Unlinking, splitting, and some other NP-hard problems in knot theory

Dale Koenig*   Anastasiia Tsvietkova†

Abstract

We prove that certain problems naturally arising in knot theory are NP-hard or NP-complete. These are the problems of determining whether a link has an unlinking or splitting number k, finding a k-component unlink as a sublink, and finding a k-component alternating sublink.

The unlinking number of a link is the minimum number of crossing changes required for it to become an unlink, minimized over all diagrams of the link. Similarly, splitting number is the number of changes for the link to become split. Problems concerning unlinking and splitting numbers have a long history in topology and knot theory. At the same time, nothing is known about computability or complexity of these problems. A recent breakthrough by Lackenby suggests an algorithm to determine whether a hyperbolic link satisfying certain restrictions has unlinking or splitting number one [11], but no algorithm for the general case is yet known. We therefore provide the first bounds on the complexity of the general unlinking and splitting number problems. Our proof of NP-hardness for unlinking and splitting numbers is built upon our other proof, that of NP-hardness of UNLINK AS A SUBLINK problem. This problem is a special case (i.e. a restriction) of the SUBLINK PROBLEM, previously proven to be NP-hard [10]. More generally, for any property X of links one can also consider decision problems of the form "Given a diagram of a link L and a positive integer k, is there a k component sublink of L with the property X?" We show that this is also NP-hard if X is the property of being an alternating link.

1 Overview

Many problems that lie at the heart of classical knot theory can be formulated as decision problems, with an algorithm being a solution. Among them, for example, is the question of equivalence (up to isotopy) of two links given by their diagrams. This can be approached in many different ways, among which applying Reidemeister moves perhaps has the longest history. Other examples are the unknotting, unlinking and splitting number questions, i.e. arriving to a diagram of an unlink, an unknot, or a split diagram from some diagram by interchanging overpasses and underpasses in a certain number of crossings.

The complexities of these basic decision problems

in knot theory are not yet well-understood. Despite the lack of polynomial algorithms, few problems in knot theory are known to be NP-hard or NP-complete at all. Those few problems are the problem of determining a bound on the genus of a knot in a general 3-manifold [1] (the genus of a knot is by definition the least genus of any Seifert surface), a closely related problem of determining a bound for the Thurston complexity of a link in S^3 [10], and the problem of detecting a sublink isotopic to a given link [10]. There is also a proof that obtaining one diagram from another one of a link in a bounded number of Reidemeister moves is NP-hard in our preprint [7], and an improvement of this result to knots in [2]. The purpose of this paper to establish that a number of other natural problems in knot theory are NP-hard. The problems described here have an advantage that they can be formulated solely in terms of link diagrams. They are described below.

In section 2 we look at the UNLINK AS A SUBLINK problem, a special case of the SUBLINK PROBLEM defined and proven to be NP-hard by Lackenby [10]. We prove that the UNLINK AS A SUBLINK problem is NP-hard. Note that restricting a problem makes it easier, so the proof that the restricted problem is hard also implies that the general problem is hard. Thus, NP-hardness of the SUBLINK PROBLEM is a corollary of the NP-hardness of the problem that we consider. The techniques used in the proof are widely used throughout this paper. More generally, for any property X of links one can also consider decision problems of the form "Given a diagram of a link L and a positive integer k, is there a k-component sublink of L with the property X?" We show in section 5 that this is also NP-hard if X is the property of being an alternating link. We expect that many other problems of this form are NP-hard.

Very little is known about unknotting, unlinking and splitting numbers in general, without restricting to particular classes of links. It is possible that these invariants are not even computable. Moreover, it is not known whether there is an algorithm to detect whether a knot has unknotting number 1. A recent breakthrough by Lackenby suggests an algorithm to determine whether a hyperbolic link satisfying certain restrictions has unlinking or splitting number one [11], but no algorithm for the general case is yet known. We

provide the first lower bounds on the complexities of the general unlinking and splitting number problems in sections 3 and 4. We note that soon after our preprint [7] with these bounds appeared online, an alternative proof of NP-hardness of the unlinking number problem was given in [2].

Every section that follows is devoted to one of the above problems. We put the sections in the order that makes it easier for the reader to follow the proofs, since some of the arguments can be seen as refinements of the others.

For an overview of complexity theory including NP-hardness and NP-completeness, see, for example, [4]. For basic notions from knot theory, e.g. knot, link, knot diagram, isotopy, link group, Wirtinger presentation, split link, prime link, meridian, longitude, incompressible torus, satellite torus, we refer the reader to [12] or [15], and provide exact page and definition numbers further in the text. The previous status of many decision and complexity problems from knot theory is discussed in [8].

2 The unlink as a sublink problem

The definitions of a knot, an n-component link, an unknot, a knot diagram are given in Chapter 4.1 of [15]. An unlink with n components is an n-component link where each component is an unknot unlinked from the rest of components. An $(n-k)$-component sublink of a given n-component link L is obtained by deleting k components of L, where k is between 0 and $n-1$.

The SUBLINK PROBLEM asks "Given diagrams of two links, is there a sublink of the first that is isotopic to the second?" Lackenby showed that this problem is NP-hard using a Karp reduction (see Section 3.1.2 in [5]) from the HAMILTONIAN PATH problem [10]. Here we examine the UNLINK AS A SUBLINK problem, in which the second link is an unlink.

UNLINK AS A SUBLINK: given a diagram for a link L and a positive integer k, is there a k-component sublink of L that is an unlink?

THEOREM 2.1. UNLINK AS A SUBLINK *is NP-complete.*

Proof. We first prove NP-hardness by providing a Karp-reduction of the 3–SAT problem to the given problem. Suppose there are variables $x_1 \ldots x_n$ and clauses $c_1 \ldots c_m$, in a formula F, where each clause is of the form

$$z_{\alpha_i} \vee z_{\beta_i} \vee z_{\gamma_i}, \ 1 \le \alpha_i < \beta_i < \gamma_i \le n,$$

and z_k represents either x_k or $\neg x_k$. We construct a $(2n+m)$-component link L^F such that there exists an $(n+m)$-component unlink as a sublink if and only if

there is a variable assignment to $x_1 \ldots x_m$ satisfying all the clauses $c_1 \ldots c_m$.

To begin, create n Hopf links, one for each variable x_α. See Figure 1. Denote this link by L_0^F.
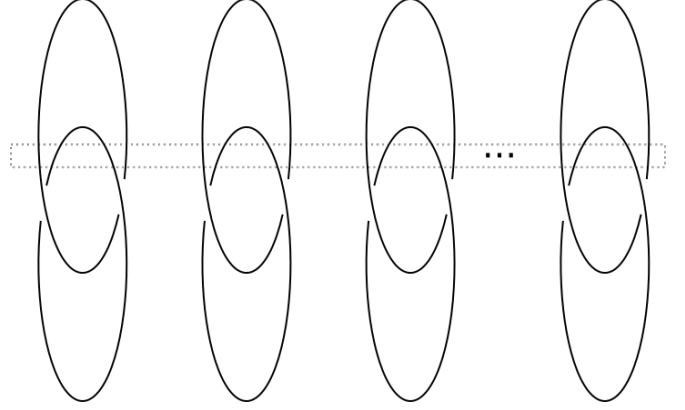

Figure 1: Hopf links, one for each variable x_i.

