PRIMALITY OF THETA-CURVES WITH PROPER RATIONAL TANGLE UNKNOTTING NUMBER ONE

KENNETH L. BAKER, DOROTHY BUCK, DANIELLE O'DONNOL, ALLISON H. MOORE, AND SCOTT TAYLOR

ABSTRACT. We show that if a composite θ-curve has (proper rational) unknotting number one, then it is the order 2 sum of a (proper rational) unknotting number one knot and a trivial θ-curve. We also prove similar results for 2-strand tangles and knotoids.

CONTENTS

1. INTRODUCTION

A spatial graph is a tame embedding of an abstract finite graph in the three-sphere S^3. A θ-curve is a spatial graph consisting of two vertices each joined by three edges. A spatial, planar graph is considered **unknotted** if it is contained in an embedded two-sphere, and the **unknotting number** u of a spatial graph is the minimum number of crossing changes required to make it unknotted. As with knots, the unknotting number is an extremely basic invariant of abstractly planar spatial graphs whose calculation can quickly become intractable. To illustrate the difficulty in understanding unknotting number, note that the additivity of unknotting number under connected sum is open for both knots and spatial graphs. However, Scharlemann showed that all unknotting number one knots are prime [Sch85]. We prove a version of Scharlemann's theorem for θ-curves. As we explain in Section 1.2, we were motivated by certain biological considerations.

Although, initially one might expect the version for θ-curves to follow easily from Scharlemann's theorem or the machinery of double-branched covers, the situation is subtler. In particular, for knots and θ-curves, there are two natural notions of connected sum: order 2 connected sum and order 3 connected sum, both depicted in Figure 1. In particular, observe that a composite θ-curve which is the order 2 connected sum of a trivial θ-curve with an unknotting number one knot, itself has unknotting number one. It turns out,

FIGURE 1. Left: an order 2 connected sum of a theta curve and a knot. Right: an order 3 connected sum of two theta curves.
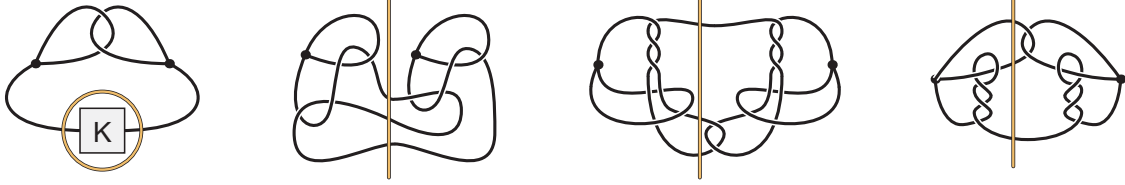

FIGURE 2. Four examples of composite handcuff curves with unknotting number one. These directly generalize to examples of unknotting proper rational tangle replacements that are not crossing changes.

however, that is the only possible exception to the claim that unknotting number one θ-curves are prime. We prove:

Theorem 1.1. *Let θ be a θ–curve. If $u(\theta) = 1$ then either*

- *θ is prime with respect to both order 2 connected sum $\#_2$ and order 3 connected sum $\#_3$, or*
- *$\theta = \theta_0 \#_2 K$ where θ_0 is the trivial θ–curve and K is a prime knot with $u(K) = 1$.*

In the latter case, up to isotopy, the unknotting crossing change occurs on K.

Theorem 1.1 follows from a similar theorem for 2-strand tangles. Deferring the definition of prime tangle to Section 2.3, we show:

Proposition 4.6. *Let T be a marked tangle sum of tangles, neither of which is integral. Then T admits no crossing change to an integral tangle.*

The arguments for Proposition 4.6 mostly adapt to using the more general proper rational tangle replacements (see Section 2.2) in place of crossing changes, though a different result from [GL06] is employed.

Proposition 4.9. *Let T be a marked tangle sum of tangles, neither of which is integral. Then T admits no proper rational tangle replacement to an integral tangle.*

Consequently, we obtain a generalization of Theorem 1.1 using proper rational tangle replacements. The **proper rational unknotting number** u_{prop} of a spatial graph is the minimum number of proper rational tangle replacements required to make it unknotted, see Section 2.2 and [ILM21].

Theorem 1.2. *Let θ be a θ–curve. If $u_{\mathrm{prop}}(\theta) = 1$ then either*

- *θ is prime with respect to both order 2 connected sum $\#_2$ and order 3 connected sum $\#_3$, or*
- *$\theta = \theta_0 \#_2 K$ where θ_0 is the trivial θ–curve and K is a prime knot with $u_{\mathrm{prop}}(K) = 1$.*

In the latter case, up to isotopy, the unknotting proper rational tangle replacement occurs on K.

Ultimately, Theorem 1.1 is a special case of Theorem 1.2 since crossing changes are proper rational tangle replacements. All of the arguments involving crossing changes readily adapt to proper rational tangle replacements with the exception of Proposition 4.6 which needs its generalization Proposition 4.9.

Proof of Theorem 1.2 (and Theorem 1.1). Suppose θ is a non-prime θ–curve with $u_{\mathrm{prop}}(\theta) = 1$.

The argument splits into two cases.

Case 1. Assume θ is locally knotted. Then Proposition 3.1 gives the desired conclusion. It shows that θ is the connected sum of a non-trivial knot K with $u_{\mathrm{prop}}(K) = 1$ and a trivial θ–curve. Moreover, any proper rational unknotting arc for θ is isotopic into the ball containing the local knot.

Case 2. Assume θ is a non-trivial vertex sum that is not locally knotted. Then Proposition 3.3 shows that this contradicts the assumption that $u_{\mathrm{prop}}(\theta) = 1$. □

2

Remark 1.3. In the statements of Theorems 1.1 and 1.2, the restriction to θ-curves is crucial. Figure 2 shows several examples of order 2 and order 3 composite **handcuff curves** of unknotting number one. In the first one, observe that K can be any knot, and so an unknotting number one handcuff curve can have arbitrarily many summands. The second and third are vertex sums of two knotted handcuff curves. The fourth is a sum of a knotted theta curve and a knotted handcuff curve. Note however, in all these examples of Figure 2, there is an unknotting crossing change disjoint from a summing sphere. Must that always be the case?

1.1. **Knotoids.** A knotoid (due to Turaev) is an equivalence class of an oriented regular immersion of an interval lying in a surface Σ, where Σ is typically S^2 or D^2. The notion of equivalence is induced by planar isotopy and the three Reidemeister moves for knots, fixing the endpoints of the interval. An embedding of the interval into Σ is a trivial knotoid. Analogously to the definition for knots, the unknotting number of a knotoid is the minimum number of crossing changes required to transform k into the trivial knotoid. We may similarly define the proper rational unknotting number of a knotoid.

Knotoids form a semigroup under the operation of knotoid multiplication, which is equivalent to vertex connected sum (i.e. an order 1 connected sum $k_1 \#_1 k_2$). Moreover, an order 2 connected sum $k_1 \#_2 K_2$ is equivalent to the knotoid product $k_1 \#_1 k_2$, where knotoid k_2 is obtained from the knot K_2 by removing a small open interval. A knotoid k is prime if any decomposition $k = k_1 \#_1 k_2$ implies that k_1 or k_2 is trivial.

By [Tur12, Theorem 6.2], there is a one-to-one correspondence between knotoids and oriented, ordered theta curves containing a fixed trivial constituent knot. In particular, given a knotoid k, the curve $\theta(k)$ is constructed by joining the two endpoints of the knotoid with two arcs above and below the surface Σ, which together form a trivial constituent. The product $k_1 \#_1 k_2$ of knotoids corresponds with the vertex sum $\theta(k_1) \#_3 \theta(k_2)$ of θ-curves.

