

# TANGLE STRUCTURE TREES

HANNO VON BERGEN AND REINHARD DIESTEL

ABSTRACT. We introduce a comprehensive data structure, tangle structure trees, which simultaneously displays all the  $\mathcal{F}$ -tangles of an abstract separation system for very general obstruction sets  $\mathcal{F}$ . It simultaneously also displays certificates  $\sigma \in \mathcal{F}$  for any non-existence of such tangles, or for the non-extendability of low-order tangles to higher-order ones.

Our theorem can be applied to produce the structures of the classical tree-of-tangles and tangle-tree duality theorems, both for graph tangles and for their known generalizations to more general separation systems. It extends those theorems to obstruction sets  $\mathcal{F}$  that need not define profiles (as they must in trees of tangles) or consist of stars of separations (as they must in tangle-tree duality).

Our existence proof for these structure trees is constructive. The construction has been implemented in open-source software available for tangle detection and further analysis.

## 1. INTRODUCTION

The notion of ‘tangles’ was originally introduced by Robertson and Seymour [20] as an abstract concept of high local connectivity in graphs, one that unifies several more concrete such notions, such as highly connected subgraphs or minors. What all these notions have in common is that, given any low-order separation of the graph, any highly cohesive substructure must lie mostly on one of its two sides: since the separation has low order, it cannot split it into two roughly equal halves. A tangle remembers only how all the low-order separations are oriented in this way, each towards that highly cohesive substructure: the collection of all these oriented separations is called a *tangle*. See [9] for a precise definition and basic facts about graph tangles.

Tangles have since been generalized to more general settings than graphs. In the setting of set partitions, they offer a precise theoretical basis for ‘fuzzy’ real-world problems such as clustering in large datasets [8]. All these, including graphs, are special cases of so-called *abstract separation systems* [7]: an axiomatic setting that assumes only the most basic properties of graph separations. This is the setting in which tangles are most easily and comprehensively treated, and we shall use this framework also in this paper.

There are two fundamental theorems about tangles, which both have their origins in [20] and are also treated in [9]. The *tree-of-tangles* theorem exposes a tree-like structure in the graph or dataset whose tangles we consider, which ‘distinguishes’ the tangles in that they are shown to ‘live in’ different areas of this tree-like structure. The other is the *tangle-tree duality* theorem, which certifies the non-existence of a tangle by exposing another tree-like structure in the graph or dataset, one

whose existence clearly precludes the existence of a tangle.<sup>1</sup> This is the more fundamental of the two theorems, in the sense that the other can be reduced to its abstract version (see below) but not vice versa [17].

Both the above theorems have been generalized to  $\mathcal{F}$ -tangles of abstract separation systems, ways of simultaneously orienting the ‘separations’ of a given structure in such a way that no three of those oriented separations form an element of  $\mathcal{F}$ .<sup>2</sup> This collection  $\mathcal{F}$  has to satisfy some constraints that depend on which of the two theorems we are considering.

For the tree-of-tangles theorem, which displays all the  $\mathcal{F}$ -tangles of a separation system simultaneously,  $\mathcal{F}$  has to include triples of oriented separations that are reminiscent of ultrafilters. If our separations are bipartitions of a fixed set  $V$ , for example, then  $\mathcal{F}$  must contain all triples of the form  $\{A, B, \overline{A \cap B}\}$ : if  $A$  and  $B$  are subsets of  $V$  deemed ‘large’ (in that the  $\mathcal{F}$ -tangle orients those partitions towards them), then the complement of  $A \cap B$  cannot also be ‘large’. Such  $\mathcal{F}$ -tangles are called *profiles*, and all existing tree-of-tangle theorems for  $\mathcal{F}$ -tangles assume that these are profiles [1–4, 6, 11, 12, 18].

For the tangle-tree duality theorem,  $\mathcal{F}$  has to satisfy another constraint: its elements have to be *stars*, nested sets of oriented separations pointing towards each other [7]. The tree-like structures by which tangle-tree duality theorems certify the non-existence of  $\mathcal{F}$ -tangles require that  $\mathcal{F}$  consist of stars: if it does not, these structures cannot be tree-like.

This state of the art leaves  $\mathcal{F}$ -tangles that are not profiles without any known way of displaying all those tangles simultaneously. And if  $\mathcal{F}$  does not consist of stars, it leaves separation systems that have no  $\mathcal{F}$ -tangles without any known way of organizing the elements of  $\mathcal{F}$  into a data structure that displays them as easily checkable certificates for the non-existence of such tangles.

The structure trees whose existence we prove in this paper do both these things: they display all the  $\mathcal{F}$ -tangles of a separation system even when they are not profiles, and for those orientations of the separation system that are not  $\mathcal{F}$ -tangles they display certificates in  $\mathcal{F}$  that show why they are not.<sup>3</sup>

Tangle structure trees display all this information in a single, comprehensive, data structure, which is maximally efficient in the following, structural, sense. A single tangle is most efficiently displayed by listing just those of its elements that are minimal in the poset of oriented separations that comes with a separation system. Indeed, all these are needed to determine the tangle, but any other elements can be deduced from them. Our structure trees display all the  $\mathcal{F}$ -tangles of a separation system  $S$  simultaneously by listing only the oriented separations that are minimal in one of those tangles, including the tangles of subsystems consisting of separations of lower order. And for orientations of  $S$  that are not  $\mathcal{F}$ -tangles it displays certificates from  $\mathcal{F}$  inside those same minimal subsets of oriented separations, in a way readily accessible in the structure tree.