Label the top components of the Hopf links $x_1 \ldots x_n$ and the bottom components $\neg x_1 \ldots \neg x_n$. Zoom in on the grey dotted box in Figure 1 and divide it vertically into m shorter boxes as in Figure 2. For each clause c_i, we will add one more link component which we call a clause component lying in the ith grey dotted box. In this box we see two strands of the link L_0^F labelled x_i and two strands labelled $\neg x_\alpha$ for each α. Pick a single strand for each label, and leave the remaining strands unlabeled.
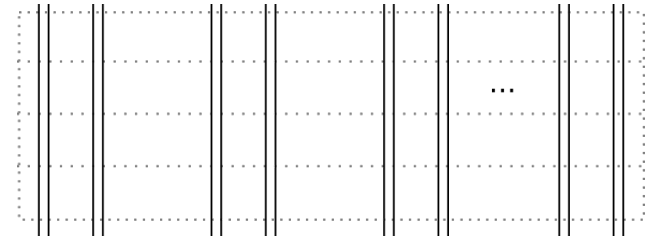

Figure 2: Dividing the box into m shorter boxes vertically, to put each clause component in.

The link group is a fundamental group of a link. The Wirtinger presentation allows to compute the group from a link diagram. For a quick algorithm, see Chapter 4.9 of [15]. Consider the Wirtinger presentation of the link group $\pi_1(S^3 - L_0^F)$. A word in the generators corresponds to a string of arcs/arrows passing under the labeled strands, passing left to right for x and $\neg x$, and right to left for x^{-1} and $\neg x^{-1}$. By placing such arrows from top to bottom in the diagram, and stringing them together using arcs which pass over the strands in the diagram, we can introduce a loop representing any word

which has no self-crossings. This construction is suggested in, for example, [14], and Figure 3 gives an example. In this way, draw a loop corresponding to the word $[z_{\alpha_i}, [z_{\beta_i}, z_{\gamma_i}]] = z_{\alpha_i}(z_{\beta_i} z_{\gamma_i} z_{\beta_i}^{-1} z_{\gamma_i}^{-1}) z_{\alpha_i}^{-1}(z_{\gamma_i} z_{\beta_i} z_{\gamma_i}^{-1} z_{\beta_i}^{-1})$ (an iterated commutator). Let the clause component corresponding to $c_i = z_{\alpha_i} \vee z_{\beta_i} \vee z_{\gamma_i}$ be this loop. For the formula F, we then obtain the link L^F by adding all clause components to L_0^F.

Each clause component links with three variable components in a Brunnian way (for Brunnian links, see p.67 of [14]). Indeed, an iterated commutator is trivial as a group element if any of the commuting elements is trivial, and the associated word can be reduced to the empty word by cancelling adjacent generators and inverses. Consider a clause $c_i = z_{\alpha_i} \vee z_{\beta_i} \vee z_{\gamma_i}$ and assume we have erased one of z_{α_i}, z_{β_i}, or z_{γ_i}, so the iterated commutator can be reduced to the empty word by such cancellations. With the way we have drawn the clause component, each cancellation in the commutator gives a sequence of type II Reidemeister moves simplifying the diagram (for Reidemesiter moves, see Chapter 4.2 of [15]). As an illustration, in our example in Figure 3, after removing one of $\neg x_1, \neg x_3$, or x_4, the clause component can be pulled away from the other components to become an unlinked unknot.

It remains only to show that the clause component links the three variable components in a nontrivial way. This follows from the fact that the fundamental group of the complement of the three variable components is free, and the clause component represents a nontrivial group element in that free group. It is therefore not freely homotopic to a trivial loop, since free homotopy corresponds to conjugation in the group, and so cannot be isotopic to a trivial loop.
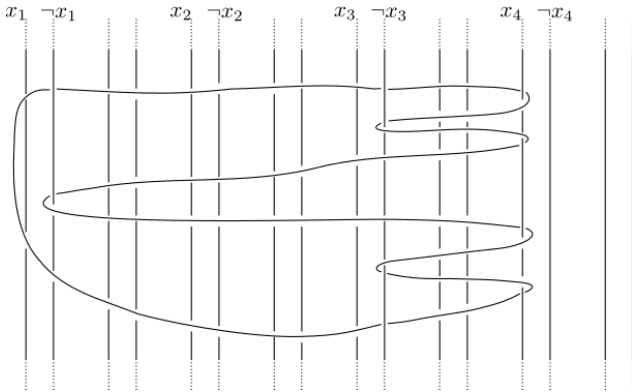

Figure 3: A link representing the commutator $[\neg x_1, [\neg x_3, x_4]]$, corresponding to the clause $\neg x_1 \vee \neg x_3 \vee x_4$.

Each clause component has precisely ten under-

crossings, and has at most $8n$ crossings between two such undercrossings. Indeed, between undercrossings it passes over each of the $4n$ strands in Figure 2 at most twice (in fact, at most once except when returning from the bottom back to the top). It follows that the number of crossings introduced for each new clause component is at most $64n + 10$. There are m clauses and $2n$ crossings corresponding to the n Hopf links, so the final diagram has at most $2n + m(64n + 10)$ crossings, which is polynomial in the size of the input formula F.

We now argue that the link L^F contains an $(n + m)$-component unlink as a sublink if and only if the corresponding formula F is satisfiable.

Suppose there is some set of variable assignments making F true. If the variable x_α is assigned to be TRUE, then delete the component labelled x_α from the link L^F, and if it is assigned FALSE, delete the component labelled $\neg x_\alpha$. For each clause component c of L^F that corresponds to a commutator $[z_{\alpha_i}, [z_{\beta_i}, z_{\gamma_i}]]$, one of z_{α_i}, z_{β_i}, or z_{γ_i} is TRUE, and the link component with the respective label was deleted. Hence the clause component c is unlinked from the rest of the diagram of L by the Brunnian property. After pulling away all the m clause components, we are left with the n unknotted components ("halves") of the Hopf links, so we have an $(n + m)$-component unlink.

