define $w(K) := \sum_{i=1}^{N-1} w(r_i)$. The Gabai width of $\mathcal{K}$ is defined as $\min_{K' \in \mathcal{K}} w(K')$, where the minimum is taken over all Morse embeddings of knots in the equivalence class $\mathcal{K}$. If K' is such that $w(K') = w(\mathcal{K})$, then we say K' is in *thin position.*

Finally, we give our formal definition of planar width. For any knot diagram D in the yz–plane that is in general position with respect to h, let $K_D \subset \mathbb{R}^3$ be any knot in the ambient isotopy class $\mathcal{K}$ such that $p(K_D) = D$. We define the *planar width* of D, denoted by $w_p(D)$, as

$$w_p(D) := \min w(K_D),$$

where the minimum is taken over all planar isotopies of D.

3 The coloring rules

In this section, we define Wirtinger width via a combinatorial method for coloring knot diagrams. Let D be a knot diagram. Let $s(D) = \{s_1, \ldots, s_J\}$ denote the set of strands of D.

Definition 3.1 A *partial coloring* is a tuple (A, f), where A is a subset of $s(D)$ and $f: A \to Z$ is a function with $Z \subset \mathbb{Z}$.

Remark Set $A_0 := \varnothing$, $Z_0 := \varnothing$, and let f_0 be the empty function. Then (A_0, f_0) is a partial coloring. We fix (A_0, f_0) to denote this vacuous partial coloring.

We define two rules for extending partial colorings. Let (A_{t-1}, f_{t-1}) denote a partial coloring, where $t \in \mathbb{N}$ and $f: A_{t-1} \to Z_{t-1}$. See Figure 3 for examples of each rule.

Seed addition We say the partial coloring (A_t, f_t) is the result of a *seed addition* to (A_{t-1}, f_{t-1}), denoted by $(A_{t-1}, f_{t-1}) \to (A_t, f_t)$, if:

- $A_{t-1} \subset A_t$ and $A_t \setminus A_{t-1} = \{s_i\}$ for some strand $s_i \in s(D) \setminus A_{t-1}$.
- $Z_t := Z_{t-1} \cup \{t\}$.
- $f_t: A_t \to Z_t$ is defined by $f_t|_{A_{t-1}} = f_{t-1}$ and $f_t(s_i) := t$.

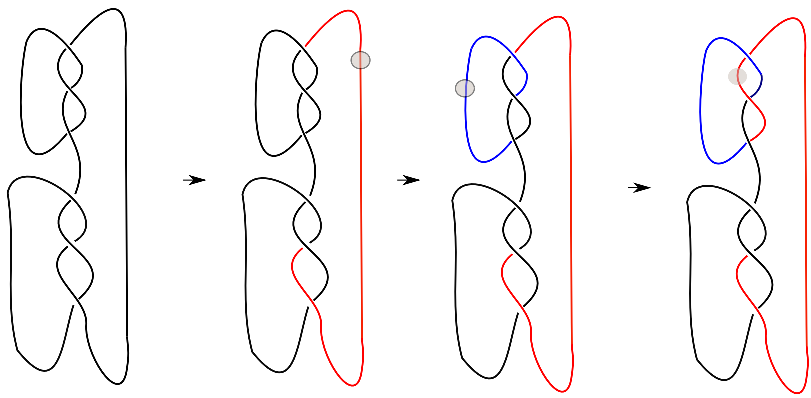

Figure 3: The first two transitions depict seed additions, the first adding the color red the second adding the color blue. The last transition depicts a coloring move extending the color red.

Coloring move We say (A_t, f_t) is the result of a *coloring move* on (A_{t-1}, f_{t-1}), denoted by $(A_{t-1}, f_{t-1}) \to (A_t, f_t)$, if:

- $A_{t-1} \subset A_t$ and $A_t \setminus A_{t-1} = \{s_q\}$ for some strand $s_q \in s(D) \setminus A_{t-1}$.
- s_q is adjacent to s_p at some crossing $x \in v(D)$ and $s_p \in A_{t-1}$.
- The over-strand s_v of x is an element of A_{t-1}.
- $Z_t := Z_{t-1}$.
- $f_t \colon A_t \to Z_t$ is defined by $f_t|_{A_{t-1}} := f_{t-1}$ and $f_t(s_q) := f_{t-1}(s_p)$.

There are two ways we refer to a coloring move. We say that s_q inherits its color from s_p, or that the coloring move was performed over the crossing x.

Remark We can always perform a seed addition to any uncolored strand. This allows us to use seed additions to extend the vacuous partial coloring (A_0, f_0).

Definition 3.2 If $(A_0, f_0) \to \cdots \to (A_t, f_t)$ is a sequence of coloring moves and seed additions on D, then we say the sequence is a *partial coloring sequence*. If we have a partial coloring sequence $(A_0, f_0) \to \cdots \to (A_J, f_J)$ such that $s(D) = A_J$, then we say the sequence is a *completed coloring sequence*. If t is an index of a partial coloring (A_t, f_t) in a specified coloring sequence, then we will refer to t as a *stage*.

Note that we can define a completed coloring sequence for any knot diagram since we can perform a seed addition to any strand.

Definition 3.3 If (A_t, f_t) is the result of a seed addition to (A_{t-1}, f_{t-1}) with $\{s_i\} = A_t \setminus A_{t-1}$, then we call s_i a *seed strand.*

Definition 3.4 Let $(A_0, f_0) \to \cdots \to (A_J, f_J)$ be a completed coloring sequence on the knot diagram D. Let $x \in v(D)$. Denote the over-strand of x by s_v and the under-strands of x by s_p and s_q. If there exists a stage t such that $s_p, s_q, s_v \in A_t$ and $f_t(s_p) \neq f_t(s_q)$, then we say x is a *multicolored crossing*. The smallest stage at which all previously stated conditions are satisfied will be referred to as the stage at which the crossing x becomes multicolored.

Completed coloring sequences allow us to extract geometric information from knot diagrams. To do this, we first record the order in which strands become colored, and crossings become multicolored.

Definition 3.5 Let $(A_0, f_0) \to \cdots \to (A_J, f_J)$ be a completed coloring sequence with multicolored crossing set $\mathcal{C}$. Let $\mathcal{C}_t$ denote the set of crossings that become multicolored at stage t. A Δ*–ordering* is an enumeration of the elements in $s(D) \cup \mathcal{C}$, $\Delta := (d_i)_{i=1}^{|s(D)|+|\mathcal{C}|}$, satisfying the following conditions:

1. For all $0 \leq t < u \leq J$, all elements colored (or multicolored) at stage t are listed before any element colored (or multicolored) at stage u.
2. For each stage $0 \leq t \leq J$, the element in $A_t \setminus A_{t-1}$ is listed, followed by all elements in $\mathcal{C}_t$ (if $\mathcal{C}_t \neq \varnothing$). That is, if at stage t a strand receives its color and a subset of crossings become multicolored, then we list the strand first, followed by all crossings that become multicolored at stage t.

Later, we use Δ–orderings to reconstruct an embedding of our knot in $\mathbb{R}^3$ from a colored knot diagram. Each seed strand will induce a single maximum and each multicolored crossing will induce a single minimum in our reconstructed embedding. The ordering of the critical points, by decreasing height with respect to h, is reflected in our Δ–ordering. We now show how to elevate this relationship into a calculation of Gabai width.

Definition 3.6 Let $(A_0, f_0) \to \cdots \to (A_J, f_J)$ be a completed coloring sequence. Let $\mathcal{S} \subseteq s(D)$, $\mathcal{C} \subseteq v(D)$ and Δ be the seed strands, multicolored crossings and Δ–ordering, respectively, of our completed coloring sequence. Let $\Delta' := (d_{i_j})_{j=1}^{|\mathcal{S}|+|\mathcal{C}|}$ denote the subsequence of Δ formed by restricting our Δ–ordering to the set $\mathcal{S} \cup \mathcal{C}$. We define the *attached sequence* $(a_i)_{i=0}^{|\Delta'|}$ to be the sequence created via the following rule:

- Set $a_0 := 0$.
- If $d_{i_j} \in \Delta'$ is a seed strand, then set $a_j := a_{j-1} + 2$.
- If $d_{i_j} \in \Delta'$ is a multicolored crossing, then set $a_j := a_{j-1} - 2$.

If the first t stages of the completed coloring involve $|S|$ total seed additions, and $|C|$ total crossings become multicolored by stage t, then we say the partial coloring sequence $(A_0, f_0) \to \cdots \to (A_t, f_t)$ induces the first $|S| + |C|$ terms of the attached sequence $(a_i)_{i=0}^{|\Delta'|}$.

Definition 3.7 Define $\mathbb{W}(D) := \min \sum_{i=0}^{N} a_i$, where the minimum is taken over all possible completed coloring sequences defined for the diagram D. Let $\mathbb{W}(\mathcal{K}) := \min \mathbb{W}(D)$, where the minimum is taken over all possible knot diagrams of knots in the isotopy class $\mathcal{K}$. We define $\mathbb{W}(\mathcal{K})$ to be the *Wirtinger width* of $\mathcal{K}$.

Remark The Δ–ordering resulting from a completed coloring sequence need not be unique. For example, if at some stage in a coloring sequence the strand s becomes colored and the crossings x_i and x_j both become multicolored, then both

$$\Delta_1 := \{\dots, s, x_i, x_j, \dots\} \quad \text{and} \quad \Delta_2 := \{\dots, s, x_j, x_i, \dots\}$$

are Δ–orderings resulting from the same coloring. In the ultimate calculation of $\mathbb{W}(D)$, such nuances do not matter as both Δ_1 and Δ_2 would induce the same attached sequence $(a_i)_{i=0}^{\Delta'}$. This is because, in each possible Δ–ordering, the crossings that become multicolored at the same stage must always be listed consecutively by the second condition in Definition 3.5.