Corollary 1.4. *Let k be a knotoid. If $u_{\mathrm{prop}}(k) = 1$, then k is prime. Hence, if $u(k) = 1$, then k is prime.*

Proof. Theorem 1.2 implies that either $\theta(k)$ is prime with respect to order 2 and 3 connected sums or that $\theta(k) = \theta_0 \#_2 K$ for some prime knot K of proper rational unknotting number one. In the the former case, k is prime, and in the latter case, the knotoid k is both prime and of knot-type, meaning the endpoints of k are contained in the same region of Σ. The last conclusion follows since a crossing change is a proper rational tangle replacement. $\square$

A **forbidden move** in a knotoid diagram is a planar move in which a single strand passes over or under a vertex. See [BG21, Figure 1.2]. Any knotoid can be made trivial via a finite sequence of forbidden moves. Consequently, the **f-distance** between knotoids is the (well-defined) minimal number of forbidden moves required to transform one knotoid into another. A forbidden move in a knotoid k corresponds directly with a crossing change in the associated θ-curve $\theta(k)$. Therefore, we obtain the immediate corollary:

Corollary 1.5. *A knotoid with f-distance 1 is prime.* $\square$

1.2. **Biological Motivation.** The study of both θ-graphs and knotoids is externally motivated by questions arising in molecular biology. Knotoids have recently been utilized to model knotting and entanglement in open proteins [GGL+17, GDBS17]. Spatial theta-curves arise as intermediary structures during the process of DNA replication in circular plasmids and in organisms with circular DNA (eg **E. coli**). Replication is the process through which a single DNA molecule reproduces to form two new identical DNA molecules. During this process the DNA double helix is partially unwound in the replication bubble; the two partially replicated strands (leading and lagging) together with the remaining parent DNA form a θ-curve structure. Intriguingly, this θ–curve can be knotted [ASZ+92, VHK+96, SdlCMR+98, LMRH+12]. During replication, packing, and other cellular processes, enzymatic complexes (e.g. topoisomerases) mediate strand-passage events that knot and unknot DNA. This leads naturally to an investigation of unknotting number in knots and θ–curves. Unknotting number one θ–curves are precisely the knotted structures that may arise as replication intermediaries following a single strand-passage event. A classification of such structures and understanding of unknotting processes in spatial θ–curves can shed light on both the knotting that occurs during replication, and the processes that drive it.
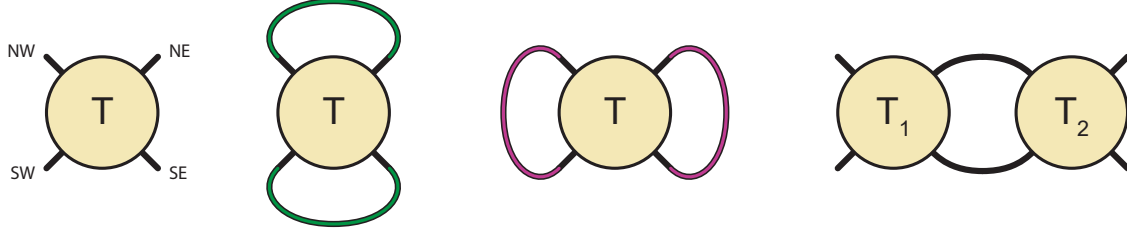
FIGURE 3. A marked two-string tangle, numerator closure indicated in green, denominator closure in magenta, and a tangle sum of marked tangles.

2. DEFINITIONS AND CONVENTIONS

We consider spatial θ-curves up to ambient isotopy. Because θ-curves are trivalent graphs, there is no difference between the notions of equivalence from rigid vertex isotopy or topological (pliant) isotopy; the results stated throughout hold in either setting.

We consider pairs (M, t) where M is a 3-manifold obtained by removing some finite number (possibly zero) of open balls from S^3 and $t \subset M$ is a properly embedded 1-dimensional CW-complex. It is a **tangle** if at least one ball has been removed from S^3 and if t is a 1–manifold. It is a **2-strand tangle** if t is the disjoint union of two arcs and M is a 3-ball. For a surface $S \subset M$, we write $S \subset (M, t)$ to mean that S is transverse to t. It is **essential** if it not a 2-sphere bounding a 3-ball disjoint from t, is incompressible, and is not ∂-parallel in the exterior of t. If S is a sphere with $|S \cap t| \leq 3$, it is a **summing sphere**. If S is a sphere with $|S \cap t| = 4$, it is a **Conway sphere**.

2.1. **Rational tangles.** If a two-string tangle (B, t) is equipped with a homeomorphism taking (B, t) to the ball of radius one in $\mathbb{R}^3$ and ∂t is taken to the NE, SE, NW and SW corners of the equatorial sphere at $z = 0$, i.e. the set of points $\{(\pm \cos(\pi/4), \pm(\sin(\pi/4), 0)\}$, we say that (B, t) is a **marked** two-string tangle. The **numerator closure**, or **denominator closure**, of a marked two-string tangle are the knot or 2-component links obtained by joining the northern and southern points, or eastern and western points, respectively, by the equatorial arcs as shown in Figure 3, left. Two marked two-string tangles $T_i = (B_i, t_i)$ for $i = 1, 2$ are **equivalent** if they are isotopic by an isotopy preserving the markings.

From two marked tangles T_1, T_2, we also may form a new marked two-string tangle $T_1 + T_2$, called the **tangle sum**, by identifying the eastern half sphere of B_1 relative the points $\{NE, SE\}$ with the western half sphere of B_2 relative the points $\{NW, SW\}$, and isotoping $t_1 \cup t_2$ relative to its boundary to be properly embedded in $B_1 \cup B_2$, as in Figure 3, right. The tangle sum is **essential** if the summing disk (the image of the half spheres that were glued together) is not parallel (via an isotopy transverse to $t_1 \cup t_2$) to the boundary of the 3-ball.

A two-string tangle (B, t) is **trivial** if t is isotopic into ∂B. It is **split** if there is a properly embedded disk separating the strands. A marked two-string trivial tangle is **rational**. Rational tangles are split. Famously, Conway [Con70] constructed a bijection between rational tangles and $\mathbb{Q} \cup \{\infty\}$ using continued fractions. Any rational number $\frac{\beta}{\alpha}$ can be expressed non-uniquely as a continued fraction

$$[a_0, a_1, \cdots, a_n] = a_0 - \cfrac{1}{a_1 - \cfrac{1}{a_2 \cdots - \frac{1}{a_n}}}.$$

The coefficients of the continued fraction determine the rational tangle. The **Conway number** of a rational tangle T is the associated element of $\mathbb{Q} \cup \{\infty\}$. Our conventions (which differ from Conway's) for constructing rational tangles are as given in Figure 4

Suppose a marked two-string rational tangle $T = T_1 + T_2$ is the Conway sum of marked two-string tangles T_1 and T_2. Observe that for each $n \in \mathbb{Z}$, we also have $T = (T_1 + [n]) + ([-n] + T_2)$. We will need two decompose two-string tangles as Conway sums and this ambiguity causes some inconvenience.

2.2. **Crossing Changes and Rational Tangle Replacement.** Suppose that M is a 3-manifold and that $\tau \subset M$ is a properly embedded graph. Given an arc $\alpha \subset M$ such that $\alpha \cap \tau = \partial \alpha$ and α is disjoint from the vertices of τ, a **rational tangle replacement** (RTR) of τ with **core arc** α is defined as follows.

$$[\,] = [0,0] = \infty \qquad [0] = 0 \qquad [3] = 3 \qquad [0,3] = -\frac{1}{3}$$

$$[a,b,c] = a - \frac{1}{b - \frac{1}{c}}$$
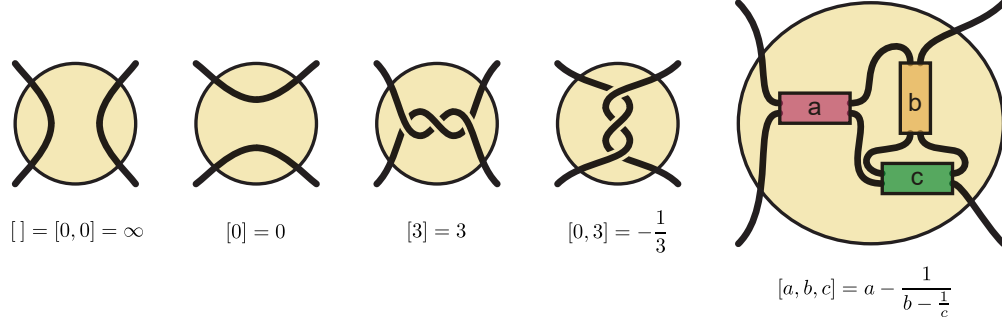
FIGURE 4. Rational tangle conventions

$$[\,] = [0,0] = \infty \qquad [0,2] = -\frac{1}{2} \qquad [0,-2] = \frac{1}{2}$$
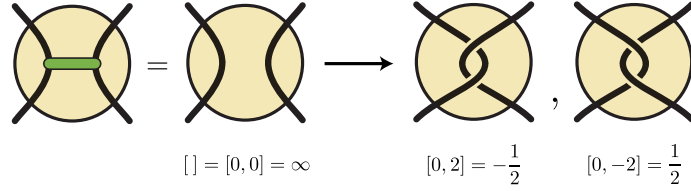
FIGURE 5. An arc α meeting τ has a regular neighborhood B meeting τ in an ∞ tangle. A crossing change of τ along α replaces this ∞ tangle with either of the $\pm\frac{1}{2}$ tangles.

Definition 2.1. Consider a regular neighborhood B of α and let $\lambda = \tau \cap B$. Then $T = (B, \lambda)$ is a trivial 2-strand tangle. By imposing a marking, without loss of generality, we may take it to be the ∞-tangle. (This choice is called a **framing** of α. Framings are parameterized by the integers.) Let $T' = (B, \lambda')$ be a rational tangle. Then $\tau' = \big((M, \tau) \setminus T\big) \cup T'$ is said to be obtained by a **RTR** on τ with (framed) crossing arc α. See Figure 5. When (B, λ') is either of the $\pm 1/2$ tangles, the RTR is a **crossing change**. When the strands λ' are properly homotopic in B to the strands λ, the RTR is a **proper rational tangle replacement**.