---

<sup>1</sup>In a graph, this would be a tree-decomposition into parts too small for a tangle to live in.

<sup>2</sup>In the case of graph tangles,  $\mathcal{F}$  consists of the sets of up to three oriented separations of the graph such that the union of the ‘back’ sides of these three separations covers the entire graph [9].

<sup>3</sup>In a follow-up paper [13] we show that if  $\mathcal{F}$  does consist of stars, our duality theorem implies the same tangle-tree dichotomy as the classical one. And we show how our structure trees give rise to trees of tangles if the  $\mathcal{F}$ -tangles happen to be profiles. Our main result thus includes the structures of both classical tangle theorems as special cases.

Our paper is organized as follows. In [Section 2](#) we provide the necessary background for tangles in abstract separation systems, the framework in which we shall construct our structure trees and prove their existence.

In [Section 3](#) we introduce basic tangle structure trees. In [Section 4](#) we prove their existence, and in [Section 5](#) we show how to make them efficient. In [Section 6](#) we collect all this information together to prove our main results, the ‘ $\mathcal{F}$ -tangle structure theorems’. In [Theorem 6.4](#) we note specifically how the main structure theorem provides certificates from  $\mathcal{F}$  for the non-existence of  $\mathcal{F}$ -tangles if there are none.

In [Section 7](#), finally, we apply our results to three particular types of  $\mathcal{F}$ -tangles. The first two of these include  $\mathcal{F}$ -tangles in graphs: those induced by  $k$ -blocks, and those that are profiles. In neither of these does the defining set  $\mathcal{F}$  consist of stars. As the third application we include  $\mathcal{F}$ -tangles that encode clusters in large datasets. These are neither profiles nor do their obstruction sets  $\mathcal{F}$  consist of stars.

## 2. TANGLES BASICS

In this section we give precise definitions and notation for abstract tangles, largely following [\[7\]](#) and indicating any deviations.<sup>4</sup>

Tangles of graphs are ways of orienting their separations, each towards one of its two sides. Abstract tangles are designed to work in scenarios where there need not be anything to ‘separate’. In order to retain our intuition from graphs, however, we continue to refer to the things of which our abstract tangles pick one of two variants (which they will indeed do) as ‘separations’. These are defined by noting some key properties of graph separations and making them into axioms, as follows.

A *separation system*  $(\vec{S}, \leq, *)$  is a set  $\vec{S}$ , whose elements we call *oriented separations*, that comes with a partial ordering  $\leq$  on  $\vec{S}$  and an order-reversing involution  $*: \vec{S} \rightarrow \vec{S}$ . Thus, for any two elements<sup>5</sup>  $\vec{r}, \vec{s}$  of  $\vec{S}$  with  $\vec{r} \leq \vec{s}$  we have  $\vec{r}^* \geq \vec{s}^*$ . We write  $\vec{s}^* =: \vec{s}$ , and call  $\vec{s}$  the *inverse* of  $\vec{s}$ . While we allow formally that  $\vec{s} = \vec{s}$ , in which case we call  $\vec{s}$  and  $s$  *degenerate*, this does not happen often in practice.<sup>6</sup>

If a separation system  $\vec{U}$  happens to be a lattice, that is, if there is a supremum  $\vec{r} \vee \vec{s}$  and an infimum  $\vec{r} \wedge \vec{s}$  in  $\vec{U}$  for every two elements  $\vec{r}, \vec{s} \in \vec{U}$ , we call  $\vec{U}$  a *universe* of separations. It is *distributive* if it is distributive as a lattice. A separation system  $\vec{S} \subseteq \vec{U}$  is *submodular* if for every two elements of  $\vec{S}$  either their infimum or their supremum in  $\vec{U}$  also lies in  $\vec{S}$ .

Very rarely we may have separations  $\vec{s} \leq \vec{s}$ ; then  $\vec{s}$  is *small* and  $\vec{s}$  is *large*.<sup>7</sup> We say that  $\vec{s}$  is *trivial* (and  $\vec{s}$  is *co-trivial*) in  $\vec{S}$  if there exists a pair of inverse  $\vec{r}, \vec{r} < \vec{s}$  in  $\vec{S}$ . Trivial separations are clearly large, so co-trivial ones are small, but the converse need not hold. See [\[7\]](#) for more on these technicalities if desired.

The set of *unoriented separations* in  $(\vec{S}, \leq, *)$  is

$$S := \{\{\vec{s}, \vec{s}\} : \vec{s} \in \vec{S}\}.$$

---

<sup>4</sup>The most important difference is that, for historical reasons, the partial ordering on  $\vec{S}$  used in [\[7\]](#) is the inverse of ours. So terms like ‘large’ and ‘small’, infima and suprema etc, are reversed.

<sup>5</sup>We often denote the elements of  $\vec{S}$  by letters with an arrow, in either direction, precisely in order to have a simple way to refer to their dual elements: by reversing the arrow. But the arrow directions have no meaning: an arbitrary element of  $\vec{S}$  could be denoted equally as  $\vec{s}$  or as  $\vec{s}$ .

<sup>6</sup>The only degenerate separation of a graph  $G = (V, E)$ , for example, is  $(V, V)$ .

<sup>7</sup>The small separations of a graph  $G$  are those of the form  $(V, A)$  with  $A \subseteq V = V(G)$ .

We call the elements  $\vec{s}, \bar{s}$  of  $s$  its *orientations*. An *orientation* of  $S$  is a set  $\tau \subseteq \vec{S}$  that contains exactly one orientation of every  $s \in S$ . For  $s \in S$  we then denote by  $\tau(s)$  the unique orientation of  $s$  contained in  $\tau$ . An orientation of a subset of  $S$  is a *partial orientation* of  $S$ .