In order to prove statements about Wirtinger width, one often needs to specify a Δ–ordering to work with. The following definition allows us to do this:

Definition 3.8 Let $\Delta = \{d_i\}_{i=1}^{|s(D)|+|\mathcal{C}|}$ be a Δ–ordering resulting from a completed coloring sequence on the knot diagram D. We define the *height function* $h_o\colon \Delta \to \mathbb{Z}$ associated to Δ by $h_o(d_t) := -t$.

The function h_o retrieves the negative of the position of d_t in the Δ–ordering. We introduce a negative sign to allow us to focus on maxima instead of minima in later constructions. The main use of h_o in later proofs will be to compare the relative positions of strands and multicolored crossings in a Δ–ordering. If d_i and d_j represent strands of a knot diagram, then the inequality $h_o(d_i) > h_o(d_j)$ should be interpreted as "d_i is colored before d_j".

Remark The name Wirtinger width comes from the fact, proved in [4], that the minimum number of seed additions necessary to obtain a completed coloring sequence on the knot diagram D is equal to the minimum number of meridional generators needed in a Wirtinger presentation of the knot group from a diagram.

4 Connections to the Wirtinger number

In this section, we prove some preliminary results that will be needed for our proof of Theorem 1.1. These results are the Wirtinger width analogues of [4, Proposition 2.2]. Let $s(D) = \{s_1, \dots, s_J\}$ denote the strands of the knot diagram D.

Definition 4.1 Let $A := \{s_1, \dots, s_n\}$ be a connected subset of $s(D)$, ordered by adjacency. Let $g\colon A \to \mathbb{Z}$. We say g has a *local maximum* at s_j if $n > 1$ and

$$g(s_j) > \begin{cases} \max\{g(s_{j-1}), g(s_{j+1})\} & \text{if } 1 < j < n, \\ g(s_2) & \text{if } j = 1, \\ g(s_{n-1}) & \text{if } j = n. \end{cases}$$

If $n = 1$, then g has a maximum at s_1.

The following is an equivalent reformulation of being *k–meridionally colorable*, and the main theorem, from [4]:

Definition 4.2 D is *k–meridionally colorable* if there exists a completed coloring sequence $(A_0, f_0) \to \cdots \to (A_J, f_J)$ containing only k seed additions.

Theorem 4.3 *Let $\mu(\mathcal{K})$ denote the minimal k such that there exists a knot diagram D of a knot in the ambient isotopy class $\mathcal{K}$ which is k–meridionally colorable. Recall $\beta(\mathcal{K})$ denotes the bridge number of $\mathcal{K}$. Then $\mu(\mathcal{K}) = \beta(\mathcal{K})$.*

Proposition 4.4 *Let $(A_0, f_0) \to \cdots \to (A_J, f_J)$ be a completed coloring sequence on a knot diagram D. Let $\Delta := (d_i)_{i=1}^M$ be a Δ–ordering on $s(D) \cup \mathcal{C}$ induced by the completed coloring sequence on D. Let $h_o\colon \Delta \to \mathbb{Z}$ be the height function on Δ defined by $h_o(d_t) := -t$. Let $x \in v(D)$ be a crossing with under-strands s_p and s_q and over-strand s_v. Let s_p and s_r be the strands adjacent to s_q.*

(1) *For all $u \in \{0, 1, \dots, J\}$ and $y \in f_u(A_u)$, $f_u^{-1}(y)$ is connected.*

(2) *For all $y \in f_J(A_J)$, h_o has a unique local maximum on $f_J^{-1}(y)$ when the set $f_J^{-1}(y)$ is ordered sequentially by adjacency. The local maximum is the unique seed strand contained in $f_J^{-1}(y)$.*

(3) *Suppose now D is a nontrivial knot diagram and $f_J(s_p) = f_J(s_q) = f_J(s_r) = y$. If k is such that $\{s_q\} = A_k \setminus A_{k-1}$, then we cannot have $\{s_p, s_r\} \subset A_{k-1}$.*

(4) *If D is a nontrivial knot diagram and $x \notin \mathcal{C}$, then $h_o(s_v) > \min\{h_o(s_p), h_o(s_q)\}$.*

(5) *If D is any knot diagram and $x \in \mathcal{C}$, then $h_o(x) < \min\{h_o(s_p), h_o(s_q), h_o(s_v)\}$.*

Proof (1) This result is a reformulation of [4, Proposition 2.2(1)] in our notation. We induct on the stage u. Recall $A_0 = \varnothing$ and f_0 is the empty function, so the claim is vacuously true for f_0.

Suppose for induction that $f_u^{-1}(y)$ is connected for all $u < t$ and $y \in f_u(A_u)$. We will show that $f_t^{-1}(y)$ is connected for all $y \in f_t(A_t)$. Say $\{s_i\} = A_t \setminus A_{t-1}$ and $f_t(s_i) = r$. We consider two cases.

First suppose (A_t, f_t) is the result of a seed addition to (A_{t-1}, f_{t-1}). By our definition of seed addition, $f_t^{-1}(r) = \{s_i\}$ and $f_t^{-1}(y) = f_{t-1}^{-1}(y)$ for all $y \in f_t(A_t) \setminus \{r\}$. Since $f_t^{-1}(r)$ is a singleton, it is connected. By our induction hypothesis, $f_{t-1}^{-1}(y)$ is connected for all $y \neq r$.

Now suppose (A_t, f_t) is the result of a coloring move on (A_{t-1}, f_{t-1}). By our definition of coloring move, $f_t^{-1}(r) = f_{t-1}^{-1}(r) \cup \{s_i\}$ and s_i must be adjacent to a strand in $f_{t-1}^{-1}(r)$. Our induction hypothesis implies $f_{t-1}^{-1}(r)$ is connected. Therefore, $f_t^{-1}(r)$ must also be connected. For all $y \in f_t(A_t) \setminus \{r\}$, we have $f_t^{-1}(y) = f_{t-1}^{-1}(y)$. Therefore, our induction hypothesis also implies $f_t^{-1}(y)$ is connected for all $y \in f_t(A_t)$. This completes the induction.

(2) This result is a reformulation of [4, Proposition 2.2(2)] in our notation. The assertion comes from the following observation. For every color $y \in f_J(A_J)$ used in the coloring of D, the set $f_J^{-1}(y)$ contains a single seed strand s_e, which is the first strand assigned the color y. All other strands $s_j \in f_J^{-1}(y)$ assigned the color y occur after s_e in the sequence Δ.

We induct on the stage u. By definition, A_1 is a singleton and $f_1 \colon A_1 \to \{1\}$. Thus h_o trivially attains a unique local maximum on the set $A_1 = f^{-1}(1)$, which contains only a seed strand.

Suppose for induction that, for all $u < t$ and all $y \in f_u(A_u)$, the seed strand of $f_u^{-1}(y)$ is the unique local maximum of h_o on the set $f_u^{-1}(y)$ when ordered sequentially by adjacency. We claim the same holds for f_t. Say $\{s_i\} = A_t \setminus A_{t-1}$ and $f_t(s_i) = r$. We consider two cases.

First suppose (A_t, f_t) is the result of a seed addition to (A_{t-1}, f_{t-1}). By our definition of seed addition, $f_t^{-1}(r) = \{s_i\}$, so h_o trivially attains a unique local maximum on this set. For all $y \in f_t(A_t) \setminus \{r\}$, we have $f_t^{-1}(y) = f_{t-1}^{-1}(y)$, so our claim follows from the induction hypothesis.

Now suppose (A_t, f_t) is the result of a coloring move on (A_{t-1}, f_{t-1}). Then there exists a strand $s_l \in A_{t-1}$ such that $f_t(s_l) = r$ and s_l is adjacent to s_i. By our definition of coloring move and h_o, since s_l is adjacent to s_i but colored before s_i, $h_o(s_i) < h_o(s_l)$. Thus s_i is not a local maximum in $f_t^{-1}(r)$. Since $f_t^{-1}(r) = f_{t-1}^{-1}(r) \cup \{s_i\}$ and $f_t^{-1}(y) = f_{t-1}^{-1}(y)$ for all $y \in f_t(A_t) \setminus \{r\}$, our claim follows from the induction hypothesis. This completes the induction.

(3) Colloquially, our assertion is that, if D is not a diagram of the unknot, then at no stage in the coloring process can we have an uncolored strand s_q adjacent to two strands s_p and s_r that were assigned the same color. Suppose for contradiction that $s_p, s_r \in A_{k-1}$. By assumption, $s_q \notin A_{k-1}$. By part (1) of this proposition, $f_{k-1}^{-1}(y)$ is connected. Since $f_J(s_p) = f_J(s_r)$, we have $\{s_p, s_r\} \subset f_{k-1}^{-1}(y)$. Since D is a knot diagram, the connectivity of $f_{k-1}^{-1}(y)$ and the inclusion $\{s_p, s_r\} \subset f_{k-1}^{-1}(y)$ implies $s(D) \setminus \{s_q\} = f_{k-1}^{-1}(y)$. Thus $s(D) = f_J^{-1}(y)$ and so our completed coloring sequence has a single seed strand. By Theorem 4.3, this implies D is a diagram of a knot with bridge number 1. But the unknot is the only knot with bridge number 1. This contradicts the nontriviality of D.

(4) Colloquially, the claim states that, if D is nontrivial and x is not multicolored, then the over-strand of x is colored before one of its under-strands. Hence, the x comes earlier in the sequence Δ than at least one of s_p or s_q.

Assume for contradiction that $h_o(s_v) < \min\{h_o(s_p), h_o(s_q)\}$. That is, the over-strand of x is colored after both under-strands s_p and s_q have been colored. Since D is a nontrivial knot diagram, the adjacent strands s_p and s_q are distinct. Without loss of generality, say s_p is colored before s_q. Let k be the stage that s_q receives its color, so $\{s_q\} = A_k \setminus A_{k-1}$.

Since $h_o(s_v) < \min\{h_o(s_p), h_o(s_q)\}$ and k is the stage at which s_q receives its color, s_v has not been colored by stage k. Therefore, no coloring move was performed over x in the completed coloring sequence.