Remark 2.2. A crossing change could also be defined via ± 1 Dehn surgery on a loop (called the **crossing loop**) bounding a disc intersecting τ in two points. However, the crossing arc perspective is more useful for us. It is possible to translate results concerning crossing arcs into results concerning crossing loops and vice versa, although care is needed. See [Lac21], for example.

Definition 2.3. Given two rational tangles (B, λ) and (B, λ') with $\partial\lambda = \partial\lambda'$, each contains a properly embedded disc separating the strands, unique up to isotopy disjoint from the strands. The **distance** Δ of an RTR is half the minimal number of intersections between the boundaries of such disks for the tangles (B, λ) and (B, λ') in Definition 2.1. This coincides with other definitions of the distance between rational tangles in the literature.

Remark 2.4. A crossing change has $\Delta = 2$. If $\Delta = 1$, the RTR is not a proper RTR.

2.3. **Sums and Summing Spheres.** We give a more precise definition of connected sum than what is in the introduction.

Definition 2.5. Suppose that K_1 and K_2 are spatial graphs in distinct copies of S^3. An order 2 connected sum is obtained as follows. For $i = 1, 2$, remove a small open balls from the interior of an edge of K_1, K_2, respectively. This results in 3-balls B_1, B_2 containing the remnants of K_1, K_2 respectively. Glue ∂B_1 to ∂B_2 by a homeomorphism taking $K_1 \cap \partial B_1$ to $K_2 \cap \partial B_2$. The result is a spatial graph $K_1 \#_2 K_2$ in S^3. We form an order k connected sum $K_1 \#_3 K_2$ similarly, except we choose each of the small open balls to be centered at a degree k vertex of K_1, K_2 respectively [Wol87]. Observe that the image of ∂B_1 and ∂B_2 is a **summing sphere** for the sum. We can also similarly define order 2 sums for tangles. A spatial graph or tangle that has a nontrivial knot as an order 2 connect summand is said to be **locally knotted**.

A θ-curve is **composite** if it is either an order 2 connected sum of a (possibly trivial) θ-curve and a nontrivial knot or an order 3 connected sum of nontrivial θ-curves. In such a case, as the summing sphere is essential, we also say that the sum is **essential**. A θ-curve which is neither trivial nor composite is **prime**.

Conversely, if $S \subset S^3$ is a sphere transverse to a θ-curve θ and intersecting it in two or three points, we can cut (S^3, θ) open along S and glue in two copies of (B^3, τ) where τ is the cone on $\theta \cap S$. If S is essential and $|S \cap \theta| = 3$, then neither component of the result is a trivial θ-curve. If $|S \cap \theta| = 2$, then one component (S^3, θ_1) of the result is a θ-curve and the other (S^3, K) is a knot.

Remark 2.6. Our definition of prime θ-curve differs slightly from that of Turaev [Tur12] and Matveev-Turaev [MT11]. They would say that the order 2 connected sum of a trivial θ-curve with a prime knot is a prime θ-curve, while we say that it is composite. Other than that difference, our definitions coincide.

2.4. **Preliminary results on Rational Tangles and Conway Sums.** In this section we collect some well-known results which will be useful to us.

Lemma 2.7. *Let T be a 2–strand tangle without local knotting. Suppose T contains essential disks D_1 and D_2 each intersecting the strands of T in zero or two points in their interior. Then*

    (1) ∂D_1 and ∂D_2 are isotopic in ∂T,
    (2) the disks have the same number of punctures, and
    (3) after an isotopy of the disks $D_1 \cap D_2 = \emptyset$.

Furthermore, if the disks are disjoint from the strands of T, then they are isotopic and T is a rational tangle.

Proof. Out of all such pairs of disks isotopic to D_1 and D_2, we may assume that D_1 and D_2 were chosen so as to minimize $|D_1 \cap D_2|$.

Suppose that there is a circle of intersection in $D_1 \cap D_2$. Let ζ be an innermost such circle on D_1, bounding an innermost disc $E_1 \subset D_1$. Let $E_2 \subset D_2$ be the disc with boundary ζ. Then $E_1 \cup E_2$ is an embedded sphere in T. Since every sphere in T separates, the parity of the punctures on E_1 and E_2 is the same. Hence, either they both have a single puncture or they both have 0 or 2 punctures. If E_1 has 2 punctures but E_2 has 0, then the sphere $E_1 \cup E_2$ bounds a ball that intersects the strands of T in a single unknotted arc (since T has no local knotting). This arc may be isotoped rel-∂ through the ball to an arc in E_1 with the same endpoints. In particular this indicates that there is another disk in this ball disjoint from the unknotted arc and meeting D_1 only in ζ. Hence this other disk is a compressing disk for D_1, contradicting that D_1 is an essential disk. If E_1 has 0 punctures but E_2 has 2, then E_1 is a compressing disk for D_2, contradicting that D_2 is an essential disk. If E_1 and E_2 both have 0 punctures or both have 1 puncture, then there is an isotopy of E_2 to E_1 through the ball they cobound due to the lack of local knotting and closed components. Pushing E_2 slightly past E_1 then achieves an isotopy of D_2 that reduces $|D_1 \cap D_2|$ contrary to the assumed minimality. Thus we are left with in the situation that any simple closed curve of the intersection $D_1 \cap D_2$ must bound a disk in each D_1 and D_2 that is punctured twice by the strands of T. In other words, any simple closed curve of $D_1 \cap D_2$ must bound an unpunctured annulus with ∂D_1 in D_1 and another with ∂D_2 in D_2. Let us set this last case aside for the moment.

So now assume $\partial D_1 \cap \partial D_2 \neq \emptyset$ so that $D_1 \cap D_2$ necessarily contains arcs. Each arc of $D_1 \cap D_2$ cuts each D_1 and D_2 into two subdisks. For each D_1 and D_2, at least one of these two subdisks is intersected by the strands of T at most once and therefore contains no simple closed curve component of $D_1 \cap D_2$. So let ζ be an arc component of $D_1 \cap D_2$ that is outermost in D_1, say, cutting off a disk E_1 that intersects D_2 only in the arc ζ. Let ϵ be the arc $\partial E_1 \backslash \zeta$. Note that ϵ lies in ∂T disjoint from the markings and connecting two points in ∂D_2. Since D_2 is essential in the complement of the strands of T, ∂D_2 separates ∂T into two disks, each with two punctures. Let F be the component of $\partial T \backslash \partial D_2$ that contains ϵ. If ϵ is inessential in F, then an isotopy reduces $|\partial D_1 \cap \partial D_2|$ taking ζ to a closed loop ζ' that bounds a disk E_1' intersecting the strands of T at most once. Yet as we have seen, such loops imply that either a disk is not essential or there is an isotopy contradicting the minimality of $|D_1 \cap D_2|$. Thus, ϵ separates the punctures of F.

Now suppose E_1 is disjoint from the strands of T and use E_1 to ∂–compress D_2 into two disks Δ and Δ'. Each puncture of $\Delta \cup \Delta'$ is also a puncture of D_2, as E_1 contains no punctures. As ϵ separates the punctures of F and as E_1 is unpunctured, this implies that Δ and Δ' are each once-punctured by the strands of T. As F is a twice-punctured disc, each of $\partial \Delta$ and $\partial \Delta'$ bounds a disc in F containing a single puncture of F.

Since no strand of T contains a local knot, the discs Δ and Δ' are inessential, being isotopic rel-∂ into F. Since D_2 may be recovered by banding Δ and Δ' together along an arc in F dual to ϵ, it follows that D_2 is isotopic rel-∂ to the twice punctured disk F in ∂T. However this contradicts that D_2 is essential.

Thus it follows that every arc component of $D_1 \cap D_2$ is an arc separating the punctures in both D_1 and D_2. In particular, if there are are arc components of $D_1 \cap D_2$, then both D_1 and D_2 are twice punctured, the arcs separate their punctures, and there are no simple closed curves of $D_1 \cap D_2$.

So now let ζ be an arc of $D_1 \cap D_2$ that is outermost in D_1, bounding a disk $E_1 \subset D_1$ that intersects D_2 only in ζ and is punctured once by the strands of T. Use E_1 to ∂–compress D_2 to disks Δ and Δ' which are each twice punctured. Again let $\epsilon = \partial E_1 \backslash \zeta$ and let F be the disk in ∂T bounded by ∂D_2 that contains ϵ. As before, ϵ must separate the punctures of F. However, this now implies that $\partial \Delta$ bounds a disk in F that is once-punctured by the strands of T. Yet since Δ is homologous through the tangle to this disk, the strands of T must also intersect Δ an odd number of times, a contradiction.

Finally, we may return to the case that $D_1 \cap D_2$ contains loops parallel to ∂D_1 and ∂D_2. However now we may assume there are no arcs of intersection. That is, we may assume $\partial D_1 \cap \partial D_2 = \emptyset$.