If  $\vec{r} \geq \vec{s}$  we say that  $\vec{r}$  *points towards*  $s$  (and that  $\bar{r}$  *points away from*  $s$ ). We say that  $\vec{r}$  *points towards* an oriented separation  $\vec{s}$  whenever it points towards  $s$ , i.e., if  $\vec{r} \geq \vec{s}$  or  $\vec{r} \geq \bar{s}$ , and similarly for ‘points away from’. A *star* is a set  $\sigma$  of non-degenerate oriented separations that point towards each other. As is easy to check, this happens if and only if  $\vec{r} \geq \vec{s}$  (and hence  $\bar{s} \geq \bar{r}$ ) for all distinct  $\vec{r}, \vec{s} \in \sigma$ .

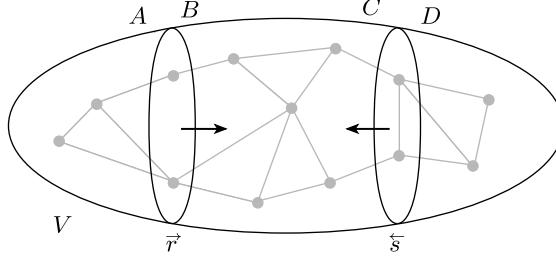


FIGURE 1. Nested separations  $r = \{A, B\}$  and  $s = \{C, D\}$  of a graph. Their orientations  $\vec{r} = (A, B)$  and  $\bar{s} = (D, C)$  point towards each other, since  $\vec{r} \geq \vec{s}$  (as  $B \supseteq D$ ) and  $\bar{s} \geq \bar{r}$  (as  $C \supseteq A$ ).

Two separations  $s, r \in S$  are *nested* if they have orientations that are comparable under  $\leq$ . Oriented separations are *nested* if their underlying unoriented separations are nested. A subset of  $\vec{S}$  is *nested* if its elements are pairwise nested.

A subset of  $\vec{S}$  is *consistent* if no pair of its elements  $\vec{r}, \vec{s}$  with  $r \neq s$  point away from each other. Stars are examples of consistent nested sets of oriented separations.

If  $\sigma \subseteq \vec{S}$  is consistent, we say that  $\vec{s} \in \vec{S}$  is *required* by  $\sigma$  if  $\vec{s} \notin \sigma$  and  $\sigma \cup \{\vec{s}\}$  is inconsistent. We shall see in [Lemma 2.2](#) that, pathological cases aside,  $\sigma \cup \{\vec{s}\}$  will then be consistent. The *closure* of  $\sigma$  is

$$[\sigma] := \sigma \cup \{\vec{s} \in \vec{S} : \vec{s} \text{ is required by } \sigma\}.$$

Note that  $\sigma$  requires  $\vec{s} \notin \sigma$  if and only if there exists an  $\vec{r} \in \sigma$  such that  $r \neq s$  and  $\vec{s} > \vec{r}$ . Thus

$$[\sigma] = \sigma \cup \{ \vec{s} \in \vec{S} : \exists \vec{r} \in \sigma \text{ such that } r \neq s \text{ and } \vec{s} > \vec{r} \},$$

which motivates the notation of  $[\sigma]$ . In particular,  $[[\sigma]] = [\sigma]$ , which justifies the term (upward) ‘closure’.

If  $\sigma$  contains no small separations, the expression above simplifies to

$$[\sigma] = \{ \vec{s} \in \vec{S} : \exists \vec{r} \in \sigma \text{ such that } \vec{r} \leq \vec{s} \}.$$

Indeed, any  $\vec{s}$  in this latter set either lies in  $\sigma$  or there is some  $\vec{r} \in \sigma$  such that  $\vec{r} < \vec{s}$ ; in that case  $r \neq s$ , since otherwise  $\vec{s} = \vec{r} < \vec{s}$ , making  $\vec{r} \in \sigma$  is small.

**Lemma 2.1.** *Let  $\sigma \subseteq \tau \subseteq \vec{S}$  be consistent sets. Then  $[\sigma] \subseteq [\tau]$ . If  $\tau$  is an orientation of all of  $S$ , then  $[\tau] = \tau$ .*

*Proof.* If  $\vec{s}$  is required by  $\sigma$ , then  $\sigma \cup \{\vec{s}\}$  is inconsistent. Since  $\sigma \subseteq \tau$ , it follows that  $\tau \cup \{\vec{s}\}$  is also inconsistent; hence  $\vec{s}$  is either an element of or required by  $\tau$ . Thus every element of  $[\sigma]$  lies in  $[\tau]$ , proving the first claim.

If  $\tau$  is an orientation of all of  $S$  then  $[\tau] = \tau$ , since for any  $\vec{s} \in [\tau] \setminus \tau$  the set  $\tau \cup \{\vec{s}\} = \tau$  would be inconsistent.  $\square$

Consistent orientations of  $S$  can contain small separations. But examples are rare and can be counter-intuitive, so we often exclude them. Co-trivial separations cannot occur in consistent orientations of  $S$ . Indeed if  $\bar{s}$  is co-trivial, witnessed by  $r \in S$ , then every orientation of  $S$  will have to orient  $r$  too. But it cannot do so consistently with  $\bar{s}$ , since both  $\vec{r}$  and  $\bar{r}$  are inconsistent with  $\bar{s}$ . Similarly,  $[\{\bar{s}\}]$  is then inconsistent, since it contains  $\vec{r}$  and  $\bar{r}$ , which are both inconsistent with  $\bar{s}$ .