Let s_p and s_r be the strands adjacent to s_q. By assumption, $x \notin \mathcal{C}$. That is, x is not multicolored, so s_p and s_q have been assigned the same color. Since s_p and s_q have been assigned the same color and are adjacent at x, but no coloring move was performed at x, s_q must have inherited its color from s_r via a coloring move. But s_p was colored before s_q. Therefore, $\{s_p, s_r\} \subset A_{k-1}$.

Since s_p and s_q were assigned the same color and s_q inherited its color from s_r, we have $f_J(s_p) = f_J(s_q) = f_J(s_r)$. But we have also showed $\{s_q\} = A_k \setminus A_{k-1}$ and $\{s_p, s_r\} \subset A_{k-1}$. Since D is a nontrivial knot diagram, we get the desired contradiction by part (3) of this proposition.

(5) The inequality is a reformulation of condition (2) in Definition 3.5 in terms of the height function h_o. In words, it states that, in the definition of sequence Δ, at each stage, the strand is listed before any crossings that become multicolored, as such a crossing does not become multicolored at stage t unless all of s_p, s_q, and s_v are in A_t. □

5 Coloring by height

In this section we describe a specific procedure for coloring diagrams of knots in thin position. It will be used to establish the inequality $\mathbb{W}(\mathcal{K}) \leq w(\mathcal{K})$. Our goal is to obtain a coloring sequence that induces a Δ–ordering which respects the ordering of the critical points of $h|_D$ by height.

For the rest of this section, let K be an embedding of the knot $\mathcal{K}$ in $\mathbb{R}^3$ that is in thin position with respect to h. Furthermore, let K be such that the knot diagram $D \subset \{yz\text{–plane}\}$, resulting from the projection p into the yz–plane is in general position with respect to h. Let $c_1 > c_2 > \cdots > c_N$ be the critical values of $h|_K$ ordered by decreasing height with respect to h. We also assume that $\mathcal{K}$ is not the ambient isotopy class of the unknot, so that D is a nontrivial diagram.

Definition 5.1 Let L be any knot diagram embedded in the yz–plane that is in general position with respect to h. Let $x \in v(L)$. Denote the under-strands of x by s_f and s_r. If $h|_{s_f}$ has a local maximum at x, then we say s_f is *the falling strand of* x. If $h|_{s_r}$ has a local minimum at x, then we say s_r is *the rising strand of* x.

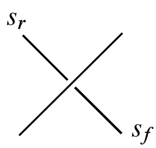

Figure 4: The rising strand and falling strand of the pictured crossing are denoted by s_f and s_r.

Recall that, for a strand s, we have defined the height of the strand to be $h(s) = \max_{y \in s} h(y)$. The assumption that D is in general position with respect to h means that all strands have distinct heights. This enables the following definition:

Definition 5.2 We say that we *color D by height* if we obtain a completed coloring sequence $(A_0, f_0) \to \cdots \to (A_J, f_J)$ by the following procedure:

Step 1 Write $s(D) = \{s_1, \ldots, s_{|s(D)|}\}$, where $h(s_1) > \cdots > h(s_{|s(D)|})$.

Step 2 Let (A_1, f_1) be the result of a seed addition to (A_0, f_0) such that $\{s_1\} = A_1 \setminus A_0$.

Step 3 Suppose we have a partial coloring sequence $(A_0, f_0) \to \cdots \to (A_{n-1}, f_{n-1})$ defined, where $A_{n-1} = \{s_1, \ldots, s_{n-1}\}$. Let x_i and x_j be the crossings incident to s_n. Say $h(x_i) < h(x_j)$. We consider two cases:

Case 1 Suppose $h|_{s_n}$ is maximized in $\operatorname{int}(s_n)$. Then we let (A_n, f_n) be the result of a seed addition to (A_{n-1}, f_{n-1}) such that $\{s_n\} = A_n \setminus A_{n-1}$.

Case 2 Suppose $h|_{s_n}$ is maximized in ∂s_n (so s_n is the falling strand of x_j). Then we let (A_n, f_n) be the result of a coloring move over x_j.

Remark When a coloring move is performed over a crossing x during the color by height process, colors must extend from the rising strand of x to the falling strand of x. Recall that, since D is assumed to be a nontrivial knot diagram, adjacent strands are distinct, so the rising and falling strands of x will always be distinct.

We first verify that knot diagrams in general position can always be colored by height.

Proposition 5.3 *If D is a knot diagram in general position with respect to h, then D can be colored by height.*

Proof We verify that each step of the color by height procedure can always be performed on D. Since D is in general position with respect to h, all strands have distinct heights. Thus, they can be ordered by decreasing height. By definition, we can always perform seed addition moves at any stage. What remains to be verified is that we can perform the coloring move stated in Step 3, Case 2 of Definition 5.2.

Let (A_n, f_n), s_n, x_i and x_j be as stated in Step 3, Case 2 of Definition 5.2. Let s_v and s_r denote the over-strand and rising strand of the crossing x_j, respectively. Since $h|_{s_n}$ is maximized in ∂s_n, we have $h(s_n) = h(x_j)$. By assumption, D is in general position with respect to h. Therefore, $h(s_n) = h(x_j) < \min\{h(s_v), h(s_r)\}$. Since the strands

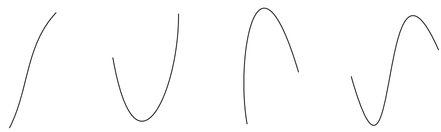

Figure 5: It will be shown that, since K is in thin position and D is in general position with respect to h, the strands of D can have at most two critical points. Moreover, if a strand has two critical points, then one must be a maximum and the other must be a minimum. This figure illustrates the stated possibilities.

were ordered by decreasing height, this implies $\{s_v, s_r\} \subset A_{n-1}$, so we can perform the desired coloring move. □

Our goal now is to show that, when we color D by height, we will get $\mathbb{W}(D) \leq w(K)$. The idea behind the upcoming technical results is that, since K is in thin position and the resulting diagram D is in general position with respect to h, the strands of D can be classified by how many critical points they contain. Figure 5 illustrates the classification, which will be used to show that the number of seed additions that occur when we color by height is equal to the number of maxima in K. Moreover, the number of multicolored crossings that occur is equal to the number of minima in K.

Lemma 5.4 *If $s \in s(D)$ and $r \in \mathbb{R}$ is a regular value of $h|_D$, then $|s \cap h^{-1}(r)| \leq 2$.*

Proof Suppose for contradiction we have a strand $s \in s(D)$ and a regular value $r \in \mathbb{R}$ of $h|_D$ such that $|s \cap h^{-1}(r)| \geq 3$. (See eg Figure 6.)

Recall that $c_1 > c_2 > \cdots > c_N$ are the critical values of $h|_K$, and say $r \in (c_{j+1}, c_j)$. Choose regular values $r_i \in (c_{i+1}, c_i)$ for $1 \leq i \leq N-1$ with $r_j = r$. Recall K is in thin position, so $w(K) = w(\mathcal{K})$. To obtain our desired contradiction, we will exhibit an isotopy on K to produce another embedding of $\mathcal{K}$ with strictly lower width.

Take three points a, b and c in $s \cap h^{-1}(r)$ that are consecutive in the strand s with respect to some orientation on s. Let $s_{a,b}$ denote the subarc of s in the yz–plane with boundary

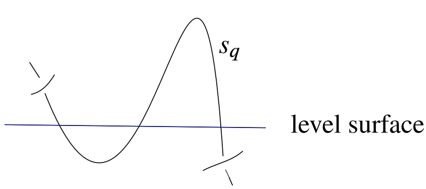

Figure 6: An example of a violation of Lemma 5.4.

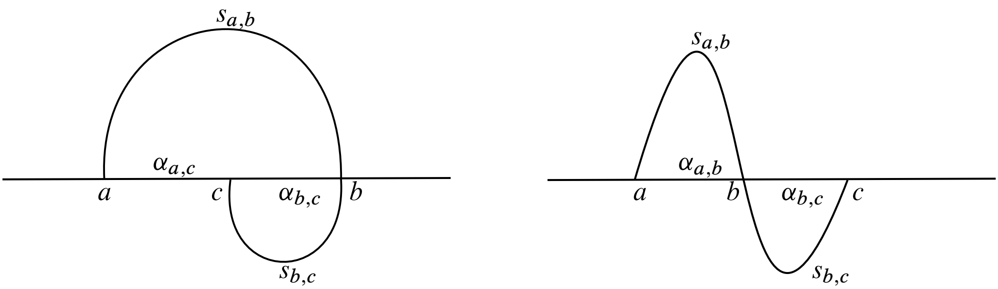

Figure 7: The setup for Cases 1 and 2 in the proof of Lemma 5.4 are on the left and right, respectively.

set $\{a, b\}$. Define $s_{a,c}$ and $s_{b,c}$ similarly. Let $\alpha_{a,b}$ be the arc in yz–plane $\cap\, h^{-1}(r)$ with boundary set $\{a, b\}$. Define $\alpha_{a,c}$ and $\alpha_{b,c}$ similarly.

Before describing the isotopy, we must consider cases based on the order of the points $\{a, b, c\}$ in yz–plane $\cap\, h^{-1}(r)$. The ordering is by the y–coordinates of the points. Up to symmetry, there are two cases to consider, as depicted in Figure 7.

Case 1 Suppose $a < c < b$. Let $D_{a,c}$ be the disk cobounded by $s_{a,c}$ and $\alpha_{a,c}$ in the yz–plane. We now define the steps of the isotopy. Let $\hat{s}_{a,c}$ be the arc component of $K \cap p^{-1}(s_{a,c})$.