Since the boundary of T is a sphere with four punctures, if $|\partial D_1 \cap \partial D_2| = 0$, then ∂D_1 is isotopic to ∂D_2 as desired. Then if one is disjoint from the strands of T it follows that the other must be as well since they are essential disks. Hence either both disks are disjoint from the strands of T or both are intersected twice. Since T has no local knotting, if the disks are disjoint from the strands, then T must be a rational tangle and so the two disks are isotopic.

So suppose the disks both intersect the strands twice. As we are presently assuming $|\partial D_1 \cap \partial D_2| = 0$, suppose $D_1 \cap D_2$ is a non-empty collection of closed curves. First consider an innermost disk E of $D_1 \cup D_2$ in the complement of the intersection $D_1 \cap D_2$. Because the disks are essential, E must be punctured by the strands of T. Because there is no local knotting in the strands of T, E must be punctured twice. Hence the curves of $D_1 \cap D_2$ are all parallel to ∂D_1 in $D_1 - T$ and also parallel to ∂D_2 in $D_2 - T$. Observe that a curve C of $D_1 \cap D_2$ outermost in D_1 must also be outermost in D_2. So C cuts off annuli A_1 from D_1 and A_2 from D_2 where $C = A_1 \cap A_2$. Then since there are no closed components of the strands of T, the annulus $A_1 \cup A_2$ is boundary parallel and guides a proper isotopy of D_1 that reduces $|D_1 \cap D_2|$, contrary to assumption. Hence, when isotoped to intersect minimally, $D_1 \cap D_2 = \emptyset$. $\qquad\square$

The following is a standard result concerning rational tangles, although we do not know of a source for the proof. It can also be proved using branched double covers.

Corollary 2.8 (cf. [KL04, page 203]). *A marked tangle that is a tangle sum is integral if and only if each summand is integral.*

Proof. It follows immediately from the definition that the tangle sum of integral tangles is integral. So suppose that T is an integral tangle that decomposes as the sum $T = T_1 + T_2$. As an integral tangle, T is a rational tangle. Thus T has no local knotting and no closed components. Furthermore being rational implies it has an essential disk D_1 disjoint from the strands of T and separating them. As a tangle sum, T contains a disk D_2 that splits T into the tangles T_1 and T_2, intersecting the strands of T twice. Observe that by the definitions of integral tangle and tangle sum, ∂D_1 and ∂D_2 are not isotopic in ∂T. Lemma 2.7 therefore implies that D_2 cannot be an essential disk. Thus D_2 is either compressible or ∂–parallel.

If D_2 were compressible, say by a compressing disk E in T_2, then the result of the compression would bound a disk D_2' in T_2 that is disjoint from the strands of T_2 and separates them. Hence it must also separate the strands of T. Thus D_2' is also an essential disk for T. However as $\partial D_2'$ is isotopic to ∂D_2, it is not isotopic to ∂D_1. This contradicts Lemma 2.7.

Therefore D_2 must be ∂–parallel. Thus either T_1 or T_2, say T_2, is an integral tangle through which the parallelism occurs. Yet then taking the tangle sum of T with the mirror $\overline{T_2}$ of T_2 would yield $T + \overline{T_2} = T_1 + T_2 + \overline{T_2} = T_1$. (Note that $T_2 + \overline{T_2}$ is the 0–tangle precisely because T_2 is an integral tangle.) As T_1 is now expressed as the sum of two integral tangles, it too is integral. $\qquad\square$

Corollary 2.9. *Suppose the marked tangle T has no local knotting or closed components and decomposes as a tangle sum along an essential summing disk. Then any other essential twice-punctured disk in T has boundary with slope ∞, the slope of the boundary of the summing disk.*

Proof. Let D_1 be the disc defining the tangle sum and let D_2 be any other essential twice-punctured disk. Since both are twice-punctured essential disks, then by Lemma 2.7 their boundaries are isotopic in the complement of the marked points. □

3. Proof of Theorem 1.1 and Theorem 1.2

3.1. **Case 1: When θ is locally knotted.**

Proposition 3.1. *If θ is a locally knotted θ–curve with $u_{\text{prop}}(\theta) = 1$, then:*
- *$\theta = K \#_2 \theta_0$ where K is a non-trivial prime knot with $u_{\text{prop}}(K) = 1$ and θ_0 is a trivial θ–curve; and*
- *any proper rational unknotting arc for θ can be isotoped to be disjoint from the summing sphere.*

Proof. Assume θ is locally knotted. Let α be a proper rational unknotting arc for θ. Then the endpoints of α lie on either one or two edges of θ. Let e be an edge of θ that does not meet α and let e_1 be another edge. Suppose that θ has a local knot K in the edge e. Without loss of generality α meets the knot $C = e \cup e_1$ at most once, along the edge e_1. Note that since K is in e, the knot C is nontrivial. After a proper RTR along α, the knot C is unchanged. Thus, $u_{\text{prop}}(\theta) \neq 1$. This is a contradiction. So θ has no local knotting in the edge e.

Consider the tangle $T = (B, t)$ that is the exterior of the edge e. Since α is disjoint from e, the tangle T contains α. Let M' be the double branched cover of B, branched over t. The arc α lifts to a knot $\tilde{\alpha}$ in M'. By assumption, performing the proper RTR on t using α converts T into a trivial tangle. Let M be the double branched cover of B with branch set the strands of this trivial tangle. Observe that M is a solid torus. Let $\tilde{\alpha}^*$ be the lift of the dual arc to α; it is a knot in M. The proper RTR along α induces a non-integral Dehn surgery on $\tilde{\alpha}$ converting M' to M and with $\tilde{\alpha}^*$ being the core of the surgery solid torus. It is non-integral in the sense that on the torus that is the boundary of a regular neighborhood of $\tilde{\alpha}$ the slope of the meridian for $\tilde{\alpha}$ and the slope of the meridian for $\tilde{\alpha}^*$ intersect minimally at least twice. Scharlemann proved[1]:

Theorem 3.2 (Abridged main theorem of [Sch90])**.** *Let k be a knot in a compact orientable 3–manifold M whose exterior is irreducible and ∂–irreducible. Let M' be a manifold obtained by Dehn surgery on k. If ∂M compresses in M or M contains a sphere not bounding a rational homology ball, then either*

> *(a),(b),(d) M' is irreducible,*
> *(c) k is cabled and the slope of the surgery is that of the cabling annulus, or*
> *(e) $M = (S^1 \times S^2) \# W$ where W is a rational homology sphere.*

□

Each summing sphere for T lifts to an essential sphere in M', so M' is reducible. Thus, neither (a), (b), nor (d) can occur. Option (c) implies that $\tilde{\alpha}$ is a cable knot and that the surgery slope is the slope of the cabling annulus. Such a slope intersects the meridian of $\tilde{\alpha}$ exactly once, rather than at least twice. Consequently, Option (c) is impossible. Option (e) is impossible since M is a solid torus.

Since the conclusions of Scharlemann's Theorem 3.2 fail, some hypothesis must be wrong. In particular, the exterior of K' (the lift of α) must be either reducible or ∂–reducible. Since the boundary of the exterior of K' is two tori, a compression of one yields a sphere that doesn't bound a ball. Hence if the exterior of K' is ∂–reducible, then it is reducible as well. Thus we assume the exterior of K' is reducible. Again by the Equivariant Sphere Theorem [MSY82] we may assume there is a reducing sphere $\tilde{S}$ such that either (i) $\tau(\tilde{S}) \cap \tilde{S} = \emptyset$ or (ii) $\tau(\tilde{S}) = \tilde{S}$ where $\tilde{S}$ is transverse to the fixed set of τ.

In (i), $\tilde{S}$ is disjoint from the fixed set of τ and projects to a sphere S' disjoint from the strands t of the tangle T. Then S' bounds a 3–ball B'. Since t has no closed components, $B' \cap t = \emptyset$. Therefore $\tilde{S}$ bounds a ball, a contradiction.

In (ii), $\tilde{S}$ projects to a sphere S' that intersects the strands t of the tangle T twice but is disjoint from α. As S' must bound a 3–ball B' in B and t has no closed components, $B' \cap t$ is a single arc. This arc must be non-trivial since otherwise $\tilde{S}$ will bound a ball. Observe that this implies S' is a summing sphere for θ.

---

[1]See the comment at the end of the paper for statement (e), although there is a typo: the first M' in the comment should be an M.

Since α is disjoint from S', either $\alpha \subset B'$ or α is disjoint from B'. If α is disjoint from B', then after the proper RTR along α, B' still meets T in a non-trivial knotted arc. This is a contradiction. Hence $\alpha \subset B'$. Then the proper RTR along α must unknot the arc $B' \cap t$. This implies that θ is the connected sum of a θ-curve θ_1 and a proper rational unknotting number one knot κ. The arc $B' \cap t$ is the result of removing an arc from κ. The argument of [Zha91] that knots with unknotting number 1 are prime (cf. [Sch85]) immediately generalizes to show that knots with proper rational unknotting number 1 are prime. Since α is a proper rational unknotting arc for θ, the curve θ_1 must be the trivial θ-curve. Consequently, S' is, up to isotopy, the unique summing sphere for θ; in particular, S is isotopic to S', as desired. $\qquad \square$

3.2. **Case 2: When θ is a vertex sum and not locally knotted.** Throughout this subsection we assume our θ–curves and tangles are not locally knotted.