**Lemma 2.2.** *Let  $\tau \subseteq \vec{S}$  be consistent, and suppose that  $\tau$  has no elements that are co-trivial in  $\vec{S}$ . Then  $[\tau]$  is consistent, and  $[\tau] \setminus \tau$  contains at most one orientation of any  $s \in S$ .*

*Proof.* If  $[\tau]$  is inconsistent, it contains orientations  $\vec{r}, \vec{s}$  of distinct  $r, s \in S$  with  $\vec{s} < \vec{r}$ . Since  $\tau$  is consistent, at least one of  $\vec{r}, \vec{s}$  is not in  $\tau$ . Without loss of generality assume  $\vec{r} \notin \tau$ : Then  $\vec{r} \in [\tau] \setminus \tau$ , so  $\vec{r}$  is required by  $\tau$  (by definition of  $[\tau]$ ), which means that  $\tau \cup \{\vec{r}\}$  is inconsistent. Thus there exists  $\vec{t} \in \tau$  with  $t \neq r$  and  $\bar{r} < \bar{t}$ . We thus have  $\bar{s} < \bar{r} < \bar{t}$ , that is  $\bar{s}$  and  $\bar{t}$  point away from each other.

If  $\bar{s} = \bar{t}$ , the above inequality yields  $\vec{r}, \bar{r} < \bar{t}$ , which makes  $\bar{t} \in \tau$  co-trivial, contradicting our assumptions. Hence  $\bar{s} \neq \bar{t}$  as well as  $\bar{s} \neq \bar{t}$ , giving  $s \neq t$ . The fact that  $\bar{s}$  and  $\bar{t} \in \tau$  point away from each other thus implies that  $\{\bar{s}, \bar{t}\}$  is inconsistent, so  $\bar{s} \notin \tau$  as well as  $\bar{t} \notin \tau$ .

As earlier with  $\vec{r}$ , the fact that  $\bar{s} \in [\tau]$  now implies that  $\vec{r} < \bar{s} < \bar{t}$  for some  $\bar{t} \in \tau$ . The two inequalities together now give  $\bar{t} < \bar{r} < \bar{t}$ , so  $\bar{t}$  and  $\bar{t}$  point away from each other. As  $\tau$  is consistent and contains both, this means that  $t' = t$ . As  $\bar{t} < \bar{t}$ , we thus have  $\bar{t} = \bar{t}$ . But now  $\bar{r} < \bar{t} = \bar{t}$  as well as  $\bar{r} < \bar{t}$ . This makes  $\bar{t} \in \tau$  co-trivial in  $\vec{S}$ , contradicting our assumptions. This proves the first assertion.

For the second, note that  $\tau$  requires any  $\bar{s} \in [\tau] \setminus \tau$ , by definition of  $[\tau]$ . This means that  $\tau \cup \{\bar{s}\}$  is inconsistent, so  $\bar{s} \notin [\tau]$  since  $[\tau]$  is consistent.  $\square$

An *order function* on  $S$  is any map  $S \rightarrow \mathbb{R}$ . Unless otherwise mentioned, we denote such order functions as  $s \mapsto |s|$ . We extend them to  $\vec{S}$  by letting  $|\vec{s}| := |\bar{s}| := |s|$ . Given an order function on  $S$  and  $k \in \mathbb{R}$ , let

$$\vec{S}_k := \{\vec{s} \in \vec{S} : |s| < k\};$$

this is again a separation system.

We sometimes require that the orientations of  $S$  shall not have certain subsets. We typically collect those together in some set  $\mathcal{F}$ , whose elements we call *forbidden subsets*. Formally, if  $\mathcal{F}$  is any set, we say that a subset of  $\vec{S}$  *avoids*  $\mathcal{F}$  if it has no subset in  $\mathcal{F}$ , i.e., which is an element of  $\mathcal{F}$ .

**Definition 2.3.** An  $\mathcal{F}$ -tangle of  $S$  is an  $\mathcal{F}$ -avoiding consistent orientation of  $S$ . The  $\mathcal{F}$ -tangles of the subsets  $S_k$  of  $S$  are the  $\mathcal{F}$ -tangles in  $\vec{S}$ .

### 3. TANGLE STRUCTURE TREES

We adopt the graph-theoretic terminology of [9]. A *tree* is a connected acyclic graph. Given nodes  $t, t'$  of a tree  $T$  we write  $tTt'$  for the unique  $t$ - $t'$  path in  $T$ . A *rooted tree* is a tree with a distinguished node called its *root*. Given a tree  $T$  with root  $r$ , we define a partial order  $\leq_r$  on  $V(T)$  by declaring  $x \leq_r y$  if  $x$  lies on the path in  $T$  from  $r$  to  $y$ . Maximal elements, including the root if  $|T| = 1$ , are called *leaves*. Any direct successors of a node in  $\leq_r$  are its *children*. We write  $E_v \subseteq E(T)$  for the set of edges from  $v$  to its children.

Let  $(\vec{S}, \leq, *)$  be a separation system, and let  $\mathcal{F}$  be any set.

**Definition 3.1.** A *separation tree*  $(T, r, \beta)$  on  $\vec{S}$  consists of a rooted tree  $(T, r)$  together with an edge labeling  $\beta: E(T) \rightarrow \vec{S}$  such that for every non-leaf  $v \in V(T)$  there exists a separation  $s_v \in S$  such that  $\beta$  restricts to a bijection  $E_v \rightarrow \{\vec{s}_v, \bar{s}_v\}$  and  $s_u \neq s_v$  whenever  $u <_r v$ .