Step 1 Perform an isotopy on K that fixes the y– and z–coordinates of all points on K, and arranges that $\hat{s}_{a,c} = p(\hat{s}_{a,c}) = s_{a,c}$ and all points in $K \setminus \hat{s}_{a,c}$ have negative x–coordinate. Note now $\hat{s}_{a,c}$ cobounds the disk $D_{a,c}$ with $\alpha_{a,c}$ in the yz–plane.

Step 2 Perform an isotopy on $\hat{s}_{a,c}$ that fixes a and c and pushes $\hat{s}_{a,c}$ across $D_{a,c}$ onto $\alpha_{a,c}$.

Step 3 After performing the isotopy, perturb the portion of K in a neighborhood of $\alpha_{a,c}$ so that $h|_K$ is Morse and has two fewer critical points than it had originally.

Let $s'_{a,c}$ and K' denote the image of $s_{a,c}$ and K, respectively, after the isotopy and perturbation. Let D' denote the diagram of K' given by projection into the yz–plane. Let $s'_{a,c}$ denote the image of $\hat{s}_{a,c}$ in D'.

Case 2 Suppose $a < b < c$. Then $s_{a,b}$ and $s_{b,c}$ cobound disks with $\alpha_{a,b}$ and $\alpha_{b,c}$, respectively, in the yz–plane. We obtain $s'_{a,c}$, K' and D' from a procedure analogous to that in Case 1. The only modification is that, in Step 2, we push across two disks instead of one.

We now claim $w(K') < w(K)$. By construction,

$$|s'_{a,c} \cap h^{-1}(r_j)| < |\hat{s}_{a,c} \cap h^{-1}(r_j)|.$$

Our procedure fixed the height of all points in K outside of a small neighborhood of $\hat{s}_{a,c}$ and did not introduce any new critical points. Therefore,

$$\sum_{i=1}^{N-1} |K' \cap h^{-1}(r_i)| < \sum_{i=1}^{N-1} |K \cap h^{-1}(r_i)| = w(\mathcal{K}).$$

The above inequality shows $w(K') < w(\mathcal{K})$. Since K was assumed to be in thin position, we get our desired contradiction. □

Proposition 5.5 *Let $(A_0, f_0) \to \cdots \to (A_J, f_J)$ be a completed coloring sequence obtained from coloring D by height.*

(1) *A seed addition is performed on the strand s if and only if $h|_s$ is maximized in the interior of s.*

(2) *Let x_i be a crossing with falling strand s_q, where x_i and x_j are the crossings incident to s_q. Then x_i is multicolored if and only if $h|_{s_q}$ is minimized in the interior of s_q and $h(x_i) < h(x_j)$.*

Proof (1) By Definition 5.2, a seed addition is performed on a strand if and only if that strand has a maximum in its interior.

(2) Let t be the stage at which s_q receives its color, so $\{s_q\} = A_t \setminus A_{t-1}$.

Suppose x_i is a multicolored crossing. Then $h|_{s_q}$ must be minimized in the interior of s_q, for otherwise, as s_q is the falling strand of x_i, it would be minimized at x_j. But, if s_q is the falling strand of x_i and $h|_{s_q}$ is minimized at x_j, then $h|_{s_q}$ would also have to be maximized at x_i, for otherwise we could find a regular value r such that $|s_q \cap h^{-1}(r)| \geq 3$, which would violate Lemma 5.4. In other words, s_q would be monotonic with respect to h. But this would mean (A_t, f_t) was the result of a coloring move on (A_{t-1}, f_{t-1}) over x_i, which is impossible because x_i is assumed to be multicolored.

In addition, if $h(x_i) > h(x_j)$, then s_q would have been colored via a seed addition, because the assumption that x_i is multicolored forbids any coloring move from being performed over x_i. The inequality $h(x_i) > h(x_j)$ would mean no coloring move was performed over x_j because we are coloring by height. By part (1) of this proposition,

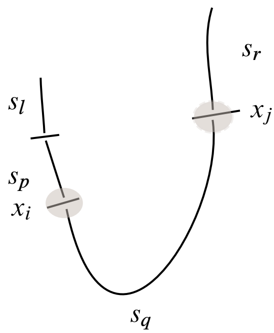

Figure 8: The setup for the proof of Proposition 5.5(2), where we want to show x_i is multicolored. It is assumed that s_q has a minimum in its interior and x_i is the lower incident crossing of s_q. The strands adjacent to s_q are s_p and s_r. The strands adjacent to s_p are s_l and s_q.

$h|_{s_q}$ would be maximized in the interior of s_q. But it was shown that $h|_{s_q}$ is also minimized in the interior of s_q. Since s_q is the falling strand of x_i and contains both a maximum and a minimum of $h|_{s_q}$ in its interior, the inequality $h(x_i) > h(x_j)$ would imply the existence of a regular value r such that $|s_q \cap h^{-1}(r)| \geq 3$, which would violate Lemma 5.4. We conclude $h(x_i) < h(x_j)$.

Conversely, suppose that $h|_{s_q}$ is minimized in the interior of s_q and $h(x_i) < h(x_j)$. We will show x_i is a multicolored crossing. Let s_p and s_r be the strands adjacent to s_q at the crossings x_i and x_j, respectively. Let s_l be the other strand adjacent to s_p. See Figure 8 for a diagram of this setup. Let u be the stage at which s_p is colored, so that $\{s_p\} = A_u \setminus A_{u-1}$.

Suppose for contradiction that x_i is not multicolored. Observe that, since $h(x_i) < h(x_j)$ and s_q is the falling strand of x_i, no coloring move could have been performed at x_i when we color D by height. We consider two cases.

Recall $\{s_q\} = A_t \setminus A_{t-1}$. First suppose (A_t, f_t) was the result of a seed addition to (A_{t-1}, f_{t-1}). By assumption, $x_i \notin \mathcal{C}$, so $f_J(s_p) = f_J(s_q)$. Thus s_p cannot also be a seed strand. Hence, s_p must have inherited its color from s_l because no coloring move could have been performed over x_i when we colored D by height. But this means $f_J(s_l) = f_J(s_p) = f_J(s_q)$ and $\{s_l, s_q\} \subseteq A_{u-1}$ must hold. This contradicts Proposition 4.4(3).

Now suppose (A_t, f_t) was the result of a coloring move on (A_{t-1}, f_{t-1}). No coloring move could have been performed over x_i when we colored D by height, so s_q must have inherited its color from s_r. But $x_i \notin \mathcal{C}$. Therefore, $f_J(s_p) = f_J(s_q) = f_J(s_r)$. If $u < t$ (that is, if s_p was colored before s_q), then $\{s_p, s_r\} \subset A_{t-1}$ and we have a contradiction to Proposition 4.4(3).

Now say $t < u$ (that is, s_q was colored before s_p). We still have $f_J(s_p) = f_J(s_q)$, so s_p cannot be a seed strand under the current assumptions. Thus s_p must have inherited its color from s_l since no coloring move could have been performed over x_i when we colored D by height. This forces $f_J(s_l) = f_J(s_p) = f_J(s_q)$ and $\{s_l, s_q\} \subset A_{u-1}$, contradicting Proposition 4.4(3).

We conclude x_i is multicolored. □

Recall that K is in thin position and D, which is the diagram of K obtained by projection into the yz–plane, has N critical points.

Corollary 5.6 *If $\mathcal{S}$ and $\mathcal{C}$ are the sets of seed strands and multicolored crossings resulting from a coloring of D by height, then $|\mathcal{S}| + |\mathcal{C}| = N$.*

Proof Proposition 5.5 implies that $\mathcal{S}$ and $\mathcal{C}$ are in bijective correspondence with the set of local maxima and the set of local minima of $h|_K$, respectively. This follows because K is assumed to be such that D is in general position with respect to h. □

Theorem 5.7 *If $\mathcal{K}$ is an ambient isotopy class of knots that does not contain the unknot, then $\mathbb{W}(\mathcal{K}) \le w(\mathcal{K})$.*

Proof Since D is a diagram of the knot K in $\mathcal{K}$, it suffices to show $\mathbb{W}(D) \le w(K)$. Let $(A_0, f_0) \to \cdots \to (A_J, f_J)$ be a completed coloring sequence on D obtained from coloring D by height. Let $(a_i)_{i=0}^N$ be the attached sequence of the coloring. We claim $\sum_{i=0}^N a_i \le w(K)$.

Note that Corollary 5.6 verifies that the number of critical points of K is equal to N, where the attached sequence $(a_i)_{i=0}^N$ resulting from coloring D by height contains $N+1$ terms. Let $r_n \in (c_{n+1}, c_n)$ denote a regular value of $h|_D$. It suffices to show $a_n \le w(r_n)$ for $1 \le n \le N$. Recall that we always have $a_0 = 0$ by definition. Fix one such n.

First we fix some notation. For all critical values c_i, let γ_i be the unique strand at which $h^{-1}(c_i)$ fails to intersect D transversely. Set $w(r_0) := 0$ for notational convenience. Write

$$a_n = \sum_{i=1}^n a_i - a_{i-1}, \quad w(r_n) = \sum_{i=1}^n w(r_i) - w(r_{i-1}),$$

so that our goal is to show

$$\sum_{i=1}^{n} a_i - a_{i-1} \leq \sum_{i=1}^{n} w(r_i) - w(r_{i-1}). \tag{1}$$

Observe that $a_i - a_{i-1} \in \{-2, 2\}$ and $w(r_i) - w(r_{i-1}) \in \{-2, 2\}$ for each i. Thus, it suffices to show that the number of positive terms in the left sum is bounded above by the number of positive terms in the right sum in equation (1).