Proposition 3.3. *Assume θ is a vertex connected sum of two non-trivial theta curves. Further assume θ has no local knots. Then $u_{\mathrm{prop}}(\theta) > 1$.*

Proof. For contradiction, we assume $u_{\mathrm{prop}}(\theta) = 1$ and α is the arc of a proper rational tangle replacement that unknots θ. Then α will meet θ in one or two edges. As before, let e be an edge of θ that does not meet α. By hypothesis, θ has an essential summing sphere S meeting θ in three points but there is no essential sphere meeting θ in just two points. For ease of exposition, we choose a fixed diagram of θ and a fixed regular neighborhood $N(e)$ of e such that near the diagram S intersects the plane of projection as in Figure 6.
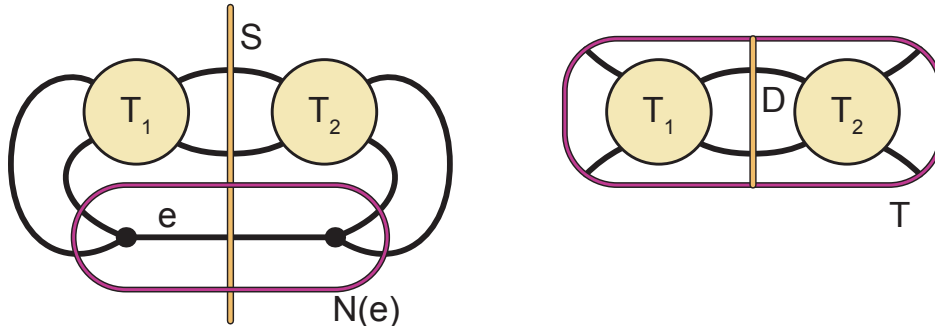

FIGURE 6. Left: The edge e and sphere S for the theta curve θ. Right: The exterior of the edge e is the tangle T. The sphere S meets T in a disk D expressing T as the tangle sum $T_1 + T_2$.

Let T be the tangle whose underlying 3-ball is $X(e)$, the closure of the complement of $N(e)$, and whose strands are the intersection of θ with this 3-ball. The intersection of S with T is a disc D intersecting the strands of T exactly twice. As in Figure 6, we may consider T as a marked tangle with the disk D decomposing T as a tangle sum of two marked tangles T_1 and T_2, thus $T = T_1 + T_2$.

Since the sphere S is essential, D is an essential summing disk of T. Consequently, neither T_1 nor T_2 is an integral tangle. Moreover, neither T_1 nor T_2 may be a split tangle since otherwise θ would contain a non-trivial knot summand contrary to hypothesis.

Since α is an arc along which a proper rational tangle replacement unknots θ, the unknotting proper rational tangle replacement along α must transform T into an integral tangle. However, as T is a non-trivial tangle sum, Proposition 4.6 below shows this is impossible in the case of a crossing change, whereas Proposition 4.9 shows this is impossible in the case of a proper rational tangle replacement of distance greater than two. Hence we arrive at a contradiction. $\qquad \square$

4. TANGLE SUMS AND PROPER RATIONAL TANGLE REPLACEMENTS

In this section, we investigate proper rational tangle replacements with respect to tangle sums. We then construct a specific tangle ρ (shown in Figure 7) whose denominator closure has sufficiently high unknotting number and proper rational unknotting number. This tangle will serve a specific role in the proofs of both Proposition 4.6 and Proposition 4.9, which handle the cases of crossing changes (distance two) and proper rational tangle replacements (distance greater than two) along marked tangle sums separately.

Proposition 4.1. *Suppose that C is a collection of mutually disjoint essential Conway spheres in an essential tangle T. If a marked tangle T is the tangle sum of tangles, neither of which is integral, then there exists an essential summing disc for T disjoint from C.*

Proof. Out of all essential summing disks for T that are transverse to C, let D be chosen to minimize $|D \cap C|$. For a contradiction, assume that $|C \cap D| > 0$.

Claim 4.2. *C cuts D into a disk with two marked points and a collection of annuli.*

Proof. A standard innermost loop argument shows that the intersection consists of some number of curves, each of which bounds an annulus in D with ∂D that is disjoint from the marked points. For completeness, we provide details. Let γ be a curve of $C \cap D$ that is innermost in D, and let $d \subset D$ be the subdisk that it bounds.

Suppose d is disjoint from the marked points on D. Then, since C is essential, γ must also bound a disk $c \subset C$ that is disjoint from the four marked points on C. Then $c \cup d$ is a sphere disjoint from the strands of T. Since T is a tangle in a ball, this sphere $c \cup d$ must bound a ball. Because the tangle T has no closed components, this ball is disjoint from the strands of T. Hence, there is an isotopy of D with support in a neighborhood of this ball that takes c through the ball to d and then just past D. Such an isotopy removes the intersection γ (and possibly other intersections) without introducing new intersections. Thus it contradicts the minimality of $|C \cap D|$.

If d contained just one of the marked points, then as every sphere in a ball separates, γ would have to bound a disk $c \subset C$ that contains just one of the four marked points in C. However then $c \cup d$ would be a sphere bounding a ball in T that intersects the strands in T in a single arc. Since T has no local knotting, this arc must be an unknotted arc running from c to d. Hence, as in the previous situation, we may isotope d along the arc through d and just past c so remove the intersection γ without introducing new intersections. Thus this again contradicts the minimality of $|C \cap D|$. $\square$

Claim 4.3. *D cuts C into two disks each with two marked points and a collection of annuli.*

Proof. We will use another innermost circle argument to show that any curve of $C \cap D$ that is innermost in C bounds a subdisk containing exactly two marked points of C. This implies that $C \cap D$ consists of a family of parallel curves that separate the four marked points of C into two pairs. From that our claim follows.

Let γ be a curve of $C \cap D$ that is innermost in C, and let $c \subset C$ be the subdisk that it bounds. Out of all such curves, we may assume that we chose γ and c to contain as few marked points as possible. Since C has four marked points, this means that c has two or fewer marked points. We show that it has exactly two.

Suppose c is disjoint from the marked points of c. Then since $\gamma = \partial c$ cobounds an annulus $a \subset D$ with ∂D that is disjoint from the marked points of D (by Claim 4.2), we may form a disk $\delta = a \cup c$ with $\partial \delta = \partial D$ that is disjoint from the strands of T. But then T is a split tangle, contradicting the assumption that T is essential.

Now suppose c contains only one marked point of C. Then since $\gamma = \partial c$ bounds a subdisk $d \subset D$ containing the two marked points of D (by Claim 4.2), we may form a sphere $c \cup d$ that meets T in three points. Since this sphere bounds a ball in the ball underlying T, this is again a contradiction. $\square$

Let a be the outermost annulus of $D \backslash C$. Then ∂a is ∂D and a component γ of $C \cap D$. Let C_γ be the component of C containing γ. Let c' and c'' be the two complementary disks of C_γ bounded by γ. Then let $D' = a \cup c'$ and $D'' = a \cup c''$. These are disks that each meet T in two points and have $\partial D = \partial D' = \partial D''$. Furthermore, a slight isotopy makes C disjoint from both D' and D''. Thus, to contradict the minimality of $|C \cap D|$, it remains to see that at least one of D' and D'' is an essential summing disk.

Observe that D' and D'' decompose T into a sum $T = T_1 + T_2 + T_3$ where C_γ bounds the tangle T_2. Hence T_2 is an essential tangle, and in particular it is not integral. If both D' and D'' were inessential summing disks, then both T_1 and T_3 would be integral by Corollary 2.8. However, if both T_1 and T_3 are integral tangles, then the component C_γ of C is isotopic to the boundary of the ball. But this contradicts that each component of C is an essential Conway sphere. Thus at least one of D' and D'' is an essential summing disk. $\square$