Conversely, assume the link L^F contains an $(n+m)$-component unlink U as a sublink. We want to prove that F is satisfiable. The sublink U must contain at most one component from each of n Hopf links, since otherwise it contains pairs of components with nonzero linking number. U must therefore contain every clause component and exactly one component of each of the n Hopf links in order to have $n+m$ components total. This corresponds to the variable assignment determined by deleting the component labelled either by x_α or by $\neg x_\alpha$ for each α. If any clause of the formula F is not satisfied as a result of this assignment, the corresponding clause component will form a non-trivial Brunnian link with the three respective "halves" of Hopf links, and thus U is not an unlink, a contradiction.

We have therefore reduced the 3–SAT problem to the UNLINK AS A SUBLINK problem. The number of crossings of the diagram of L^F is bounded by a polynomial in the size of F. This implies we have a Karp reduction, so UNLINK AS A SUBLINK is NP-hard.

To see that the problem is NP, we observe that a polynomial length certificate consists of a choice of k link components followed by a sequence of Reidemeister moves (and resulting diagrams) converting the corresponding k-component link diagram to the trivial diagram of a k-component sublink. The number of Reidemeister moves needed to split the diagram is polynomial

as per [6, 9], and this splitting needs to be performed up to $k-1$ times. The number of moves needed to show each component is an unknot is polynomial by [6, 9] as well. Moreover, the size of the diagram at each step is bounded by a polynomial in the size of the original diagram, since the number of crossings is fewer than the original number of crossings plus the number of Reidemeister moves performed. Thus there is a polynomial length certificate showing that the original link contains a k-component unlink as a sublink, and UNLINK AS A SUBLINK is in NP. □

Note that NP-hardness of the SUBLINK PROBLEM as proven in [10] follows immediately as a corollary, since every instance of UNLINK AS A SUBLINK is also an instance of the SUBLINK PROBLEM.

3   Unlinking number

We next consider the problem of calculating the unlinking number of a link. The *unlinking number* of a link is the minimum number of crossing changes required to become an unlink, minimized over all diagrams of the link. The definition can also be formulated without reference to knot diagrams as is described below.

DEFINITION 3.1. *Suppose L is a link and α a simple arc with endpoints lying on L and disjoint from L otherwise. An* unlinking move *corresponding to the arc α is a homotopy of L that is the identity outside of a regular neighborhood $N(\alpha)$ of α, and within $N(\alpha)$ performs a move in which the link passes over itself once as in Figure 4.*
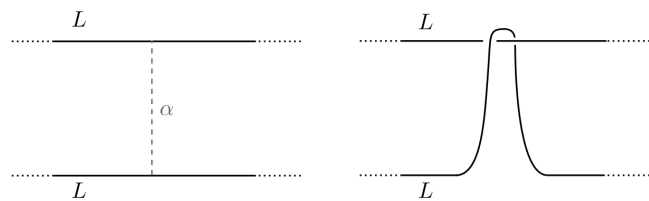
Figure 4: An example of an unlinking move replacing the fragment of a link on the right with the fragment of a link on the left.

An arc α alone does not uniquely determine a given unlinking move, as the link in Figure 4 (B) might twist around the arc. (An example with twisting is depicted in Figure 5.) Rather, there are infinitely many different unlinking moves that can be performed for any given

arc. However, for our purposes distinguishing these moves is not important; we only need a way to represent where an unlinking move occurs.
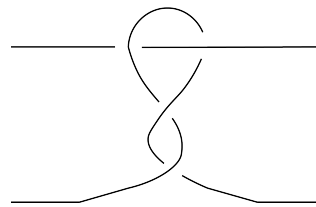
Figure 5: Link twisting around the arc α.

A finite collection of unlinking moves can be represented by a collection of disjoint arcs with disjoint regular neighborhoods. Then each unlinking move does not affect the neighborhoods of the arcs associated with the other unlinking moves. Therefore, all unlinking moves in such a collection can be performed simultaneously or in any order without changing the resulting link.

With this definition, the unlinking number of a link can be defined as the smallest number of unlinking moves needed to change the link into the unlink. Since unlinking moves are defined by paths in the link complement, they are independent of the link diagram. Hence one can take the smallest number of unlinking moves for one given link diagram, unlike with crossing changes, where the number of changes is minimized across all link diagrams.

Nonetheless, the two definitions of unlinking number are equivalent. Any crossing change is an unlinking move corresponding to the vertical arc connecting the two points of the crossing. For the other direction, given a collection of unlinking moves, first perform an isotopy of L to make all of the unlinking move arcs perpendicular to a plane Π. This can be done by first choosing a thin regular neighborhood of each arc, performing an isotopy that shrinks the arcs until they lie in small disjoint Euclidean balls, followed by an ambient isotopy that is the identity outside the set of balls to straighten the arcs. These isotopies will also move the link L. Once all arcs are straight and perpendicular to Π, project L to Π. After applying a small perturbation of L away from the arcs if necessary, the result of this projection will be a link diagram by general position arguments. Then each unlinking move's arc is a crossing arc in the diagram. However, it is not yet true that the unlinking moves correspond to the obvious crossing changes. If the unlinking move twists about its arc, we must additionally perform an isotopy to remove this twisting. Such an isotopy will introduce new crossings, and by slightly perturbing the diagram we can ensure again that all crossings are distinct. At the end of this process, each

unlinking move corresponds to a crossing change in the diagram as desired.

Thus, given a diagram of L and a set of n crossing changes turning L into the unlink, there is a set of n unlinking moves making L into an unlink. Conversely, if there is a set of n unlinking moves making L into an unlink then there is a diagram for which n crossing changes make L an unlink. Therefore, the minimum number of crossing changes needed to make L an unlink is the same as the minimum number of unlinking moves needed.

UNLINKING NUMBER: Given a link diagram and an integer n, does the link have unlinking number n?

It is not known whether UNLINKING NUMBER is computable, and it is not expected to be in NP [8]. Below, we prove that it is NP-hard (Theorem 3.1). The difficulty with finding a polynomial certificate lies in the fact that crossing changes required to efficiently get an unlink may not be visible in a given diagram. However one can consider the following restricted problem.

DIAGRAMMATIC UNLINKING NUMBER: Given a link diagram D and an integer n, can the link be made into an unlink by n crossing changes in the diagram D?

The restricted problem is NP-complete, which we will prove as a corollary to the proof of Theorem 3.1.

THEOREM 3.1. UNLINKING NUMBER *is NP-hard.*

We will prove this by modifying our construction for the UNLINK AS A SUBLINK PROBLEM. In L^F, we replace each variable component with its untwisted Whitehead double. Figure 6 demonstrates the replacement for a single Hopf link. We continue to refer to the Whitehead doubled variable components as variable components.
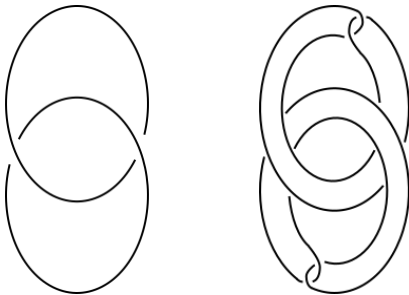

Figure 6: (A) A Hopf link and (B) the link obtained by replacing each component with a Whitehead double.