Let t be the stage such that $s \in A_t$ if and only if $r_n < h(s)$. That is, a strand is colored by stage t if and only if its height is greater than r_n. We can acquire such a t because our completed coloring sequence was obtained from coloring D by height. To count the number of positive terms in the sums for equation (1), define

$$S_n := \{i \mid a_i - a_{i-1} = 2, 1 \leq i \leq n\}, \quad M_n := \{i \mid w(r_i) - w(r_{i-1}) = 2, 1 \leq i \leq n\}.$$

The value $|M_n|$ is the number of maxima above r_n. The value $|S_n|$ is related to the number of seed additions that have been performed by stage t. When coloring by height, it is possible that the lower incident crossing corresponding to a minimum below r_n becomes multicolored by stage t. Therefore, we cannot guarantee the equality of $|S_n|$ and $|M_n|$. However, we have the following claim, which suffices for our desired result:

Claim $$|S_n| \leq |M_n|.$$

Proof By Proposition 5.5(1), each strand containing a maximum with height above r_n must have been colored via a seed addition by stage t. Since D is in general position with respect to h, for all c_j above r_n corresponding to a minimum of a strand γ_j, the over- and under-strands of the lower incident crossing of γ_j have height greater then c_j, and hence r_n. Therefore, by Proposition 5.5(2), each minimum above r_n corresponds to a crossing that becomes multicolored by stage t. Since there are n critical points above r_n, we conclude that $(A_0, f_0) \to \cdots \to (A_t, f_t)$ induces at least the first $n+1$ terms $(a_i)_{i=0}^n$ in the attached sequence $(a_i)_{i=0}^N$.

By Definition 5.2, of coloring by height, $|M_n|$ is the number of seed additions in the partial coloring sequence $(A_0, f_0) \to \cdots \to (A_t, f_t)$. Since $(A_0, f_0) \to \cdots \to (A_t, f_t)$ induces at least the first $n+1$ terms $(a_i)_{i=0}^n$ in the attached sequence $(a_i)_{i=0}^N$, $|S_n|$ is bounded above by the number of seed additions in $(A_0, f_0) \to \cdots \to (A_t, f_t)$. Therefore, $|S_n| \leq |M_n|$, as desired. □

This claim shows that the number of positive terms in $\sum_{i=1}^n a_i - a_{i-1}$ is bounded above by the number of positive terms in $\sum_{i=1}^n w(r_i) - w(r_{i-1})$, which verifies the inequality in equation (1). □

6 Lifting a colored diagram

In this section we give a method for obtaining a Morse embedding of a knot from a colored knot diagram such that the ordering of the maxima and minima by height matches the Δ–ordering of seed strands and multicolored crossings. Then we use this method to show $\mathbb{W}(\mathcal{K}) \geq w(\mathcal{K})$.

For the rest of this section, let D be a diagram of a knot in the ambient isotopy class $\mathcal{K}$ such that $\mathbb{W}(D) = \mathbb{W}(\mathcal{K})$. Assume $\mathcal{K}$ is not the ambient isotopy class of the unknot, so that D is a nontrivial diagram. Let $(A_0, f_0) \to \cdots \to (A_J, f_J)$ be a completed coloring sequence on D with attached sequence $(a_i)_{i=0}^N$. Let $\mathcal{S}$, $\mathcal{C}$ and $\Delta = \{d_i\}_{i=1}^M$ denote the set of seed strands, multicolored crossings and the Δ–ordering on $s(D) \cup \mathcal{C}$ induced by our completed coloring sequence, respectively. Let $\Delta' := \{d_{i_j}\}_{j=1}^N$ be the subsequence of Δ formed by restricting our Δ–ordering to $\mathcal{S} \cup \mathcal{C}$. Let $h_o \colon \Delta \to \mathbb{Z}$ be the height function associated to Δ, defined by $h_o(d_t) := -t$.

In this section, we embed our diagram into the plane $z = -M - 1$. Recall that D is defined as a four-valent graph with labels at each vertex containing over/under information. The labels take the form of deleting parts of the edges in the graph corresponding to under-strands. We now want to view D as a disjoint union of arcs in the plane. To this end, for all $d_i \in \Delta$ representing a strand, let d_i^* be the strand d_i with neighborhoods of the boundary of d_i removed, as dictated by the labels on the vertices of D. For each $d_i \in \Delta$ representing a multicolored crossing, let $d_i^* := d_i$. This switch in perspective on knot diagrams, from a four-valent graph to a disjoint union of arcs in the plane, is necessary to adapt the proof of the main theorem in [4] to our situation.

Theorem 6.1 *There exists a knot $\hat{K}$ in the ambient isotopy class $\mathcal{K}$ embedded so that $h|_{\hat{K}}$ has N critical values $c_1 > c_2 > \cdots > c_N$. For all critical values, c_j is a maximum if and only if d_{i_j} is a seed strand. In addition, c_j is a minimum if and only if d_{i_j} is a multicolored crossing.*

Proof For all $d_t \in \Delta$, let $\hat{d}_t$ denote the copy of d_t^* embedded in the plane $z = h_o(d_t)$ so that the orthogonal projection of $\hat{d}_t$ onto the plane $z = -M - 1$ is d_t^*. Recall that the crossings of a knot diagram are by definition just points on the plane, so, if d_t is a crossing, then d_t^* is the point in the plane $z = h_0(d_t)$ projecting orthogonally onto d_t. We call $\hat{d}_t$ the *lift* of d_t.

In what follows, we show that the lifts $\hat{d}_t$ can be connected in such a way that the resulting knot has D as the diagram of its projection onto the plane $z = -M - 1$. Let

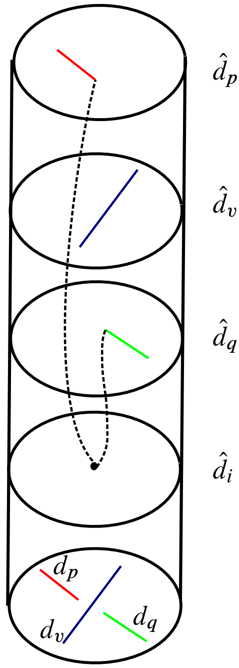

Figure 9: The construction of s_{pq} (the black dashed line) at the multicolored crossing d_i.

d_p and d_q be strands adjacent at the crossing x. Let d_v be the over-strand of x. Let $\epsilon > 0$ be such that the ball, denoted by $B(x,\epsilon)$, in the plane $z = -M-1$ has nonempty connected intersection with the strands d_p, d_q and d_v and empty intersection with all other strands. Then the cylinder $B(x,\epsilon)\times\mathbb{R}$ (where $\mathbb{R}$ denotes the z–direction) has nonempty connected intersection with $\hat{d}_p$, $\hat{d}_q$, and $\hat{d}_v$. The cylinder $B(x,\epsilon)\times\mathbb{R}$ is disjoint from all other lifts. At the crossing x, we embed an arc connecting the lifts $\hat{d}_p$ and $\hat{d}_q$, denoted by s_{pq}, via the following rule based on whether or not x is multicolored:

Connection case 1 Suppose x is a multicolored crossing. Say $x = d_i$. By Proposition 4.4(5), $h_o(d_i) < \min\{h_o(d_p), h_o(d_q), h_o(d_v)\}$. This means the plane $z = h_o(d_i)$ is below the planes containing the lifted under- and over-strands of x. Therefore, we can let s_{pq} be the union of two smooth monotone arcs connecting the endpoints of $\hat{d}_p$ and $\hat{d}_q$ in $B(x,\epsilon)\times\mathbb{R}$ to the point $\hat{d}_i$. This means $\hat{d}_i$ is the unique minimum of $h|_{s_{pq}}$. Moreover, we can choose s_{pq} such that it is contained in $B(x,\epsilon)\times\mathbb{R}$, disjoint from $\operatorname{int}(\hat{d}_v)$, and such that the orthogonal projection of

$$(\hat{d}_p \cup s_{pq} \cup \hat{d}_q \cup \hat{d}_v \cup \hat{d}_i) \cap (B(x,\epsilon)\times\mathbb{R})$$

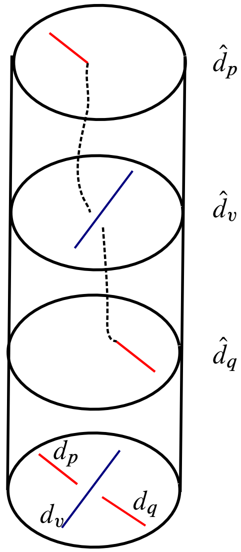

Figure 10: The construction of s_{pq} (the black dashed line) at crossings that are not multicolored.

onto the plane $z = -M - 1$ is $B(x,\epsilon) \cap D$, where s_{pq} projects to the deleted portions of the under-strands of x in D. See Figure 9 for a diagram of this construction.

Connection case 2 Suppose x is not a multicolored crossing. By Proposition 4.4(4), $h_o(s_v) > \min\{h_o(s_p), h_o(s_q)\}$. This means the plane $z = h_o(d_v)$ containing the lifted over-strand of x is above at least one of the planes containing the lifted under-strands of x. Therefore, we can let s_{pq} be a smooth monotone arc that connects the endpoints of $\hat{d}_p$ and $\hat{d}_q$ that intersect $B(x,\epsilon)\times\mathbb{R}$. Moreover, we can choose s_{pq} such that it is contained in $B(x,\epsilon)\times\mathbb{R}$, disjoint from $\mathrm{int}(\hat{d}_v)$, and such that the orthogonal projection of

$$(\hat{d}_p \cup s_{pq} \cup \hat{d}_q \cup \hat{d}_v) \cap (B(x,\epsilon)\times\mathbb{R})$$

onto the plane $z = -M - 1$ is $B(x,\epsilon) \cap D$, where s_{pq} projects to the deleted portions of the under-strand of x in D. See Figure 10 for a diagram of this construction.

Performing the above procedure at each crossing of D to connect all the lifts gives us a knot. Let $\widetilde{K} := \left\{\bigcup_t \hat{d}_t\right\} \cup \left\{\bigcup_{p,q} s_{pq}\right\}$. Since we respected the crossings under projection when defining each s_{pq}, D is a diagram of $\widetilde{K}$ under orthogonal projection onto the plane $z = -M - 1$. Hence, $\widetilde{K}$ is in the ambient isotopy class $\mathcal{K}$. However, $\widetilde{K}$ does not have the desired local extrema because the lifted strands are parallel to the xy–plane.