Let α continue to denote an unknotting arc for θ, or more generally, the arc of a proper rational tangle replacement that unknots θ.
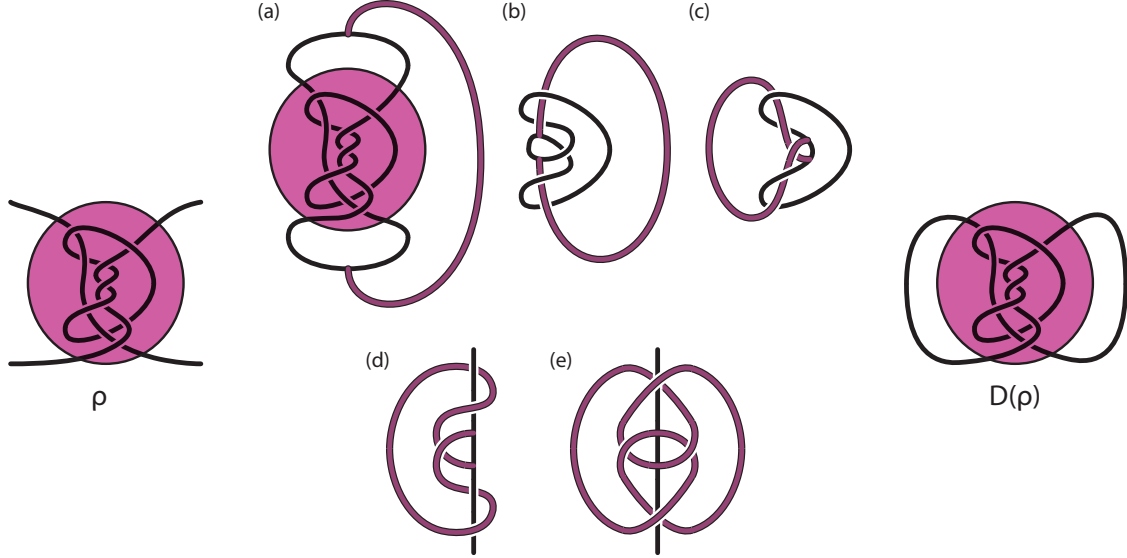
FIGURE 7. Left: The tangle ρ. Center: The normal closure of ρ with its core arc (a) is isotopic to an unknotted axis with an arc (b)–(d). The double branched cover along the axis (e) is the figure eight knot, 4_1. Right: The denominator closure $D(\rho)$ of ρ is observed to be the knot 10_{139}.

Lemma 4.4. *Every essential summing disk for T must intersect α.*

Proof. Suppose an essential summing disk for T that is disjoint from α expresses T as the the sum $T = T_1 + T_2$. Thus neither T_1 nor T_2 is integral, and we may assume α is contained in T_2.

Say performing the crossing change (or proper RTR) along α in T_2 produces T_2^*. Then after the crossing change (proper RTR) along α in T, the summing disk expresses the resulting tangle as $T^* = T_1 + T_2^*$. Since T^* is an integral tangle, Corollary 2.8 implies that both T_1 and T_2^* are integral tangles. However this contradicts that T_1 is non-integral. $\qquad\square$

Claim 4.5. *The tangle ρ shown in Figure 7 has unknotted numerator closure $N(\rho)$, is not a non-trivial tangle sum, and contains no essential torus, essential Conway sphere, or essential annulus. In particular, the double branched cover of ρ is the exterior of the hyperbolic knot 4_1, the figure eight knot. Furthermore, its denominator closure is the knot 10_{139} which has unknotting number 4 and proper rational unknotting number ≥ 2.*

Proof. Ignoring the extra arc, Figure 7(a)–(c) shows that $N(\rho)$ is the unknot. This extra arc, however, is the core arc of the trivial rational tangle joined to ρ to make $N(\rho)$. Figure 7(a)–(d) shows the straightening of the unknot $N(\rho)$ at the expense of twisting up this arc. Figure 7(e) shows the knot that is the lift of the arc to the double branched cover along the unknot. One readily observes that this is the figure eight knot, 4_1. Since ρ was the exterior tangle of the core arc, this shows that the double branched cover of ρ is the exterior of the figure eight knot 4_1.

If ρ were to contain an essential summing disk, an essential torus, an essential Conway sphere, or essential annulus then the double branched cover of ρ would contain either an essential annulus or an essential torus and hence be non-hyperbolic. However, as the double branched cover of ρ is the exterior of the hyperbolic knot 4_1, this cannot be.

The denominator closure of ρ, $D(\rho)$, is a 10 crossing knot which SnapPy [CDGW] identifies as the knot 10_{139} (which one may also confirm by hand). As informed by KnotInfo [LM21], this knot has unknotting number 4, see [Kaw98, GI02]. It also has proper rational unknotting number ≥ 2 [ILM21]. $\qquad\square$

11

4.1. **Sums of non-integral tangles admit no crossing changes to integral tangles.** In what follows, we study crossing changes on 2-strand tangles, making heavy use of work of Eudave-Muñoz and Gordon–Luecke. Eudave-Muñoz used tangles to construct a family of strongly invertible hyperbolic knots having distance 2 Dehn surgeries to toroidal 3-manifolds. Gordon and Luecke show that unknotting number one knots with essential Conway spheres have a close relationship to Eudave-Muñoz's construction.

Proposition 4.6. *Let T be a marked tangle sum of tangles, neither of which is integral. Then T admits no crossing change to an integral tangle.*

Proof of Proposition 4.6. Let the marked tangle T be expressed as $T = T_1 + T_2$ for some non-integral tangles T_1 and T_2. By the definition of tangle sum, the summing disks for this decomposition of T has boundary of slope ∞. Indeed, by Corollary 2.9, every essential twice-punctured disk in tangle T has boundary slope ∞.

Assume for contradiction that T admits a crossing change along an arc α to an integral tangle. By adding the mirror of that integral tangle to T, say along T_2, and updating T to be that result, we may assume the crossing change along α transforms T into the 0–tangle.

Let $K_\rho = N(\rho + T)$, the numerator closure of $\rho + T$ where ρ is the tangle shown in Figure 7. Since a crossing change along α transforms T to the 0–tangle, it also transforms $K_\rho = N(\rho + T)$ into $K_\rho^* = N(\rho + 0) = N(\rho)$ which is the unknot by Claim 4.5. Hence K_ρ is a knot with unknotting number 1 and unknotting arc $\alpha \subset T$. Because both T and ρ are essential tangles, it follows that the sphere $\partial \rho$ (which is isotopic to ∂T) is an essential Conway sphere for K_ρ. Thus K_ρ is a knot with unknotting number 1 and an essential Conway sphere.

Therefore, by [GL06, Theorem 6.2], up to mirroring K_ρ, either

(1) any unknotting arc for K_ρ may be isotoped to be disjoint from all essential Conway spheres and essential tori in the exterior of K_ρ,
(2) K_ρ is an EM-knot, or
(3) K_ρ is the union of an essential tangle P and an essential EM-tangle P_0.

Furthermore, in case (3), every unknotting arc for K_ρ is isotopic into the EM-tangle P_0 and the standard unknotting arc for P_0 continues to be an unknotting (trivializing) arc for K_ρ. Figure 8 shows the two families of EM-tangles, $A_1(\ell, m)$ and $A_2(\ell, m)$, with their standard unknotting arcs and essential summing disks, see [GL06, Figure 4.3]. They are depicted, up to adding integral tangles, so that the summing disk decomposes the EM-tangles as a non-trivial tangle sum. So that the summing disks are essential and the tangle sum is non-trivial: for $A_1(\ell, m)$ we assume $|\ell| > 1$, $m \neq 0$ and $(\ell, m) \neq (2, 1)$ or $(-2, -1)$; for $A_2(\ell, m)$ we assume $|\ell| > 1$, $m \neq 0, 1$. Also shown are the rational tangles resulting from the crossing change along the standard unknotting arc. In these tangles, the summing disk is unique (see eg. [GL06, Addendum to Theorem 4.3]).

Case (1). Assume the unknotting arc α for K_ρ may be isotoped to be disjoint from all essential Conway sphere and essential tori in the exterior of K_ρ.

Claim 4.7. *At least one of T_1 or T_2 is an essential tangle.*

Proof. If not, then both are rational tangles since we are assuming T is not locally knotted. Since neither T_1 nor T_2 is an integral tangle, then $D(T)$, the denominator closure of T, is a connected sum of two non-trivial two-bridge knots. Then the crossing change along α taking T to the 0–tangle transforms $D(T)$ to the unknot. Yet this contradicts that unknotting number one knots are prime [Sch85]. $\square$

Say T_2 is an essential tangle. Then since $\rho + T_1$ is not split (as it is neither rational nor locally knotted), ∂T_2 is an essential Conway sphere in K_ρ. After a slight push into the interior of T we may assume ∂T_2 is disjoint from ρ. Then, by the hypothesis of this case, we may assume that α is disjoint from both $\partial \rho$ and ∂T_2.

If α were contained in T_2, then the crossing change would change T_2 into T_2^*. Thus $T_1 + T_2^*$ must be the 0–tangle since that is the only way to fill ρ to get the unknot. However the 0–tangle, indeed any integral tangle, only decomposes as a sum of two tangles if the two tangles are integral by Corollary 2.8. Yet this implies that T_1 is an integral tangle, contrary to assumption. Hence α is not contained in T_2.

If α were contained in T but not in T_2, then the crossing change does not alter that each of the four strands of the tangle between ρ and T_2 connect $\partial\rho$ to ∂T_2. Hence the crossing change cannot make either of these Conway spheres become inessential.

Indeed, suppose there were a compressing disk D for one of the spheres after the crossing change. Isotop it rel–∂ to intersect the two spheres minimally. Being a compressing disk for one of the spheres, any intersections in the interior of D must be with the other sphere. If there are any such intersections, then they will cut off an innermost subdisk D' from D. By the presumed minimality, D' must be a compressing disk too. Hence we may restrict attention to a compressing disk D that meets the two spheres only in its boundary. Such a disk D would have to be contained in the $S^2 \times I$ tangle between ρ and T_2 since both those tangles are essential. Yet then D must separate a strand of the tangle from the other strands, preventing it from connecting $\partial\rho$ to ∂T_2. This is a contradiction.