Thus, every non-leaf node  $v$  of a separation tree has either one or two children. If there is no need to refer to  $r$  or  $\beta$  explicitly, we usually abbreviate  $(T, r, \beta)$  to  $T$ . For every node  $v$ , we write  $\beta_v$  for the set  $\beta(E(rTv))$  of edge labels on the path  $rTv$ .

**Lemma 3.2.** *Let  $(T, r, \beta)$  be a separation tree on  $\vec{S}$ . Every orientation  $\tau$  of  $S$  contains  $\beta_\ell$  as a subset for some unique leaf  $\ell \in V(T)$ .*

*Proof.* Since  $\beta_r = \emptyset \subseteq \tau$ , there exists a node  $v \in V(T)$  that is maximal with respect to  $\leq_r$  subject to  $\beta_v \subseteq \tau$ . If  $v$  is not a leaf, let  $v'$  be its child with  $\beta(vv') = \tau(s_v) =: \vec{s}$ . As  $\beta_v \subseteq \tau$ , and  $\vec{s} \in \tau \setminus \beta_v$  since  $s_u \neq s_v$  for all  $u < v$ , we see that  $\beta_{v'} = \beta_v \cup \{\vec{s}\} \subseteq \tau$  contradicts the maximality of  $v$ . Hence  $v =: \ell$  is our desired leaf.

For a proof that  $\ell$  is unique, let  $\ell'$  be another leaf and let  $v$  be the greatest common ancestor of  $\ell$  and  $\ell'$ . Then  $v$  has distinct children  $w$  and  $w'$  with  $\beta(vw) \in \beta_\ell$  and  $\beta(vw') \in \beta_{\ell'}$ . Since  $\beta(vw) \neq \beta(vw')$ , they cannot both lie in  $\tau$ . As  $\beta(vw) \in \beta_\ell \subseteq \tau$  we must have  $\beta(vw') \notin \tau$ , and hence  $\beta_{\ell'} \not\subseteq \tau$ .  $\square$

A separation tree  $T$  is *consistent* if  $\beta_v \subseteq \vec{S}$  is consistent for every node  $v \in T$ .

**Lemma 3.3.** *Let  $(T, r, \beta)$  be a consistent separation tree on  $\vec{S}$ , and let  $v$  be a non-leaf node of  $T$ . Then  $\vec{s}_v$  and  $\bar{s}_v$  are  $\leq$ -minimal in  $\beta_v \cup \{\vec{s}_v\}$  and in  $\beta_v \cup \{\bar{s}_v\}$ , respectively.*

*Proof.* Suppose not. Without loss of generality assume there exists  $\vec{s} \in \beta_v$  with  $\vec{s} < \vec{s}_v$ . Let  $w$  be the child of  $v$  with  $\beta(vw) = \bar{s}_v$ . Then  $\{\vec{s}, \bar{s}_v\} \subseteq \beta_w$  is inconsistent, contradicting the consistency of  $(T, r, \beta)$ .  $\square$

**Corollary 3.4.** *For every consistent separation tree and non-leaf node  $v$  we have  $\{\vec{s}_v, \bar{s}_v\} \cap \lfloor \beta_v \rfloor = \emptyset$ . That is,  $s_v$  is not oriented by  $\lfloor \beta_v \rfloor$ .*

*Proof.* Suppose, say, that  $\vec{s}_v \in \lfloor \beta_v \rfloor$ . Since  $s_u \neq s_v$  for all nodes  $u < v$ , we know that  $\vec{s}_v \notin \beta_v$ . Our assumption that  $\vec{s}_v \in \lfloor \beta_v \rfloor$  thus means that  $\vec{s}_v$  is required by  $\beta_v$ . Hence there exists  $\vec{r} \in \beta_v$  such that  $\vec{r} < \vec{s}_v$  (and  $r \neq s_v$ ). This contradicts Lemma 3.3.  $\square$

Lemma 3.3 and Corollary 3.4 imply that consistent separation trees offer an efficient data structure for storing consistent orientations of  $S$ , in particular,  $\mathcal{F}$ -tangles. Indeed, any such orientation  $\tau$  labels the tree's edges between the root and some leaf  $\ell$ , but this  $\beta_\ell$  is only a small subset of  $\tau$ : elements  $\vec{s}$  of  $\tau$  required by  $\beta_\ell$  for some  $v < \ell$  will not appear in  $\beta_\ell \setminus \beta_v$ . But they can be reconstructed from  $\beta_v$  by the consistency of  $\tau$ .

By definition of separation trees, the separations labelling their edges do not repeat along any root-to-leaf path: if  $u < v$  then  $s_u \neq s_v$ . Hence the length of any path from the root to a leaf is at most  $|S|$ . Since every node has at most two children, this bounds the size of the tree:

**Corollary 3.5.** *Separation trees on  $\vec{S}$  have at most  $2^{|S|}$  leaves and fewer than  $2^{|S|+1}$  nodes.*  $\square$

A leaf  $\ell$  of a consistent separation tree on  $\vec{S}$  is an  $\mathcal{F}$ -tangle leaf if the closure  $[\beta_\ell]$  of  $\beta_\ell$  is an  $\mathcal{F}$ -tangle of  $S$ . A leaf  $\ell$  is *forbidden* (by  $\mathcal{F}$ ) if  $\beta_\ell$  contains a member of  $\mathcal{F}$  as a subset. Note that  $\mathcal{F}$ -tangle leaves are never forbidden.