In the resulting link, any variable component can be made into an unknot unlinked from the remainder of the diagram using a single crossing change. In

fact, there are two such distinct crossing changes, but they determine only one unlinking move. This is the same unlinking move that is determined by any crossing change in a standard diagram of a Whitehead link. We call this unlinking move an *unclasping move.* Using n such unlinking moves followed by Reidemeister moves we can replace n variable components from the connected diagram that we have with an n-component unlink, disjoint from the rest of the diagram. Moreover, if there exists a set of n unlinking moves resulting in the unlink, then Lemma 3.1(a) asserts that there exists another set of n unlinking moves, also resulting in the unlink, such that any move involving a variable component is an unclasping move for that component. The second part of the lemma (Lemma 3.1(b)) will be used in Section 4.

Here and further, if L is a link and K is a component of L, denote by $L\backslash K$ the sublink of L consisting of components other than K.

LEMMA 3.1. *Suppose L is a link, and K is a component of L. Let $C = c_1, c_2, ..., c_n$ be a set of n unlinking moves of L, at least one of which involves the component K. Let c be an unlinking move of L that involves the component K and results in K becoming an unknot split from the other components of L. Then we can replace C with another sequence C' of unlinking moves beginning with c such that*

> *(a) C' has length equal or less than that of C.*
> *(b) If C results in L being an unlink, so does C'.*

> *(c) If P is some component of $L\backslash K$ and C results in P being split from $L\backslash P$, so does C'.*

Proof. Since the first unlinking move of C' is c, the component K is unknotted and unlinked from $L\backslash K$ after c is applied, and $L\backslash K$ is unchanged. Let D be the disk bounded by K after c is applied. We now construct C' iteratively by modifying moves from C.

For each move of C, if the move involves the component K, we do not include the move in C'. We will obtain the remaining moves of C' by modifying the moves of K that do not involve K. Since at least one move of C is assumed to involve K, this ensures that condition (a) is upheld. If a move c_i does not involve K, and the path determining c_i does not intersect the disk D, we add c_i to C' unmodified. If the path does pass through D, we isotope the path relative its boundary in the complement of $L\backslash K$ to make it disjoint from D. Figure 7 shows a link with a component K that is unlinked and unknotted. In the diagram on the left, the green arc corresponds to an unmodified unlinking move (a crossing change), and passes through the disk D bounded by K. In the diagram on the right, the

green arc is isotoped in the complement of $L\backslash K$ (but not in the complement of L) to correspond to a modified unlinking move that avoids D. This modification is always possible when K is unlinked from $L\backslash K$. Note that since we isotope in the complement of $L\backslash K$, the path must intersect K at some point in time during the isotopy. We add the modified unlinking move to C', and repeat for each unlinking move c_i of C
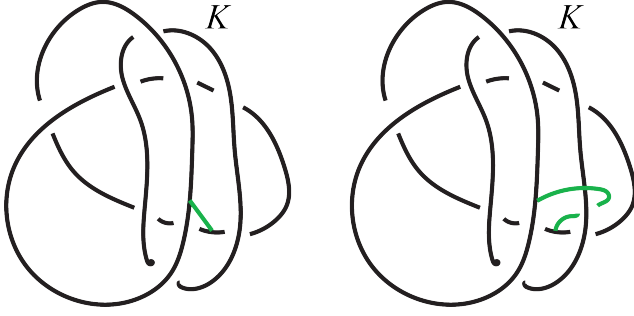

Figure 7: Modifying an unlinking move from C.

Let $C' = c, c'_1, ..., c'_m$, where for notational convenience we do not assume that c'_i is the result of modifying c_i (for example, if c_1 involved K then we would have skipped it, so c'_1 would be a result of modifying c_2). Since the paths determining the modified moves were constructed to be disjoint from D, the disk D is disjoint from $L\backslash K$ after performing C'. Therefore, after performing all unlinking moves of C', K must be an unlinked unknot. To evaluate conditions (b) and (c), we therefore only need to analyze how the moves of C' affect $L\backslash K$. If we restrict both C and C' to the subset of moves that do not involve K, then the sets of restricted moves act similarly on $L\backslash K$, since the paths determining the moves only differ by an ambient isotopy of $L\backslash K$. So (b) and (c) hold. □

We can now prove Theorem 3.1.

Proof. [Proof of Theorem 3.1] For a given 3–SAT problem F with n variables and m clauses, take the associated link L^F constructed for UNLINK AS A SUBLINK, and modify it by replacing every component with an untwisted Whitehead double. Denote this link by L^F_*. In this diagram, unclasping moves for each Whitehead doubled component are visible as crossing changes in the diagram. Each Whitehead doubled component bounds an obvious disk in the diagram after performing such an unclasping move.

Take a sequence C of n unlinking moves of L^F_* resulting in an unlink. A Whitehead doubled Hopf link cannot be unlinked without an unlinking move that only involves the components of that Whitehead

doubled Hopf link. Indeed, moves involving other components will not split the two components of a Whitehead doubled Hopf link from each other. Note however that once a component of Whitehead doubled Hopf link is a split from all other components, it is an unlinked unknot. Denote the Whitehead doubled Hopf sublinks by $W_1, W_2, ..., W_n$. Since there are n distinct Whitehead doubled Hopf sublinks of L^F_* and C consists of n unlinking moves, C must consist of one move involving each of W_i. Note that such sequence C is not unique: we just choose one such C.

We claim that by repeatedly applying Lemma 3.1 to L^F_* and one of the components of each W_i, we obtain another set of unlinking moves C_* such that the ith move of C_* is an unclasping move for either the x_i or $\neg x_i$ component of L^F_*. Indeed, first find an unlinking move of C involving one of the components K_1 of W_1. Applying Lemma 3.1 to L and K_1 gives a sequence of unlinking moves C' of equal or lesser length than C. The first unlinking move c'_1 of C' is an unclasping move for K_1. By Lemma 3.1(b), the new sequence of moves C' must still result in L being an unlink. Note that C' must have a move involving one of the components K_2 of W_2, since C had a move involving only W_2, and in the proof of Lemma 3.1 we did not exclude such moves. So we can again apply Lemma 3.1 to L and the component K_2 of W_2. Recall that we are working with a particular link diagram of L^F_*, described in the beginning. After performing an unclasping move c'_2 for K_2, the component K_2 bounds a disk visible in the link diagram, and the path determining c'_1 does not intersect this disk. Therefore, the application of Lemma 3.1 does not modify c'_1, and we obtain a sequence beginning with c'_2, c'_1. Repeating this process, we can get a sequence of unlinking moves such that, up to reordering, its ith unlinking move is an unclasping move for either the x_i or $\neg x_i$ component of L^F_*.