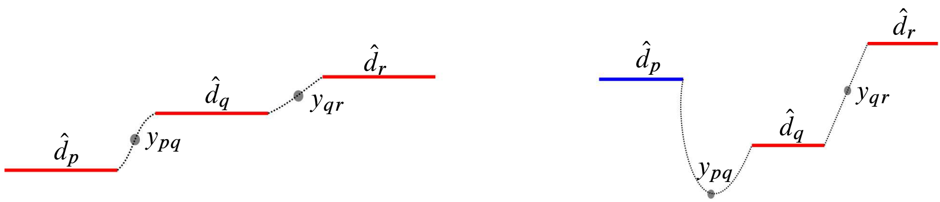

Figure 11: The setup of perturbation case 1, divided into subcases based on whether y_{pq} does (right) or does not (left) orthogonally project onto a multicolored crossing. Here d_q is not a seed strand. The idea is to perturb $[y_{pq}, y_{qr}]$, the subarc from y_{pq} to y_{qr} containing $\hat{d}_q$, into a monotonic arc with endpoints y_{pq} and y_{qr}.

We now show how to perturb the lifted strands contained in $\widetilde{K}$ so that we have the desired local extrema. For all s_{pq}, let y_{pq} denote the point in ∂s_{pq} that orthogonally projects to the corresponding crossing. Let d_p and d_r be the strands adjacent to d_q. Let $[y_{pq}, y_{qr}]$ denote the subarc of $s_{pq} \cup \hat{d}_q \cup s_{qr}$ from y_{pq} to y_{qr}. We consider cases based on whether d_q is a seed strand.

Perturbation case 1 Suppose d_q is not a seed strand. See Figure 11 for diagrams of what the lifts and $[y_{pq}, y_{qr}]$ could look like in this case. By Proposition 4.4(2), d_q is not the local maximum of h_o on $f_J^{-1}(f_J(d_q))$.

Claim $$\min\{y_{pq}, y_{qr}\} < h_o(d_q) < \max\{y_{pq}, y_{qr}\}.$$

Proof We consider cases based on whether the points y_{pq} and y_{qr} orthogonally project onto multicolored crossings. First suppose neither y_{pq} nor y_{qr} orthogonally projects onto multicolored crossings. Then d_p, d_q and d_r have all been assigned the same color. That is, $d_p, d_q, d_r \in f_J^{-1}(f_J(d_q))$. Since D is assumed to be nontrivial, if k denotes the stage at which d_q receives its color, then Proposition 4.4(3) asserts that $\{d_p, d_r\} \not\subset A_{k-1}$. That is, either d_p or d_r is uncolored at stage k. This implies $\min\{h_o(d_p), h_o(d_r)\} < h_o(d_q)$. But d_q is not the local maximum of h_o. Therefore, $h_o(d_q) < \max\{h_o(d_p), h_o(d_r)\}$. By the proof of connection case 2 of this theorem, the strands s_{pq} and s_{qr} are monotonic, so

$$\begin{aligned}\min\{h_o(d_p), h_o(d_r)\} < \min\{y_{pq}, y_{qr}\} < h_o(d_q) < \max\{y_{pq}, y_{qr}\} \\ < \max\{h_o(d_p), h_o(d_r)\},\end{aligned}$$

which gives the claim in this case.

Now say y_{pq} orthogonally projects onto a multicolored crossing. Then there exists some d_i such that $\hat{d}_i = y_{pq}$ and $h_o(d_i) = y_{pq}$. Proposition 4.4(5) implies

Figure 12: The setup of perturbation case 2, divided into subcases based on whether y_{pq} does (right) or does not (left) orthogonally project onto a multicolored crossing. Here d_q is a seed strand. The idea is to perturb $[y_{pq}, y_{qr}]$, the subarc from y_{pq} to y_{qr} containing $\hat{d}_q$, into an arc with a single maximum at the midpoint of $[y_{pq}, y_{qr}]$.

$y_{pq} = h_o(d_i) < h_o(d_q)$. Since $f_J^{-1}(f_J(d_q))$ is connected by Proposition 4.4(1) and d_q is not a seed strand, y_{qr} does not orthogonally project onto a multicolored crossing. Therefore, d_p must have inherited its color from d_r via a coloring move, so $h_o(d_q) < h_o(d_r)$. Since $h_o(d_q) < y_{qr} < h_o(d_r)$, we get the claim in this case. The argument for if y_{qr} orthogonally projects onto a multicolored crossing is similar. □

By the above claim, we can let the subarc $[y_{pq}, y_{qr}]'$ be an arbitrarily small perturbation of $[y_{pq}, y_{qr}]$ into a smooth monotonic arc, strictly increasing or decreasing as dictated by the values of $h_o(d_p)$ and $h_o(d_r)$. The perturbation is assumed to fix y_{pq}, y_{qr} and the projection to the plane $z = -M - 1$.

Perturbation case 2 Suppose d_q is a seed strand. See Figure 12 for diagrams of what the lifts and $[y_{pq}, y_{qr}]$ could look like in this case. By Proposition 4.4(2), d_q is the unique local maximum of h_o on $f_J^{-1}(f_J(d_q))$.

Claim $$\max\{y_{pq}, y_{qr}\} < h_o(d_q).$$

Proof If y_{pq} orthogonally projects onto a multicolored crossing, then $y_{pq} < h_o(d_q)$ by the same reasoning as in the proof of the claim for perturbation case 1. So suppose y_{pq} does not orthogonally project onto a multicolored crossing. Then d_p and d_q received the same color. That is, $d_p \in f_J^{-1}(f_J(d_q))$. Since d_q is the unique local maximum of h_o on $f_J^{-1}(f_J(d_q))$, the plane $z = h_o(d_p)$ containing $\hat{d}_p$ lies below the plane $z = h_o(d_q)$ containing $\hat{d}_q$. Hence, $y_{pq} < h_o(d_q)$. We have $y_{qr} < h_o(d_q)$ by similar reasoning. □

Let m_q be the midpoint of $\hat{d}_q$. By the previous claim, we can let $[y_{pq}, y_{qr}]'$ be an arbitrarily small perturbation of $[y_{pq}, y_{qr}]$ that fixes y_{pq}, m_q and y_{qr}. In addition, we arrange $[y_{pq}, y_{qr}]'$ so that $h|_{[y_{pq},y_{qr}]'}$ strictly increases from y_{pq} to m_q and strictly decreases from m_q to y_{qr} while fixing the projection to the plane $z = -M - 1$.

Perform a perturbation on the set of subarcs $\{[y_{pq}, y_{qr}]\}$ of $\widetilde{K}$ as dictated above. Let $\widehat{K}$ denote the resulting knot. Note $\widehat{K}$ is ambient isotopic to $\widetilde{K}$. Recall $\Delta' := \{d_{i_j}\}$ is the restriction of our Δ–ordering to $\mathcal{S} \cup \mathcal{C}$.

By perturbation case 2, each lifted seed strand $\hat{d}_{i_j}$ results in a maximum of $\widehat{K}$. The critical point corresponding to this maximum is the midpoint m_{i_j} of $\hat{d}_{i_j}$. Therefore, $\widehat{K}$ has a single maximum for every seed strand d_{i_j} with height $h_o(d_{i_j})$. By perturbation case 1, all other lifted strands become monotonic after perturbation.

By connection case 1, each multicolored crossing results in a minimum of $\widehat{K}$. The critical point corresponding to this minimum is the lifted multicolored crossing. Therefore, $\widehat{K}$ has a single minimum for every multicolored crossing d_{i_j} with height $h_o(d_{i_j})$.

Since the monotonicity of the subarcs of s_{pq} from y_{pq} to $\partial \hat{d}_q$ is preserved by our perturbation, $\widehat{K}$ has only $|\mathcal{S} \cup \mathcal{C}| = N$ local extrema. Ordering the critical values $c_1 > c_2 > \cdots > c_N$ of $h|_{\widehat{K}}$ by decreasing height for each j between 1 and N, c_j is a maximum if and only if d_{i_j} is a seed strand and c_j is a minimum if and only if d_{i_j} is a multicolored crossing, as desired. □

Corollary 6.2 *If $\mathcal{K}$ is an ambient isotopy class of knots that does not contain the unknot, then $\mathbb{W}(\mathcal{K}) \geq w(\mathcal{K})$*

Proof Let D be a diagram of a knot in the ambient isotopy class $\mathcal{K}$ such that $\mathbb{W}(D) = \mathbb{W}(\mathcal{K})$. Then there exists a completed coloring sequence on D with attached sequence $(a_i)_{i=0}^{N}$ such that $\sum_{i=0}^{N} a_i = \mathbb{W}(\mathcal{K})$. Let $\Delta' = \{d_{i_j}\}_{j=1}^{N}$ denote the Δ–ordering resulting from this coloring, restricted to the resulting seed strands and multicolored crossings. By Theorem 6.1, there exists a knot $\widehat{K}$ in the ambient isotopy class $\mathcal{K}$ with N local extrema that satisfy the following property: if $c_1 > c_2 > \cdots > c_N$ are the critical values of $h|_{\widehat{K}}$ ordered by decreasing height, then c_j is a maximum if and only if d_{i_j} is a seed strand and c_j is a minimum if an only if d_{i_j} is a multicolored crossing. This property ensures that, if $r_i \in (c_{i+1}, c_i)$ is a regular value of $h|_{\widehat{K}}$, then $a_i = |\widehat{K} \cap h^{-1}(r_i)|$. Therefore,

$$\mathbb{W}(\mathcal{K}) = \mathbb{W}(D) = w(\widehat{K}) \geq w(\mathcal{K}).$$ □

7 Proof of the main theorems

In this section we summarize previous results to prove our main theorems. Note that most results of Sections 5 and 6 do not apply to the unknot, so we must handle that case separately.