**Definition 3.6.** An  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$  is a consistent separation tree on  $\vec{S}$  in which every leaf is either an  $\mathcal{F}$ -tangle leaf or forbidden, and for every non-leaf  $v$  the set  $\beta_v$  has no subset in  $\mathcal{F}$ .

Every  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$  displays all the  $\mathcal{F}$ -tangles of  $S$ , not just some of them:

**Theorem 3.7.** Let  $\vec{S}$  be a separation system, let  $\mathcal{F}$  be any set, and let  $T$  be any  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$ .

- (i) For every  $\mathcal{F}$ -tangle  $\tau$  of  $S$  there is a leaf  $\ell$  of  $T$  such that  $[\beta_\ell] = \tau$ .
- (ii) If all the leaves of  $T$  are forbidden, then  $S$  has no  $\mathcal{F}$ -tangle.

*Proof.* (i) By [Lemma 3.2](#) there exists a leaf  $\ell$  with  $\beta_\ell \subseteq \tau$ . This leaf is not forbidden, since  $\beta_\ell \subseteq \tau$  has no subset in  $\mathcal{F}$ . As  $T$  is an  $\mathcal{F}$ -tangle structure tree, this means that  $\ell$  is an  $\mathcal{F}$ -tangle leaf: that  $[\beta_\ell]$  is an  $\mathcal{F}$ -tangle of  $S$ . As  $[\beta_\ell] \subseteq \tau$ , this tangle can only be  $\tau$ .

(ii) is immediate from (i). □

In the next section we shall determine the sets  $\mathcal{F}$  for which  $S$  admits an  $\mathcal{F}$ -tangle structure tree whose tree-order is compatible with a given order function on  $S$ , in that lower-order separations label edges further down on the tree. As we shall see in [Theorem 6.3](#), such trees display not only all the tangles of the entire set  $S$ , as all  $\mathcal{F}$ -tangle structure trees do by [Theorem 3.7](#), but all the  $\mathcal{F}$ -tangles in  $\vec{S}$ : the  $\mathcal{F}$ -tangles of the sets  $S_k = \{s \in S : |s| < k\}$  with  $k \in \mathbb{N}$ .

#### 4. EXISTENCE OF TANGLE STRUCTURE TREES

Let  $\vec{S}$  be a separation system with an order function  $||$  on  $S$ . Let  $\mathcal{F}$  be any set. We shall need two conditions on  $\mathcal{F}$  to ensure that  $\vec{S}$  has an  $\mathcal{F}$ -tangle structure tree.

The first is that  $\{\vec{s}\} \in \mathcal{F}$  for every  $\vec{s} \in \vec{S}$  that is trivial in  $\vec{S}$ . This holds for all sets  $\mathcal{F}$  of interest, and if it does we call  $\mathcal{F}$  *standard for*  $\vec{S}$ . Intuitively, co-trivial separations  $\vec{s}$  point to places in our graph or structure that are tiny – too small to house any tangle. More formally, we already saw in [Section 2](#) that no consistent orientation of  $S$  can contain co-trivial separations. Assuming that  $\mathcal{F}$  is standard for  $\vec{S}$  therefore places no restrictions on the  $\mathcal{F}$ -tangles of  $S$ .

The second condition has more substance. If  $T$  is any separation tree on  $\vec{S}$  then, by [Lemma 3.2](#), every orientation  $\tau$  of  $S$  contains  $\beta_\ell$  for a unique leaf  $\ell$ . If  $\tau$  is consistent and  $[\beta_\ell]$  orients all of  $S$ , then  $[\beta_\ell] = \tau$  by [Lemma 2.1](#). If  $\tau$  is not an  $\mathcal{F}$ -tangle, then  $\tau = [\beta_\ell]$  has a subset  $\sigma$  in  $\mathcal{F}$  that witnesses this. If  $T$  is in fact an  $\mathcal{F}$ -tangle structure tree, we know that our consistent orientation  $\tau$  of  $S$  is an  $\mathcal{F}$ -tangle (without having to assume that  $[\beta_\ell]$  orients all of  $S$ ) unless it contains such a set  $\sigma \in \mathcal{F}$  not only in  $[\beta_\ell]$  but even in  $\beta_\ell$ : among the edge labels of  $T$ .

In order for this to be possible, we therefore need to make some ‘richness’ assumption about  $\mathcal{F}$ : an assumption which, in our example, ensures that  $\mathcal{F}$  has enough elements to contain a subset also of  $\beta_\ell$  as soon as it contains a subset of  $[\beta_\ell]$ .

We shall identify an essentially weakest-possible such richness condition below, in [Definition 4.3](#). This will need some more preparation. For motivation, readers are

invited to peek at [Lemma 4.4](#) and the definition preceding it now. That definition clearly implies that  $\mathcal{F}$  contains a subset of  $\beta_\ell$  whenever it contains a subset of  $\lfloor \beta_\ell \rfloor$ . The notion of richness from [Definition 4.3](#) will still imply this, but is weaker.

A separation tree  $T$  on  $\vec{S}$  with an order function  $\|$  on  $S$  is *ordered* if  $|s_v| \leq |s_w|$  whenever  $v$  and  $w$  are non-leaves of  $T$  with  $v \leq w$ . It is *thoroughly ordered* (in  $\vec{S}$ ) if, for every non-leaf node  $v$ , the separation  $s_v$  is not oriented by  $\lfloor \beta_v \rfloor$ <sup>8</sup> and has minimum order among the separations in  $S$  not oriented by  $\lfloor \beta_v \rfloor$ .