Deleting an unlinked unknot does not change whether a link is an unlink. So every unlinking move of C_* can be seen as deleting a component of L^F_*, as in the proof of NP-hardness for UNLINK AS A SUBLINK (Theorem 2.1). Hence from the proof of Theorem 2.1, a set of unlinking moves C_* that results in L^F_* becoming an unlink corresponds to a variable assignment satisfying the 3–SAT problem F.

The number of crossings in L^F_* is at most four times the number of crossings in L^F plus twice the number of components (the number of components is bounded by the number of crossings of L^F). In the proof of Theorem 2.1 we show that the number of crossings of L^F is polynomial in the size of F, therefore so is the number of crossings of L^F_*. Hence the reduction above is polynomial in the size of F. □

COROLLARY 1. DIAGRAMMATIC UNLINKING NUMBER *is NP-complete.*

Proof. In this proof, for every 3SAT formula F we need to construct a link diagram (rather than a link) that can be unlinked with n unlinking moves if and only if F is satisfiable. We will use the initial diagram of L_*^F from above. With this, the fact that DIAGRAMMATIC UNLINKING NUMBER is NP-hard follows from the proof of Theorem 3.1. In the proof, a set of n unlinking moves C assumed to take L_*^F to the unlink is modified to obtain a new set C_* that also changes L_*^F into an unlink. In particular, all moves of C_* consist of changing a crossing of a Whitehead double of an unknot. All such moves are actually realized in the diagram, so there exists a set of n crossing changes changing the diagram of L_*^F into a diagram of the unlink if and only if the formula F has a solution. Thus, the restricted problem is NP-hard.

For proving that the problem is in NP, note that L_*^F can be changed to a diagram of an unlink with n crossing changes, though the diagram might not yet be the standard unlink diagram consisting of n circles with no crossings. Hence, we need to take a polynomial length certificate consisting of a choice of these n crossings followed by one of the known polynomial length certificates for the unlink. This proves the problem is NP-complete. ◻

REMARK 2. *An alternative construction for the link L_*^F is to take the link L^F and replace only the Hopf link components with the Whitehead doubles.*

While we do not use this directly, we note that Theorem 1.2 from [3] shows that there is a unique way, up to a certain natural notion of equivalence, to unknot by a crossing change a Whitehead double of any knot but the figure-eight.

4 Splitting and Unknotting numbers

There are two other natural problems related to UN-LINKING NUMBER. The first is the unknotting number problem, which is the restriction of the unlinking number problem to single component links.

UNKNOTTING NUMBER: Given a knot diagram and an integer n, does the knot shown have unknotting number n?

The next is the problem of calculating the splitting number of a link. A split link is a link that has a (topological) 2-sphere in its complement separating one or more link components from the others.

DEFINITION 4.1. *The* splitting number *of a link L is the minimal number of crossing changes required to* make the link a split link, minimized over all diagrams.

As with the unlinking number, the splitting number can also be formulated without reference to diagrams by using unlinking moves. Specifically, it is the minimum number of unlinking moves needed to change a link into a split link. Variations of the splitting number have also been analyzed in the literature (see, for example, [8] for details). For example, while the splitting number only requires one separating 2-sphere in the link complement, the *total splitting number* asks for the minimum number of crossing changes needed to ensure that every component is split by a 2-sphere from the other components.

SPLITTING NUMBER: Given a link diagram and an integer n, does the link have splitting number n?

The UNKNOTTING NUMBER and SPLITTING NUMBER problems can also be rephrased as upper bound problems with "number at most n". Although rephrasing in this way changes the problems, it does not affect the property of being NP-hard. Indeed, n calls to an oracle solving the exact value problem is sufficient to solve the upper bound problem, while two calls to an oracle for the upper bound problem will solve the exact value problem. Since both the crossing number and splitting number are bounded by the crossing number and therefore the size of the input, both reductions are polynomial.

If one tries to extend the proof of Theorem 3.1 to the unknotting number, the main difficulty is that an unknotting number is known only for few specific classes of knots. Therefore, when constructing a new knot corresponding to a given 3–SAT instance, it is hard to show that there are no unexpected ways of unknotting it. Nonetheless, we conjecture:

CONJECTURE 3. UNKNOTTING NUMBER *is NP-hard.*

We will use methods similar to the ones in Theorem 3.1 to prove the following theorem.

THEOREM 4.1. SPLITTING NUMBER *is NP-hard.*

We start with two preliminary lemmas.

LEMMA 4.1. *Let $x_1, \cdots, x_n$ be generators for the free group F_n and let g be a nonempty product of nontrivial twice iterated commutators in $x_1, \cdots, x_n$, so $g = [x_{\alpha_1}, [x_{\beta_1}, x_{\gamma_1}]][x_{\alpha_2}, [x_{\beta_2}, x_{\gamma_2}]] \cdots [x_{\alpha_m}, [x_{\beta_m}, x_{\gamma_m}]]$, where $1 \leq \alpha_i < \beta_i < \gamma_i \leq n$ for all i. Then g is nontrivial in F_n.*

Proof. A word in the free group is trivial if and only if it can be reduced to the empty word by cancelling

adjacent $x_\alpha x_\alpha^{-1}$ or $x_\alpha^{-1} x_\alpha$ pairs. Within a given iterated commutator

$$[x_{\alpha_i}, [x_{\beta_i}, x_{\gamma_i}]] = x_{\alpha_i}(x_{\beta_i} x_{\gamma_i} x_{\beta_i}^{-1} x_{\gamma_i}^{-1}) x_{\alpha_i}^{-1}(x_{\gamma_i} x_{\beta_i} x_{\gamma_i}^{-1} x_{\beta_i}^{-1})$$

there are no such pairs since $1 \le \alpha_i < \beta_i < \gamma_i \le n$. Between adjacent commutators, we can have at most two pairs of elements that cancel. Indeed, each iterated commutator begins with the product $x_{\alpha_i} x_{\beta_i} x_{\gamma_i}$ of three distinct generators but ends with the expression $x_{\beta_i} x_{\gamma_i}^{-1} x_{\beta_i}^{-1}$ using only two distinct generators. It follows that whatever is between those two expressions in each iterated commutator will never be cancelled in the product g of iterated commutators. Hence the word g cannot be reduced to the empty word. □

LEMMA 4.2. *Let F_n and g be as in Lemma 4.1 and let S be a subcollection of the collection of generators $x_1, \cdots, x_n$. Let $N(S)$ be the smallest normal subgroup of F_n containing S, and G be the quotient $F_n/N(S)$. Then the image of g under the quotient map is trivial if and only if at least one element of each triple $x_{\alpha_i}, x_{\beta_i}, x_{\gamma_i}$ is contained in S.*

Proof. Denote the image of g under the quotient map by $\bar{g}$. First suppose that at least one element from each triple $x_{\alpha_i}, x_{\beta_i}, x_{\gamma_i}$ is contained in S. Then the image of each iterated commutator $[x_{\alpha_i}, [x_{\beta_i}, x_{\gamma_i}]]$ is trivial in G, and so $\bar{g}$ is trivial in G as the product of these commutators.