Since the result of the crossing change on K_ρ has an essential Conway sphere, it cannot be the unknot. Hence α is not contained in T.

Thus α may be isotoped into ρ. Suppose the crossing change transforms ρ into ρ^* as it changes K_ρ to K_ρ^*. By assumption, K_ρ^* is the unknot. If ρ^* is not a rational tangle, then again $\partial\rho^*$ will be an essential Conway sphere in K_ρ^* which should be the unknot, a contradiction. Hence ρ^* must be rational. Because ρ is homotopic to the 1–tangle and a crossing change will not alter the homotopy type, ρ^* is also homotopic to the 1–tangle. In particular, ρ^* is a rational tangle that is not the ∞–tangle.

Suppose T_1 is an essential tangle. Since ρ^* is not the ∞–tangle, $\rho^* + T_1$ must again be an essential tangle. Then, as T_2 is also an essential tangle, $\partial(\rho^* + T_1) = \partial T_2$ is an essential Conway sphere for K_ρ^*, a contradiction. Hence T_1 must be an inessential tangle. By assumption it is not locally knotted and non-integral, so T_1 must be a non-integral rational tangle.

If ρ^* is also a non-integral rational tangle, then $\rho^* + T_1$ is a tangle sum of two non-integral (and non-∞) rational tangles. Hence $\rho^* + T_1$ is an essential tangle. Therefore, as before, $\partial(\rho^* + T_1) = \partial T_2$ is an essential Conway sphere for K_ρ^*, a contradiction.

Therefore ρ^* must be an integral tangle. Hence $D(\rho^*)$, the denominator closure of ρ^*, is an unknot. However, since $D(\rho^*)$ is obtained from the knot $D(\rho)$, the denominator closure of ρ, by a crossing change along α, then $D(\rho)$ has unknotting number 1. This contradicts that $D(\rho)$ has unknotting number 2 as determined in Claim 4.5.

Case (2). An EM-knot has a unique Conway sphere by [EMn97] (see also [GL06, Theorem 6.2(2)(b)]). This sphere splits the EM-knot into two Montesinos tangles, each a tangle sum of two non-integral rational tangles. While the essential Conway sphere $\partial\rho$ splits K_ρ into the tangles T and ρ, the tangle ρ is not a tangle sum as shown in Claim 4.5. Hence K_ρ cannot be an EM-knot as in case (2) above.

Case (3). Thus K_ρ must be as in case (3) above: there is an EM-tangle P_0 contained in K_ρ so that ∂P_0 is an essential Conway sphere and every unknotting arc for K_ρ may be isotoped into P_0. Furthermore, the standard unknotting arc for P_0 must continue to be an unknotting arc for K_ρ. Hence we shift out attention upon the standard unknotting arc of P_0, regardless of whether α is isotopic to it in K_ρ.

Claim 4.8. *Either $P_0 \subset T$ where ∂P_0 is an essential Conway sphere or $P_0 = T$.*

Proof. Assume the Conway spheres $\partial\rho$ and ∂P_0 have been isotoped in K_ρ to intersect minimally. Since both $\partial\rho$ and ∂P_0 are essential Conway spheres in K_ρ, the curves of $\partial\rho \cap \partial P_0$ must all be isotopic in each $\partial\rho$ and ∂P_0 and bound disks in both that intersect the strands twice. In particular, the components of $\rho \cap \partial P_0$ must either be essential twice-punctured disks or essential annuli. However by Claim 4.5, ρ may contain neither. Hence $\partial\rho \cap \partial P_0 = \emptyset$.

If $\partial\rho$ is contained in P_0, then since $\partial\rho$ is essential in K_ρ, either $\partial\rho$ is also essential in P_0 or it cobounds a product tangle with ∂P_0. The latter implies that we may take $\partial\rho = \partial P_0$ by an isotopy through the product tangle. Since the double branched cover of an EM-tangle such as P_0 is a graph manifold (see [GL06, p2069–2070]), neither is it a hyperbolic manifold nor does it contain any essential torus bounding a hyperbolic manifold. Since the double branched cover of ρ is a hyperbolic manifold by Claim 4.5, $\partial\rho$ may only be contained in P_0 if $\partial\rho = \rho \cap P_0$. If ∂P_0 is contained in ρ, then since ∂P_0 is essential in K_ρ, we similarly conclude that either ∂P_0 is also essential in ρ or $\partial P_0 = \partial\rho$ after an isotopy. However since ρ contains no
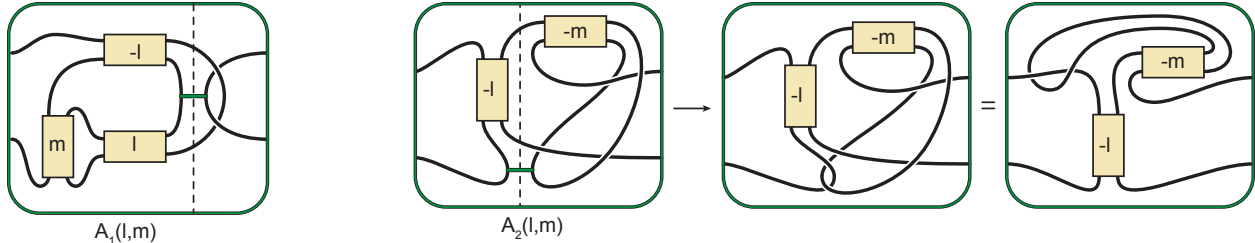
FIGURE 8. Left: EM-tangle $A_1(\ell, m)$ with its summing disk and standard unknotting arc. Right: EM-tangle $A_2(\ell, m)$ with its summing disk and standard unknotting arc, followed by the unknotting crossing change along this arc, and an isotopy to more visibly be a rational tangle. (EM-tangels are shown up to adding integral tangles.)

essential Conway spheres by Claim 4.5, this implies $\partial P_0 = \partial \rho$. Since we have already ruled out the possibility that $P_0 = \rho$, we conclude that $P_0 = T$ if $\partial P_0 \subset \rho$. Thus P_0 must be contained within T.

Finally, observe that if ∂P_0 is not an essential Conway sphere in T, the either ∂P is isotopic to ∂T through a product tangle in T or there is a compressing disk d for ∂P_0 in T disjoint from the strands. In the former, we may then assume $P_0 = T$. In the latter, the disk d would need to become trivial upon inclusion of T into $K_\rho = N(\rho + T)$ since ∂P_0 is essential in K_ρ; yet this cannot occur since ∂d does not bound a disk in ∂P_0 disjoint from the strands. Hence either ∂P_0 is an essential Conway sphere in T or $P_0 = T$. □

If P_0 is contained in T where ∂P_0 is an essential Conway sphere, then by Proposition 4.1 there is an essential summing disk D_0 for T that is disjoint from P_0. Assume D_0 decomposes T into the sum $T_1 + T_2$ and, say, that $P_0 \subset T_2$. While α must intersect D_0 as an arc in T by Lemma 4.4, there must be an isotopy of α through K_ρ to bring it to an arc α_0 in $P_0 \subset T_2$ since we are in case (2). The crossing change on α then becomes a crossing change on α_0 taking T_2 to a tangle T_2'. Thus the crossing changes on α_0 takes $K_\rho = N(\rho + T_1 + T_2)$ to $N(\rho + T_1 + T_2')$ which should be an unknot. However as ρ is a non-trivial tangle whose numerator closure is the unknot by Claim 4.5, we must have $T_1 + T_2'$ be the 0–tangle. Yet this requires both T_1 and T_2' to be integral tangles by Corollary 2.8. As D_0 is an essential summing disk, T_1 is not an integral tangle, a contradiction.

If $P_0 = T$ then the summing disk of P_0 must be a summing disk of T. Since T is homotopic to the 0–tangle, so must be P_0. Thus P_0 cannot be an EM-tangle $A_1(\ell, m)$ since they are homotopic to tangles disjoint from their summing disks and hence to the ∞–tangle. (See Figure 8.) Thus P_0 must be an EM-tangle $A_2(\ell, m)$. Since T admits a crossing change to the 0–tangle, P_0 must also admit a crossing change to the 0–tangle. In particular, the standard unknotting arc for P_0 must change P_0 to the 0–tangle. However, the unknotting crossing change along the standard unknotting arc for an EM-tangle $A_2(\ell, m)$ produces the rational tangle $[0, -\ell, -m, -2] = \frac{2m-1}{\ell(2m-1)-2}$. This is integral only if $\ell(2m-1) - 2 = \pm 1$ which implies that either $\ell = \pm 1$ or $m = 0, 1$. Yet these choices for ℓ or m make the summing disk of P_0 inessential so that P_0 is not a non-trivial tangle sum. This is a contradiction. □

4.2. Sums of non-integral tangles admit no proper RTRs to integral tangles.

Proposition 4.9. *Let T be a marked tangle sum of tangles, neither of which is integral. Then T admits no proper rational tangle replacement to an integral tangle.*

Proof. This proof follows similarly to that of Proposition 4.6 which proves the result for crossing changes. Since proper rational tangle replacements have distance $\Delta \geq 2$ and those of distance $\Delta = 2$ are realized as crossing changes, we may restrict attention to proper rational tangle replacements of distance $\Delta > 2$.