**Lemma 4.1.** *Every thoroughly ordered separation tree is ordered.*

*Proof.* Let  $T$  be a thoroughly ordered separation tree on  $\vec{S}$ . If it is not ordered, it has nodes  $v, w$  with  $v < w$  such that  $|s_v| > |s_w|$ . Since  $s_w$  is not oriented by  $\lfloor \beta_v \rfloor \subseteq \lfloor \beta_w \rfloor$ , this contradicts the requirement that  $s_v$  have minimum order among the separations not oriented by  $\lfloor \beta_v \rfloor$ .  $\square$

We say that  $\vec{s} \in \vec{S}$  is *weakly eclipsed* by  $\vec{r} \in \vec{S}$  if  $\vec{r} < \vec{s}$  and  $|r| \leq |s|$ , and *eclipsed* by  $\vec{r}$  if  $\vec{r} < \vec{s}$  and  $|r| < |s|$ . Given any set  $\tau \subseteq \vec{S}$ , a subset  $\sigma \subseteq \tau$  is *efficient* (in  $\tau$ ) if no element of  $\sigma$  is eclipsed by any other element of  $\tau$ . It is *strongly efficient* if no element of  $\sigma$  is weakly eclipsed by any other element of  $\tau$ . Note that if the order function on  $S$  is injective and  $\tau$  is a partial orientation of  $S$ , then every efficient subset of  $\tau$  is strongly efficient in  $\tau$ .<sup>9</sup>

**Lemma 4.2.** *Let  $v$  be a node of a thoroughly ordered separation tree  $T$  on  $\vec{S}$ . Then*

- (i) *every strongly efficient subset of  $\lfloor \beta_v \rfloor$  is contained in  $\beta_v$ ;*
- (ii) *if  $\beta_v$  is consistent, it is efficient in any partial orientation  $\tau$  of  $S$  that includes  $\lfloor \beta_v \rfloor$ .*

*Proof.* (i) Let  $\sigma$  be any strongly efficient subset of  $\lfloor \beta_v \rfloor$ , and let  $\vec{s} \in \sigma$  be given. Suppose that  $\vec{s} \notin \beta_v$ . Then  $\vec{s} \in \lfloor \beta_v \rfloor \setminus \beta_v$  is required by  $\beta_v$ , so there exists  $\vec{r} \in \beta_v$  with  $\vec{r} < \vec{s}$  (and  $r \neq s$ ). As  $\vec{r} \in \beta_v$ , we have  $r = s_u$  for some  $u < v$ . Choose such an  $\vec{r}$  with  $u$  minimal in  $\prec_r$ .

Our aim is to show that  $|r| \leq |s|$ : then  $\vec{r} \in \lfloor \beta_v \rfloor$  eclipses  $\vec{s} \in \sigma$  weakly, contradicting the strong efficiency of  $\sigma$  as a subset of  $\lfloor \beta_v \rfloor$ . Since  $T$  is thoroughly ordered, we shall have  $|r| \leq |s|$  as desired if  $\lfloor \beta_u \rfloor$  does not orient  $s$ : this would make  $s$  a candidate for  $s_u$ , so  $|r| > |s|$  would contradict the fact that  $r = s_u$ .

So let us show that neither  $\vec{s}$  nor  $\vec{r}$  lies in  $\lfloor \beta_u \rfloor$ . For  $\vec{s} \notin \lfloor \beta_u \rfloor$ , recall first that  $\vec{s} \notin \beta_v \supseteq \beta_u$ . Hence if  $\vec{s} \in \lfloor \beta_u \rfloor$ , there exists an  $\vec{r}' \in \beta_u$  such that  $\vec{r}' < \vec{s}$ . This satisfies  $r' = s_{u'}$  for some  $u' <_r u$ , so  $\vec{r}' < \vec{s}$  contradicts our original choice of  $\vec{r}$  given  $\vec{s}$ . Hence  $\vec{s} \notin \lfloor \beta_u \rfloor$  as desired. Suppose now that  $\vec{s} \in \lfloor \beta_u \rfloor$ . As  $\vec{s} > \vec{r}$ , any  $\vec{r}' \leq \vec{s}$  in  $\beta_u$  satisfies  $\vec{r}' \leq \vec{s} < \vec{r}$ . So  $r$  was already oriented by  $\lfloor \beta_u \rfloor$ , contradicting the fact that  $r = s_u$ . Hence neither  $\vec{s}$  nor  $\vec{r}$  lies in  $\lfloor \beta_u \rfloor$ , as desired.

(ii) Suppose not; then some  $\vec{r} \in \tau$  eclipses some  $\vec{s} \in \beta_v$ . Let  $u < v$  be such that  $s = s_u$ . As  $|r| < |s|$  and  $T$  is thoroughly ordered, we know that  $\lfloor \beta_u \rfloor$  orients  $r$ .

Since  $\tau$  is a partial orientation of  $S$  containing  $\vec{r}$ , it cannot also contain  $\vec{r}$ . We thus cannot have  $\vec{r} \in \lfloor \beta_u \rfloor$ , since  $\lfloor \beta_u \rfloor \subseteq \lfloor \beta_v \rfloor \subseteq \tau$  by [Lemma 2.1](#). So  $\vec{r} \in \lfloor \beta_u \rfloor$ . As  $\vec{r} < \vec{s}$ , this implies  $\vec{s} \in \lfloor \beta_u \rfloor$  unless  $\vec{r} = \vec{s} \in \beta_u$ . Both these contradict the fact that  $s = s_u$ .  $\square$

<sup>8</sup>If  $T$  is consistent, then this holds by [Corollary 3.4](#).