Now we prove that if some triple contains no element of S, then g must be nontrivial. For each i such that one of $x_{\alpha_i}, x_{\beta_i}, x_{\gamma_i}$ is in S, the image of $[x_{\alpha_i}, [x_{\beta_i}, x_{\gamma_i}]]$ is trivial in G, so we can remove $[x_{\alpha_i}, [x_{\beta_i}, x_{\gamma_i}]]$ from g without changing its image. Due to this, we may assume that no triple of g contains an element of S. Next, observe that G is a free group on $n - |S|$ generators. Indeed, define a map from F_n to $F_{n-|S|}$ by crossing out all appearances of elements of S and their inverses. It is simple to check that this is a surjective group homomorphism with kernel $N(S)$, and image G. The generators of G can be taken to be the set of x_α not contained in S. The word g is a sequence of iterated commutators in these generators, so we can apply Lemma 4.1 to see that the image of g is nontrivial. □

We now begin the construction.

Proof. [Proof of Theorem 4.1]

Let F be a 3-SAT formula with n variables and m clauses. We construct a link L^F that has splitting number n if and only if F has a solution. First, add in additional variables $x_{n+1} \ldots x_{2n}$ and clauses $x_i \vee \neg x_i \vee x_{n+i}$ and $x_i \vee \neg x_i \vee \neg x_{n+i}$ for $1 \le i \le n$. These new clauses and variables do not change the satisfiability of the 3-SAT formula. Indeed, all new clauses are satisfied by any TRUE/FALSE variable assignment. The new formula F' has $n' = 2n$ variables and $m' = m + 2n$ clauses.

To construct the link, begin with a $2n'$-component unlink, and label the components $x_1, \cdots, x_{n'}, \neg x_1, \cdots, \neg x_{n'}$. The complement of this unlink has fundamental group $F_{2n'}$, and Wirtinger generators

$$x_1, \cdots, x_{n'}, \neg x_1, \cdots, \neg x_{n'}$$

each corresponding to a loop going once around the strand of the link component with the respective label. Suppose F' has clauses $c_i = z_{\alpha_i} \vee z_{\beta_i} \vee z_{\gamma_i}$ for $1 \le i \le m'$ and $1 \le \alpha_i, \beta_i, \gamma_i \le n'$, where z_{α_i} is either x_{α_i} or $\neg x_{\alpha_i}$ appearing in c_i. Then we add a single component P corresponding to the word

$$[z_{\alpha_1}, [z_{\beta_1}, z_{\gamma_1}]][z_{\alpha_2}, [z_{\beta_2}, z_{\gamma_2}]] \cdots [z_{\alpha_{m'}}, [z_{\beta_{m'}}, z_{\gamma_{m'}}]]$$

in the fundamental group. Once the above product of the iterated commutators is written as a word in $x_\alpha^{\pm 1}$ and $(\neg x_\alpha)^{\pm 1}$, P is drawn using the same method as in the proof of Theorem 2.1 with respect to the link strands/components labeled by the generators. Note however that the strands do not belong to Hopf sublinks anymore as in Theorem 2.1, but rather to unlinked components. See Figure 8 for a simple example.
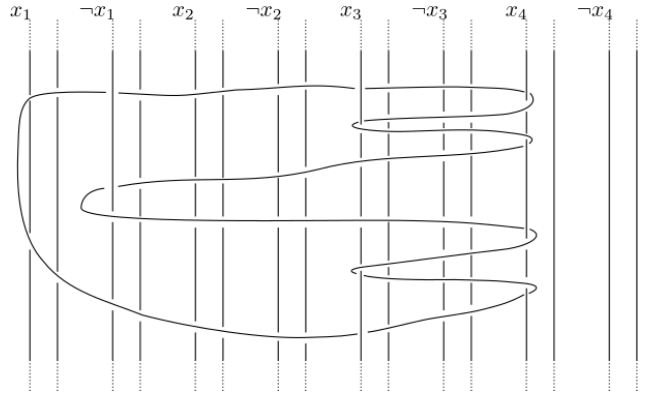

Figure 8: A link representing the commutator $[\neg x_1, [x_3, x_4]]$, corresponding to the clause $\neg x_1 \vee x_3 \vee x_4$.

We claim P is not split from the collection of other components. In fact, suppose S is some subset of $x_1, \cdots, x_{n'}, \neg x_1, \cdots, \neg x_{n'}$ and we delete the components with labels corresponding to elements of S to get a link L_S. Then by Lemma 4.2, P is split from the other components in L_S if and only if each triple $\alpha_i, \beta_i, \gamma_i$ contains some element of S.

The next step is to construct an intermediate link L^d by replacing each of the components $x_1, ..., x_{n'}, \neg x_1, ..., \neg x_{n'}$ (but not the combined clause component P) with an untwisted Whitehead double. We call these Whitehead doubles the *variable components* of L^d, in contrast to the clause component P, and continue to refer to them by the same labels. Note that as a Whitehead double, each variable component can be made into an unknot unlinked from the other components by a single crossing change (or unlinking move, in the diagram free terminology). We again call this crossing change an *unclasping move*. Thus, the splitting number of L^d is always one, although splitting the link in this way will in general not split P from the other components.

We will show that in L^d, the component P can be split from all other components in n unlinking moves if and only if there is a solution for F. More specifically, we claim that if P can be made split from the rest of L^d by n unlinking moves, then there is a set of n unclasping moves for variable components that also splits P from the rest of L^d, and that this set of n unclasping moves corresponds to an assignment satisfying F. Here unclasping a component labelled z_{β_k} means assigning TRUE to z_{β_k} in the variable assignment for F. Note that such an assignment also corresponds to a partial assignment of variables of F', but leaves n of the variables of F' unassigned. However, the construction of F' ensures that if such an assignment satisfies all clauses of F, then any arbitrary choice of TRUE/FALSE for the last n variables of F' will satisfy all clauses of F'. That is, every clause of F' is assured to be true by the assignment given to one of the first n variables. It follows that in the group element g corresponding to the component P, every iterated commutator will collapse to the identity in the group quotient determined by the partial variable assignment as per Lemmas 4.1 and 4.2.

Suppose P can be split from the other components of L^d in n unlinking moves. Denote this set of unlinking moves by C. We now apply Lemma 3.1 to every variable component of L^d such that there is a move in C involving this component. The chosen component serves as K in the lemma. Then using Lemma 3.1 (a) and (c) for such components we can replace any move in C involving a variable component with an unlinking move unclasping this component. After all possible such replacements in C, the new set of unlinking moves C' consists of some moves unclasping variable components of L^d and some moves that only involve P. We will see that the latter moves cannot appear in C'.