Before proving Theorem 1.1, we need one more technical lemma. Colloquially, it states that, at any stage of a coloring sequence, the number of multicolored crossings that have occurred is bounded above by the number of colors (seed strands) that have appeared.

Lemma 7.1 *Let $(A_0, f_0) \to \cdots \to (A_t, f_t)$ be a partial coloring sequence on the knot diagram D. Let $C := \{x_1, \dots, x_m\} \subset v(D)$ be the set of crossings of D that have become multicolored by stage t. Then $|C| \leq |f_t(A_t)|$.*

Proof We define a graph associated to the partial coloring sequence. Let $V := \{v_1, \dots, v_m\}$ be the vertex set, where we have one vertex for every multicolored crossing. Recall from Proposition 4.4(1) that, for all $y \in f_t(A_t)$, the set $f_t^{-1}(y)$ is connected. This means that, for all $y \in f_t(A_t)$, there are at most two multicolored crossings with under-strands assigned the color y. That is, the set $f_t^{-1}(y)$ contains the under-strands of at most two multicolored crossings. For each $y \in f_t(A_t)$ where $f_t^{-1}(y)$ contains the under-strands of two distinct multicolored crossings $x_i, x_j \in C$ (so $i \neq j$), let e_{ij} be an edge that joins the vertices v_i and v_j. For each $y \in f_t(A_t)$ where $f_t^{-1}(y)$ contains the under-strand of a single multicolored crossing $x_i \in C$, let e_{ii} be a loop based at the vertex v_i. That is, e_{ii} is an edge with both endpoints at v_i. Let E be the set of all edges obtained by this procedure.

Let $G := (V, E)$ denote the resulting graph. From the definition of G, it is clear that $|E| \leq |f_t(A_t)|$ and $|C| = |V|$. Let $\deg(v)$ denote the number of edges incident to v, where any loop based at v is counted twice. The *handshaking lemma*, which is a standard result in graph theory, states that

$$\sum_{v \in V} \deg(v) = 2|E|.$$

The under-strands of each multicolored crossing must be assigned different colors, and loops based at v are counted twice in the definition of $\deg(v)$, so $2 \leq \deg(v) \leq 4$ for all $v \in V$. Therefore,

$$2|V| \leq \sum_{v \in V} \deg(v).$$

But $|C| = |V|$ and $|E| \leq |f_t(A_t)|$. Therefore,

$$2|C| = 2|V| \leq \sum_{v \in V} \deg(v) = 2|E| \leq 2|f_t(A_t)|,$$

which gives the desired inequality. □

We now restate and prove our main theorems.

Theorem 1.1 *If $\mathcal{K}$ is an ambient isotopy class of knots, then $\mathbb{W}(\mathcal{K}) = w(\mathcal{K})$.*

Proof We begin with the case where $\mathcal{K}$ is not the ambient isotopy class of the unknot. Theorem 5.7 gives $\mathbb{W}(\mathcal{K}) \leq w(\mathcal{K})$. Corollary 6.2 gives $w(\mathcal{K}) \leq \mathbb{W}(\mathcal{K})$, so we get the desired equality.

Now suppose that $\mathcal{K}$ is the ambient isotopy class of the unknot. Then $w(\mathcal{K}) = 2$. We can obtain a completed coloring sequence on the standard diagram of the unknot, with no crossings, by performing a single seed addition. This shows $\mathbb{W}(\mathcal{K}) \leq 2$. We now verify that $\mathbb{W}(\mathcal{K}) \geq 2$. Let U be a diagram of the unknot. Let $(A_0, f_0) \to \cdots \to (A_J, f_J)$ be a completed coloring sequence on U with attached sequence $(a_i)_{i=0}^N$.

Let $a_n := \min\{a_i\}_{i=0}^N$. Then there exists a stage t such that the partial coloring sequence $(A_0, f_0) \to \cdots \to (A_t, f_t)$ induces the first n terms, $(a_i)_{i=0}^n$, in our attached sequence. Write

$$a_n = \sum_{i=1}^{n} a_i - a_{i-1}. \tag{2}$$

Define

$$S := \{i \mid a_i - a_{i-1} = 2,\, 1 \leq i \leq n\}, \quad C := \{i \mid a_i - a_{i-1} = -2,\, 1 \leq i \leq n\}.$$

The quantity $|S|$ is equal to the number of seed additions that have been performed by stage t. Thus, $|S| = |f_t(A_t)|$. The quantity $|C|$ is the number of crossings that have become multicolored by stage t, because $a_n = \min\{a_i\}_{i=0}^N$. By Lemma 7.1, $|C| \leq |f_t(A_t)| = |S|$. We have $a_i - a_{i-1} \in \{-2, 2\}$ for all i between 1 and n, so $|S|$ is also the number of positive terms in equation (2), and $|C|$ is also the number of negative terms in equation (2). Therefore, Lemma 7.1 implies that the number of negative terms is bounded above by the number of positive terms in equation (2). We conclude $a_n \geq 0$.

Since $a_n = \min\{a_i\}_{i=0}^N$, all terms in the attached sequence are nonnegative. Any completed coloring sequence on a knot diagram must start with a seed addition. Therefore, $a_0 = 0$ and $a_1 = 2$. Hence, our conclusion verifies that $\mathbb{W}(U) \geq 2$. But U was arbitrary, so $\mathbb{W}(\mathcal{K}) \geq 2$. Therefore, $\mathbb{W}(\mathcal{K}) = 2 = w(\mathcal{K})$. □

Theorem 1.2 *For any ambient isotopy class $\mathcal{K}$ of knots and any positive integer n, there exist infinitely many diagrams D of knots in $\mathcal{K}$ such that $\mathbb{W}(D) = w(\mathcal{K})$ but $w_p(D) \geq \mathbb{W}(D) + n$.*

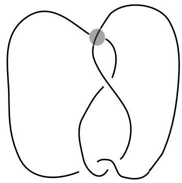

Figure 13: The diagram of the unknot U with a highlighted crossing.

Proof Let U be the diagram of the unknot depicted in Figure 13, contained in the yz–plane. Let E be the diagram obtained by performing a crossing change to the highlighted crossing in Figure 13. See Figure 14. Let $\mathcal{E}$ denote the ambient isotopy class of the figure 8 knot and K_E denote a knot in $\mathcal{E}$ such that $p(K_E) = E$. (Recall $p\colon \mathbb{R}^3 \to \mathbb{R}^2$ is the standard projection into the yz–plane.)

By Theorem 1.1, there exists a diagram D' of a knot $K_{D'}$ in $\mathcal{K}$ such that $\mathbb{W}(D') = w(\mathcal{K})$. Let

$$D = D' \# U \# \cdots \# U,$$

where there are m terms in the connected sum, and the connected sum is performed as shown in Figure 15.

We take the strand of D' on which we surger to form D to be a seed strand of a completed coloring sequence on D' which realizes the equality $\mathbb{W}(D') = w(\mathcal{K})$. After performing a seed addition to the strand of D labeled s in Figure 15, we can use coloring moves to extend the color to all other strands of D which correspond to components of U. Since D was formed by surgering the aforementioned seed strand of D', it is easy to see $\mathbb{W}(D) = \mathbb{W}(D') = w(\mathcal{K})$. These equalities are independent of m.

By performing a crossing change at each crossing of D highlighted in Figure 15, we get a diagram of the knot $K_{D'} \# K_E \# \cdots \# K_E$. See Figure 16.

Without loss of generality, we can perform an arbitrarily small perturbation on the knot $K_{D'} \# K_E \# \cdots \# K_E$, which descends to a planar isotopy on $D' \# E \# \cdots \# E$, such that

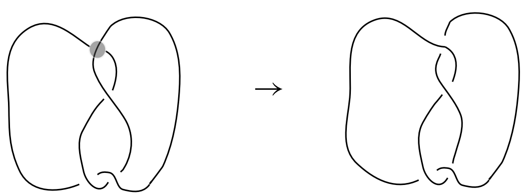

Figure 14: The crossing change performed on U (left) at the highlighted crossing to obtain E (right).

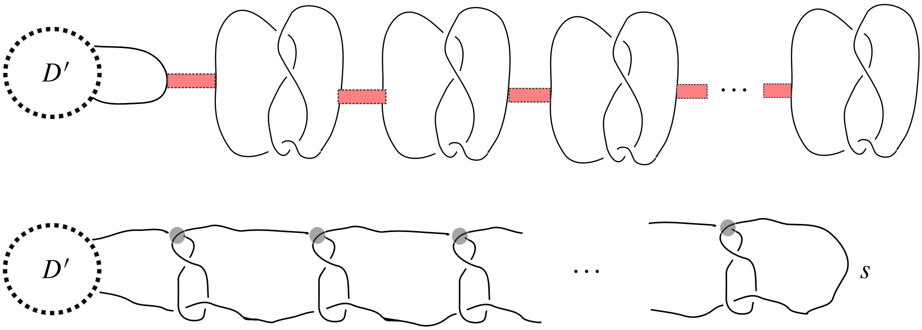

Figure 15: Top: D' with m copies of U. The rectangles along which we surger to form D are in red. Our calculations of width are independent of the orientations of the diagrams, so we assume each diagram is oriented to make the depicted connected sum well defined. Bottom: D with some crossings highlighted and a strand labeled s.