As in the proof of Proposition 4.6, assume for contradiction that a tangle sum $T = T_1 + T_2$ admits a proper rational tangle replacement (of distance $\Delta > 2$) along an arc α to the 0–tangle. Then, again using the tangle ρ from Figure 7, we may form the knot $K_\rho = N(\rho + T)$ which the proper rational tangle replacement transforms into the unknot $K_\rho^* - N(\rho + 0)$. Let α^* be the arc on K_ρ^* dual to α (ie. the core arc of the replacing rational tangle). Hence K_ρ is a knot with proper rational unknotting number 1 along the arc $\alpha \subset T$. As before, $\partial \rho$ is an essential Conway sphere for K_ρ.

14

Now we shall show that α cannot be isotoped to be disjoint from all essential Conway spheres for K_ρ. This argument follows exactly the same as **Case (1)** of Proposition 4.6 for crossing changes using that

- a proper rational tangle replacement also preserves the connectivity of the strands (ie. it preserves the homotopy type of the rational tangle), and
- $D(\rho)$ actually has proper rational unknotting number ≥ 2 by Claim 4.5 instead of just the regular unknotting number ≥ 2.

Now, instead of [GL06, Theorem 6.2], we must apply [GL06, Theorem 5.2] to the double branched cover of $X = K_\rho \backslash \mathcal{N}(\alpha) = K_\rho^* \backslash \mathcal{N}(\alpha^*)$, the exterior of the arc α on the knot K_ρ, which is also the exterior of the arc α^* on the unknot K_ρ^*. In particular, since K_ρ^* is the unknot, α^* lifts to a knot k in S^3 whose exterior is $X = S^3 \backslash \mathcal{N}(k)$. In the context of [GL06, Theorem 5.2], $S^3 = X(\mu)$ and the double branched cover of K_ρ is $X(\gamma)$. Then the distance $\Delta(\mu, \gamma)$ between the curves μ and γ in ∂X is the distance between the Dehn fillings, this agrees with the distance Δ of the rational tangle replacement. Therefore, since $\Delta > 2$, only cases (1) and (4) of [GL06, Theorem 5.2] apply. However, in each of these two cases, the knot that is the lift of α is disjoint from the essential tori of $X(\gamma)$. Since we have shown that α cannot be isotoped to be disjoint from all of the essential Conway spheres of K_ρ, its lift cannot be isotoped to be disjoint from all of the essential tori of $X(\gamma)$ that are the lifts of these essential Conway spheres. This contradicts [GL06, Theorem 5.2], and thus there can be no proper rational tangle replacement from K_ρ to the unknot K_ρ^*. Consequently, T admits no proper rational tangle replacement to an integral tangle. $\qquad\square$

References

[ASZ⁺92]   David E. Adams, Eugene M. Shekhtman, E. Lynn Zechiedrich, Molly B. Schmid, and Nicholas R. Cozzarelli. The role of topoisomerase IV in partitioning bacterial replicons and the structure of catenated intermediates in DNA replication. *Cell*, 71(2):277–288, 10 1992.

[BG21]   Agnese Barbensi and Dimos Goundaroulis. f-distance of knotoids and protein structure. *Proc. A.*, 477(2246):Paper No. 20200898, 18, 2021.

[CDGW]   Marc Culler, Nathan M. Dunfield, Matthias Goerner, and Jeffrey R. Weeks. SnapPy, a computer program for studying the geometry and topology of 3-manifolds. Available at http://snappy.computop.org (05/11/2021).

[Con70]   J. H. Conway. An enumeration of knots and links, and some of their algebraic properties. In *Computational Problems in Abstract Algebra (Proc. Conf., Oxford, 1967)*, pages 329–358. Pergamon, Oxford, 1970.

[EMn97]   Mario Eudave-Muñoz. Non-hyperbolic manifolds obtained by Dehn surgery on hyperbolic knots. In *Geometric topology (Athens, GA, 1993)*, volume 2 of *AMS/IP Stud. Adv. Math.*, pages 35–61. Amer. Math. Soc., Providence, RI, 1997.

[GDBS17]   Dimos Goundaroulis, Julien Dorier, Fabrizio Benedetti, and Andrzej Stasiak. Studies of global and local entanglements of individual protein chains using the concept of knotoids. *Scientific Reports*, 7(1):6309, 2017.

[GGL⁺17]   Dimos Goundaroulis, Neslihan Gügümcü, Sofia Lambropoulou, Julien Dorier, Andrzej Stasiak, and Louis Kauffman. Topological models for open-knotted protein chains using the concepts of knotoids and bonded knotoids. *Polymers*, 9(9), 2017.

[GI02]   William Gibson and Masaharu Ishikawa. Links and Gordian numbers associated with generic immersions of intervals. *Topology Appl.*, 123(3):609–636, 2002.

[GL06]   C. McA. Gordon and John Luecke. Knots with unknotting number 1 and essential Conway spheres. *Algebr. Geom. Topol.*, 6:2051–2116, 2006.

[ILM21]   Damian Iltgen, Lukas Lewark, and Laura Marino. Khovanov homology and rational unknotting, 2021.

[Kaw98]   Tomomi Kawamura. The unknotting numbers of 10_{139} and 10_{152} are 4. *Osaka J. Math.*, 35(3):539–546, 1998.

[KL04]   Louis H. Kauffman and Sofia Lambropoulou. On the classification of rational tangles. *Adv. in Appl. Math.*, 33(2):199–237, 2004.

[Lac21]   Marc Lackenby. Links with splitting number one. *Geom. Dedicata*, 214:319–351, 2021.

[LM21]   Charles Livingston and Allison H. Moore. Knotinfo: Table of knot invariants. URL: knotinfo.math.indiana.edu, November 2021.

[LMRH+12]   V. Lopez, M. L. Martinez-Robles, P. Hernandez, D. B. Krimer, and J. B. Schvartzman. Topo IV is the topoisomerase that knots and unknots sister duplexes during DNA replication. *Nucleic Acids Res.*, 40(8):3563–3573, Apr 2012.

[MSY82]   William Meeks, III, Leon Simon, and Shing Tung Yau. Embedded minimal surfaces, exotic spheres, and manifolds with positive Ricci curvature. *Ann. of Math. (2)*, 116(3):621–659, 1982.

[MT11]   Sergei Matveev and Vladimir Turaev. A semigroup of theta-curves in 3-manifolds. *Mosc. Math. J.*, 11(4):805–814, 822, 2011.

[Sch85]   Martin G. Scharlemann. Unknotting number one knots are prime. *Invent. Math.*, 82(1):37–55, 1985.

[Sch90]   Martin Scharlemann. Producing reducible 3-manifolds by surgery on a knot. *Topology*, 29(4):481–500, 1990.

[SdlCMR+98]   D. Santamaria, G. de la Cueva, M. L. Martinez-Robles, D. B. Krimer, P. Hernandez, and J. B. Schvartzman. DnaB helicase is unable to dissociate RNA-DNA hybrids. Its implication in the polar pausing of replication forks at ColE1 origins. *J. Biol. Chem.*, 273(50):33386–33396, Dec 1998.

[Tur12]   Vladimir Turaev. Knotoids. *Osaka J. Math.*, 49(1):195–223, 2012.

[VHK+96]   E. Viguera, P. Hernandez, D. B. Krimer, A. S. Boistov, R. Lurz, J. C. Alonso, and J. B. Schvartzman. The ColE1 unidirectional origin acts as a polar replication fork pausing site. *J. Biol. Chem.*, 271(37):22414–22421, Sep 1996.

[Wol87]   Keith Wolcott. The knotting of theta curves and other graphs in S^3. In *Geometry and topology (Athens, Ga., 1985)*, volume 105 of *Lecture Notes in Pure and Appl. Math.*, pages 325–346. Dekker, New York, 1987.

[Zha91]   Xingru Zhang. Unknotting number one knots are prime: a new proof. *Proc. Amer. Math. Soc.*, 113(2):611–612, 1991.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MIAMI
Email address: k.baker@math.miami.edu

DEPARTMENT OF MATHEMATICS & DEPARTMENT OF BIOLOGY, DUKE UNIVERSITY
Email address: dbuck@math.duke.edu

DEPARTMENT OF MATHEMATICS, MARYMOUNT UNIVERSITY
Email address: dodonnol@marymount.edu

DEPARTMENT OF MATHEMATICS & APPLIED MATHEMATICS, VIRGINIA COMMONWEALTH UNIVERSITY
Email address: moorea14@vcu.edu

DEPARTMENT OF MATHEMATICS, COLBY COLLEGE
Email address: scott.taylor@colby.edu