Note that in C', the moves involving variable components do not yet determine an assignment of variables of F in the right way. In particular, a variable might be assigned both TRUE and FALSE if there is an unclasping move for both x_i and for $\neg x_i$. A variable of F may also be not assigned anything at all. In what follows, the extra clauses that we added to F to obtain F' will be used to determine a variable assignment for F fully. Therefore our assignment for F will not come from the current sequence C'. Rather, we will modify C'.

Suppose a link component that corresponds to x_{n+i} or $\neg x_{n+i}$ is unclasped for $1 \leq i \leq n$ by an unlinking move v in C'. We might have one of two scenarios.

In the first scenario, there is no unclasping move for the link components labeled either x_i or $\neg x_i$. In this case v from C' can be replaced by either such unclasping move. Indeed, in Lemma 4.2, replacing x_{n+i} or $\neg x_{n+i}$ with x_i or $\neg x_i$ in the set S still yields a trivial image of g, since at least one element of every triple $x_{\alpha_i}, x_{\beta_i}, x_{\gamma_i}$ is still contained in S after the replacement. Therefore, such a replacement of v in C' yields the sequence that still splits P.

In the second scenario, there is already another unclasping move u in C' for either x_i or $\neg x_i$. This means that either (1) we applied Lemma 4 to x_i or $\neg x_i$ when we constructed C' from C, to obtain this unclasping move; or (2) the move u was one of the original linking moves in C; or (3) we already replaced another unclasping move of x_{n+i} or $\neg x_{n+i}$ in C' with u as in the first scenario. In all cases, the move c can be deleted from C' without changing whether F is satisfied.

Note that if we already have an unclasping move u for x_i or $\neg x_i$ in C', then u yields a part of a variable assignment. All clauses of F' involving the component x_{n+i} or $\neg x_{n+i}$ are already satisfied by assigning x_i to either TRUE or FALSE, so the TRUE/FALSE assignment of x_{n+i} corresponding to the unclasping move c is redundant. Indeed, no clauses of the original formula F involve x_{n+i} or $\neg x_{n+i}$, and the added clauses $x_i \vee \neg x_i \vee x_{n+i}$ and $x_i \vee \neg x_i \vee \neg x_{n+i}$ will both still be satisfied after such a replacement or deletion.

After all modifications, denote the resulting sequence by C''. As explained above, C'' splits P from the rest of L^d. At the end of this process, we can assume that all unclasping moves in C'' are of link components corresponding to x_i or $\neg x_i$ for $1 \leq i \leq n$. The clauses $x_i \vee \neg x_i \vee x_{n+i}$ then ensure that either x_i or $\neg x_i$ is unclasped for all $1 \leq i \leq n$. Since there are at most n unclaspings, x_i and $\neg x_i$ cannot both be unclasped for any i. The choice of which variable components to unclasp therefore corresponds to a TRUE/FALSE assignment to $x_1 \ldots x_n$, and these unclaspings will result in the clause component being split if and only each clause of F is satisfied.

It follows that there must be precisely n unclasping

moves in C''. Since C'' is a set of n unlinking moves, all unlinking moves must therefore be unclasping moves, and there can be no moves involving P. The set of unlinking moves C'' therefore corresponds to a variable assignment in F, and the clause component can be split from the other components in n unlinking moves if and only if there is a solution for F. Note again that the splitting number of L^d is always 1, since unclasping one side of any variable produces a split link.

To obtain the final link L_F, we add one more component to L^d. Consider the curve E drawn in grey in Figure 9. This curve has linking number 1 with every link component other than the combined clause component, and we can choose this curve to have no crossings with the combined clause component. Thicken the curve E, and let E' be a curve going around the boundary torus of the thickened curve $n + 1$ times longitudinally and once meridianally (p.106 of [15] defines meridian and longitude of a knot). Then E' links with every non-clause component $n + 1$ times.
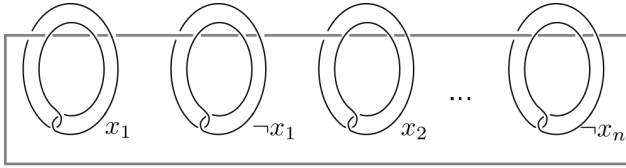

Figure 9: A curve E in thick grey linking every non-clause link component.

Let L_F be the union of the variable components, the combined clause component, and E'. The component E' cannot be unlinked from any of the variable components in fewer than $n + 1$ moves, so the only way to get a split link from L_F in at most n unlinking moves is to split the combined clause component from the remaining components. As argued above, this is possible if and only if F has a solution. Thus, L_F can be split in n unlinking moves if and only if there is a solution to the formula F.

The size of the diagram of L_F is polynomial in the size of the input for F. Indeed, before adding E' the number of crossings is less than the number of crossings involved in the construction for UNLINKING NUMBER, which was shown to be polynomial in the size of the input in Theorem 3.1. K has n self-crossings, and $8n(n+1)$ crossings with other components, so adding the component E' adds polynomially many more crossings. Therefore, we have a polynomial reduction and the theorem follows. $\square$

REMARK 4. *The method of combining all clause link components into a single component can also be applied to the* UNLINK AS A SUBLINK, UNLINKING NUMBER,

and ALTERNATING SUBLINK, *though we use different methods there. While combining all components allows the number of components in the link corresponding to a given 3-SAT formula to be reduced, the proofs become more complicated.*

5 Alternating Sublinks

We next consider another variation of the SUBLINK PROBLEM. Definition 4.4.4 in [15] explains what an alternating link is.

ALTERNATING SUBLINK: Given a diagram of a link L and a positive integer k, does L have a k-component sublink that is alternating?

Note that unlike the SUBLINK PROBLEM as proposed by Lackenby and the special case UNLINK AS A SUBLINK analyzed in Theorem 2.1, the problem above asks not for a specific sublink, but a sublink with a given property.

THEOREM 5.1. ALTERNATING SUBLINK *is NP-hard.*

We need the following lemma:

LEMMA 5.1. *Let L_s be a link that is not the unlink, and L'_s the result of replacing every component of L_s with a Whitehead double. Then L'_s is not alternating.*

Proof. The link L_s is split if and only if L'_s is split. A satellite or companion torus is defined on p.121 of [15]. A splitting sphere for L_s also splits L'_s, while a splitting sphere for L'_s can be isotoped to obtain a splitting sphere disjoint from all satellite tori, and is therefore a splitting sphere for L_s.