$h|_{K_{D'} \# K_E \# \cdots \# K_E}$ is Morse. Since planar width is unaffected by crossing changes, we get

$$w_p(D) = w_p(D' \# E \# \cdots \# E) \geq w(K_{D'} \# K_E \# \cdots \# K_E).$$

Recall Schubert's theorem on the additivity of bridge number (see [11, Theorem 1]), which states that, for any two ambient isotopy classes of knots $\mathcal{K}_1$ and $\mathcal{K}_2$,

$$\beta(\mathcal{K}_1 \# \mathcal{K}_2) = \beta(\mathcal{K}_1) + \beta(\mathcal{K}_2) - 1.$$

For any ambient isotopy class of knots, bridge number is a lower bound on Gabai width. By inductively applying Schubert's theorem with this observation, and the fact that $\beta(\mathcal{E}) = 2$ (recall $\mathcal{E}$ is the ambient isotopy class of the figure 8 knot), we get

$$w(K_{D'} \# K_E \# \cdots \# K_E) \geq \beta(K_{D'} \# K_E \# \cdots \# K_E) \geq \beta(\mathcal{K}) + m\beta(\mathcal{E}) - m = \beta(\mathcal{K}) + m,$$

where we got the second inequality because m is just the number of copies of E that we used in the connected sum to form D. Since the equalities $\mathbb{W}(D) = \mathbb{W}(D') = w(\mathcal{K})$ are independent of m, we can take m arbitrary large. Taking $m = \mathbb{W}(D) + n - \beta(\mathcal{K})$ in particular gives $w_p(D) \geq \mathbb{W}(D) + n$. $\square$

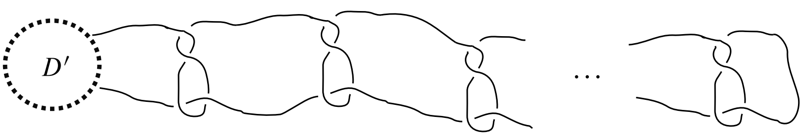

Figure 16: The resulting diagram of the knot $K_{D'} \# K_E \# \cdots \# K_E$ after performing a crossing change at each highlighted crossing in Figure 15.

8 Applications and further questions

In this section, we demonstrate how Theorem 1.1 can be used to write algorithms for calculating Gabai width. We will describe an algorithm we wrote that calculated the Gabai width of a large subset of tabulated knots from [9]. The data and code for our calculation are available at [10].

Our strategy was to modify the code in [14], which is the original algorithm for calculating Wirtinger number developed by the authors in [4], so that, given a Gauss code, it will output a completed coloring sequence for Wirtinger width. The modification is easy because the coloring moves for Wirtinger number and Wirtinger width are the same. Our modifications were motivated by the following lemma:

Lemma 8.1 *If $K \subset S^3$ is a* 4*–bridge prime knot in thin position*, *and thin position for K is not bridge position*, *then K has Gabai width* 28*.*

Proof Consider $\mathbb{R}^3$ now as in $S^3 = \mathbb{R}^3 \cup \{\infty\}$, with h the same height function as before. A thin position embedding of a 4–bridge knot must have four maxima and four minima. Since K is prime, $S^3 \setminus \eta(K)$ does not contain any essential 2–punctured spheres, where $\eta(K)$ is a tubular neighborhood of K. Wu [15] showed that the thinnest thin level of a knot that is in thin position but not bridge position is an essential surface in $S^3 \setminus \eta(K)$. Therefore, $|K \cap h^{-1}(r)| \neq 2$ for any regular value r of $h|_K$. For any regular value r of $h|_K$ at the thinnest level, the number of maxima above $h^{-1}(r)$ must be greater than or equal to the number of minima above $h^{-1}(r)$. These facts mean that the only possible orderings of the critical points of a prime 4–bridge knot are

$$M > M > M > M > m > m > m > m \quad \text{and} \quad M > M > M > m > M > m > m > m,$$

where the M's represent maxima and m's represent minima. The first ordering corresponds to a Gabai width of 32 while the second corresponds to a Gabai width of 28. However, the first ordering also corresponds to a bridge position embedding of a 4–bridge knot. Since bridge position of K is not thin position, the ordering of the critical points of K must be as in the second ordering above, so K has Gabai width 28. □

We focused on a subset of tabulated knots from [9] that are known to be prime with bridge number 4, with Gauss codes such that the code in [14] can actually detect bridge number 4. A prime knot with bridge number 4 such that thin position is bridge position must have Gabai width 32. Therefore, given a Gauss code representing a prime knot with bridge number 4, Lemma 8.1 implies that such a knot must have Gabai width 32 or 28. By Theorem 1.1, such a knot must have Wirtinger width 32 or 28. So every time we can find a completed coloring sequence on such a Gauss code giving Wirtinger width 28, we know the Gauss code represents a knot with Gabai width 28. Whenever our algorithm outputs an upper bound of 32 on the Wirtinger width for a given Gauss code, we unfortunately do not get any new information about Gabai width for the corresponding knot.

In light of these observations, we modified the code in [14] to search for a completed coloring sequence that starts with three seed additions, followed by coloring moves until we get a multicolored crossing, then finishes coloring the diagram with a seed addition that comes before three more multicolored crossings appear. Recall that seed strands correspond to maxima and the multicolored crossings correspond to minima, so such a coloring sequence corresponds to an embedding of the knot with Gabai width 28.

Our code implemented the above strategy and was able to verify that 54 756 tabulated knots have Gabai width 28, out of 86 981 knots that were tested. This is the first time a systematic calculation of Gabai width has been performed on this collection of Gauss codes. The appendix of [4] states that the code we modified in [14] for our algorithm runs in factorial time. Our modifications are such that our algorithm also runs in factorial time. However, our algorithm ran fast in practice since we had such specific information about the ordering of the seed strands and multicolored crossings in the completed coloring sequence we desired. In general, whenever bridge number is much less than the crossing number, the code in [14] runs fast in practice.

We remark that it was important to know the Gauss codes we were working on had diagrams such that the code in [14] can actually detect Wirtinger number 4 (and hence bridge number 4). In general, this does not always happen. In [3], the authors give examples of prime, reduced, alternating diagrams of a knot such that the Wirtinger number is strictly greater then the bridge number.

We briefly describe how we knew the bridge number. In [1], the authors give a method of establishing bridge number based on homomorphisms from the knot group to Coxeter groups. In ongoing work [2], the authors use computational methods to find homomorphisms as described in [1] to verify that each of the knots tested in our code [10] have bridge number 4.

Our implementation depended heavily on the Wirtinger number of a knot diagram. In general, the search for the minimum $\mathbb{W}(D)$ over all possible diagrams D is subtle. We took great advantage of the fact that the diagrams we worked on actually realized the Wirtinger number $\mu(D)$. In order to find a more robust implementation of our notions, it is important to understand how Wirtinger number and Wirtinger width interact. This leads to the following natural questions:

Question How can we determine whether or not a diagram D realizes the minimal $\mathbb{W}(D)$ without knowing beforehand that it realizes the minimal $\mu(D)$, the Wirtinger number?

Question If the knot diagram D realizes the Wirtinger number, then does D also realize the Wirtinger width?

One expects the answer to the second question to be no, since in [5] the authors exhibit a knot $\mathcal{K}$ such that the thin position embedding has more that $\beta(\mathcal{K})$ many maxima. However, finding a knot diagram which disproves our question seems difficult. An obvious first step is to check our knot data for a knot such that our algorithm outputs an upper bound of 32 for Gabai width, and try to show that the Gabai width of such a knot is actually 28.

References

[1] **S Baader**, **R Blair**, **A Kjuchukova**, *Coxeter groups and meridional rank of links*, Math. Ann. 379 (2021) 1533–1551 MR Zbl

[2] **R Blair**, **A Kjuchukova**, **N Morrison**, *Coxeter quotients of knot groups through* 16 *crossings*, preprint (2022) arXiv 2208.09032

[3] **R Blair**, **A Kjuchukova**, **M Ozawa**, *The incompatibility of crossing number and bridge number for knot diagrams*, Discrete Math. 342 (2019) 1966–1978 MR Zbl

[4] **R Blair**, **A Kjuchukova**, **R Velazquez**, **P Villanueva**, *Wirtinger systems of generators of knot groups*, Comm. Anal. Geom. 28 (2020) 243–262 MR Zbl

[5] **R Blair**, **M Tomova**, *Width is not additive*, Geom. Topol. 17 (2013) 93–156 MR Zbl

[6] **D Gabai**, *Foliations and the topology of* 3*–manifolds, III*, J. Differential Geom. 26 (1987) 479–536 MR Zbl

[7] **H Goda**, **M Scharlemann**, **A Thompson**, *Levelling an unknotting tunnel*, Geom. Topol. 4 (2000) 243–275 MR Zbl

[8] **C M Gordon**, **J Luecke**, *Knots are determined by their complements*, J. Amer. Math. Soc. 2 (1989) 371–415 MR Zbl

[9] **J Hoste**, **M Thistlethwaite**, **J Weeks**, *The first* 1 701 936 *knots*, Math. Intelligencer 20 (1998) 33–48 MR Zbl

[10] **R Lee**, *Wirtinger width*, Python code (2019) Available at `https://github.com/LeeRicky/Wirtinger-Width`

[11] **J Schultens**, *Additivity of bridge numbers of knots*, Math. Proc. Cambridge Philos. Soc. 135 (2003) 539–544 MR Zbl

[12] **A Thompson**, *Thin position and the recognition problem for* S^3, Math. Res. Lett. 1 (1994) 613–630 MR Zbl

[13] **A Thompson**, *Thin position and bridge number for knots in the* 3*–sphere*, Topology 36 (1997) 505–507 MR Zbl

[14] **P Villanueva**, *Wirtinger number*, Python code (2018) Available at `https://github.com/pommevilla/calc_wirt`

[15] **Y-Q Wu**, *Thin position and essential planar surfaces*, Proc. Amer. Math. Soc. 132 (2004) 3417–3421 MR Zbl

Department of Mathematics, UC Santa Barbara
Goleta, CA, United States

rickylee@ucsb.edu

Received: 10 December 2019 Revised: 10 April 2022

Algebraic & Geometric Topology

msp.org/agt

See inside back cover or msp.org/agt for submission instructions.

The subscription price for 2023 is US \$650/year for the electronic version, and \$940/year (+\$70, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Algebraic & Geometric Topology is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

Algebraic & Geometric Topology (ISSN 1472-2747 printed, 1472-2739 electronic) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840.

AGT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebraic & Geometric Topology

Volume 23 Issue 6 (pages 2415–2924) 2023