We will assume L'_s is alternating and arrive at a contradiction. Consider a reduced alternating diagram D of L'_s. If L'_s is split, L'_s can be seen to be split in D as was proved by Menasco [13]. Moreover, every connected piece of D is alternating, and at least one of such pieces, say P, is not an unknot, since L'_s is not an unlink. The piece P is alternating and has a connected diagram by assumption, and hence P is non-split as a link in S^3.

We claim that P is then prime as a link in S^3 (Definition 4.3 in [12] explains what a prime link is). Menasco's results [13] imply that a prime link complement cannot contain an incompressible non-boundary parallel torus, but there is a satellite torus in the complement $S^3 - P$, a contradiction.

Let us now prove the claim that P is prime. Suppose Σ is an embedded 2-sphere intersecting P in two points. Let K' be the link component of P that Σ intersects, and let T be the companion torus for K'. Let K be the component of L_s of which K' is a Whitehead double, so T can be thought of as the boundary of a regular neighborhood of K. See Figure 10.
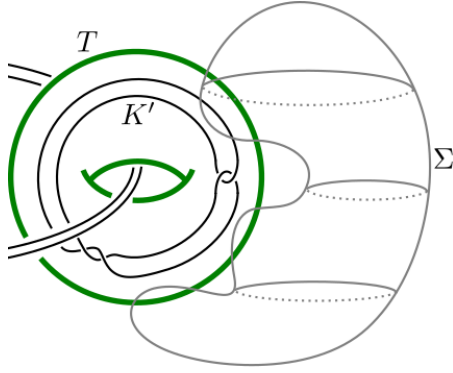
Figure 10: The link component K' (black) in it's companion torus T (thick green), and the sphere Σ (grey) intersecting K' in two points. Fragments of other components of L are also depicted in black.

Figure 11: The link component K' (black) in it's companion torus T (thick green), and the sphere Σ (grey) intersecting K' in two points. Fragments of other components of L are also depicted in black.
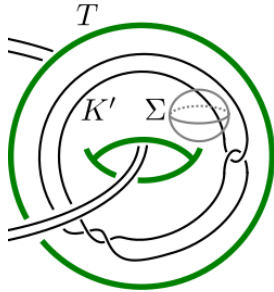
If the 2-sphere Σ intersects T in an essential meridianal curve, there must be at least two such curves of intersection. Therefore, the number of intersections of Σ and the component K' is at least 4, contradicting the definition of Σ. If Σ intersects T in an inessential curve, we can isotope Σ to remove this intersection, since P is not split. Since Σ intersects K' by assumption, Σ lies entirely on the same side of T as K', and any other components of P lie on the other side of T, since T is a companion torus for K, as shown in Figure 11. Therefore, the connected component of the manifold $S^3 - \Sigma$ that does not contain T intersects K' and no other components of P. We claim this intersection with K' is an unknotted arc.

Let V be the solid torus bounded by T and containing K'. Consider a homeomorphism h of V such that the resulting solid torus $h(V)$ is embedded and unknotted in S^3, and K' inside V' is an untwisted whitehead double of an unknot. The map h can be visualized, for exam-

ple, as cutting V along a meridianal disk, unknotting and untwisting it if necessary, and regluing along the same disk again. Then there is a 2-sphere $h(\Sigma)$ inside $h(V)$ intersecting $h(K')$ in two points. The sphere $h(\Sigma)$ splits S^3 into two 3-balls. Since $h(K')$ is the Whitehead double of an unknot, and is therefore unknotted itself, it must intersect both 3-balls in unknotted arcs. On the other hand, h did not change the interior of V, and therefore pulling back by h shows that Σ bounds an unknotted arc in V as well. We have shown that any 2-sphere intersecting Σ in two points bounds an unknotted arc on one side, and hence P is prime. □

Proof. [Proof of Theorem 5.1] Given a 3-SAT formula F, we use the same construction for a link L_*^F as for UNLINKING NUMBER in Theorem 3.1. Recall that each component of the link L_*^F is the Whitehead double of a component of the respective link L^F used for UNLINK AS A SUBLINK in the proof of Theorem 2.1. Let L_s be a k-component sublink of L^F, before replacing all components with their Whitehead doubles, and L_s' be the respective sublink of L_*^F after replacing the components with Whitehead doubles.

We proved in Theorems 2.1 and 3.1 that each of L_s and L_s' is the unlink if and only if it corresponds to a solution of the 3–SAT problem F. If L_s' does not correspond to a solution of F, the sublink L_s is not an unlink, and hence L_s' cannot be alternating by Lemma 5.1. On the other hand, if L_s' is an unlink, it is trivially alternating. Therefore, L_s' is alternating if and only if it corresponds to a solution of F. This Karp reduction is the same as the one used for UNLINKING NUMBER, so it is polynomial in the size of input for F. It follows that ALTERNATING SUBLINK is NP-hard. □

References

[1] Ian Agol, Joel Hass, and William Thurston, *The computational complexity of knot genus and spanning area*, Trans. Amer. Math. Soc. 358 (2006), no. 9, 3821–3850.

[2] Arnaud de Mesmay, Yo'av Rieck, Eric Sedgwick, Martin Tancer, *The unbearable hardness of unknotting*, Symposium on Computational Geometry 2019: 49:1–49:19.

[3] Alexander Coward and Mark Lackenby, *Unknotting genus one knots*, Comment. Math. Helv. 86 (2011), no. 2, 383–399.

[4] Michael R. Garey and David S. Johnson, *Computers and Intractability: A guide to the Theory of NP-Completeness*, W.H. Freeman, 1979.

[5] Oded Goldreich, *P, NP, and NP-completeness: The Basics of Computational Complexity*, Cambridge University Press, 2010.

[6] Joel Hass, Jeffrey Lagarias, and Nicholas Pippenger, *The computational com- plexity of knot and link problems*, JACM 46 (1999), no. 2, 185–211.

[7] Dale Koenig and Anastasiia Tsvietkova, *NP-hard problems naturally arising in knot theory*, preprint, arXiv:1809.10334.

[8] Mark Lackenby, *Elementary knot theory*, Lectures on Geometry (Clay Lecture Notes), Oxford University Press (2017).

[9] ——, *A polynomial upper bound on Reidemeister moves*, Ann. of Math. 182 (2015), no. 2, 491–564.

[10] ——, Some conditionally hard problems on links and 3-manifolds, Discrete Comput. Geom. 58 (2017), no. 3, 580–595.

[11] ——, *Links with splitting number one*, preprint, arXiv:1808.05495.

[12] W. B. Raymond Lickorish, An introduction to knot theory, p. 201, Springer- Verlag, 1997.

[13] William Menasco, *Closed incompressible surfaces in alternating knot and link complements*, Topology 23 (1984), no. 1, 37–44.

[14] Dale Rolfsen, *Knots and links*, American Mathematical Society, 1976.

[15] Jennifer Schultens, *Introduction to 3-manifolds*, American Mathematical Society, 2014.