<sup>9</sup>We need the assumption on  $\tau$ , since  $\vec{s}$  eclipses  $\vec{r}$  weakly if  $\vec{s} < \vec{r}$ . So we do not want  $\vec{s}, \vec{r} \in \tau$ .

A typical application of [Lemma 4.2 \(ii\)](#) is that  $\beta_v$  is efficient in any tangle of  $S$  that contains it. Such a tangle will be consistent and hence include  $[\beta_v]$  by [Lemma 2.1](#).

A typical application of [Lemma 4.2 \(i\)](#) will be that witnesses  $\sigma \in \mathcal{F}$  to the fact that some  $[\beta_v]$  fails to extend to an  $\mathcal{F}$ -tangle can be found not only in  $[\beta_v]$  but in  $\beta_v$  itself. Then our structure trees will display such witnesses  $\sigma \in \mathcal{F}$  in the label sets  $\beta_\ell$  of their forbidden leaves  $\ell$ . For this to work with the help of [Lemma 4.2](#), we need  $\mathcal{F}$  to contain witnesses that are strongly efficient in their  $[\beta_v]$ .

This motivates our formal richness condition on  $\mathcal{F}$ :

**Definition 4.3.** A set  $\mathcal{F}$  is *rich for  $\vec{S}$*  if every consistent orientation of  $S$  that has a subset in  $\mathcal{F}$  also has a strongly efficient<sup>10</sup> subset in  $\mathcal{F}$ .

The assumption that  $\mathcal{F}$  is rich will be central to our results. It is needed to ensure that our  $\mathcal{F}$ -tangle structure trees exist, and we shall see after [Lemma 4.7](#) that no weaker condition on  $\mathcal{F}$  will ensure the same.

Although it may look a bit technical at first glance, this ‘richness’ requirement on  $\mathcal{F}$  is quite natural, given the role of these  $\mathcal{F}$  in tangle theory. For example, given a consistent orientation  $\tau$  of  $S$  that has a subset  $\sigma \in \mathcal{F}$ , we can often obtain a strongly efficient subset  $\sigma' \in \mathcal{F}$  of  $\tau$  simply by replacing every element  $\vec{s}$  of  $\sigma$  by some  $\vec{s}' \leq \vec{s}$  that is minimal in  $\tau$ . One still has to check then that  $\sigma'$  is indeed in  $\mathcal{F}$ . But as the idea behind those forbidden triples in  $\mathcal{F}$  is that they identify areas in our graph or other structure that are ‘too small to be home to a tangle’, it is not unnatural for this particular  $\sigma'$  to be in  $\mathcal{F}$  if  $\sigma$  was.

Let us cast this example in the form of a lemma. Let us call a set  $\mathcal{F}$  *closed under minimization in  $\vec{S}$*  if it contains every set  $\sigma' \subseteq \vec{S}$  obtained from some  $\sigma \subseteq \vec{S}$  in  $\mathcal{F}$  by replacing every  $\vec{s} \in \sigma$  with some  $\vec{s}' \leq \vec{s}$ .

**Lemma 4.4.** *If  $\mathcal{F}$  is closed under minimization in  $\vec{S}$ , then  $\mathcal{F}$  is rich for  $\vec{S}$ .*

*Proof.* Let  $\tau$  be any consistent orientation of  $S$  that has a subset  $\sigma$  in  $\mathcal{F}$ . We have to find a set  $\sigma' \subseteq \tau$  in  $\mathcal{F}$  that is strongly efficient in  $\tau$ .

Let  $\sigma'$  be obtained from  $\sigma$  by replacing every  $\vec{s} \in \sigma$  by some  $\vec{s}' \leq \vec{s}$  that is minimal in  $\tau$ . The set  $\sigma'$  is strongly efficient in  $\tau$ , since any  $\vec{r} < \vec{s}' \in \sigma'$  in  $\tau$  contradicts the minimal choice of  $\vec{s}' \leq \vec{s}$ . Since  $\mathcal{F}$  is closed under minimization, we have  $\sigma' \in \mathcal{F}$  as required.  $\square$

**Lemma 4.5.** *Let  $\mathcal{F}$  be rich and standard for  $\vec{S}$ . Let  $T$  be a thoroughly ordered separation tree on  $\vec{S}$ , and let  $v \in V(T)$ . Assume that  $\beta_v$  is consistent and avoids  $\mathcal{F}$ , and that  $[\beta_v]$  orients all of  $S$ . Then  $[\beta_v]$  is an  $\mathcal{F}$ -tangle of  $S$ .*

*Proof.* As  $\beta_v$  avoids  $\mathcal{F}$ , which is standard,  $\beta_v$  has no element that is co-trivial in  $\vec{S}$ . As  $\beta_v$  is consistent,  $[\beta_v]$  is consistent by [Lemma 2.2](#). Since  $[\beta_v]$  orients all of  $S$ , it can thus only fail to be an  $\mathcal{F}$ -tangle of  $S$  if it has a subset  $\sigma$  in  $\mathcal{F}$ . As  $\mathcal{F}$  is rich for  $\vec{S}$ , we can choose  $\sigma$  to be strongly efficient as a subset of  $[\beta_v]$ .<sup>11</sup> Then even  $\sigma \subseteq \beta_v$  by [Lemma 4.2 \(i\)](#), which contradicts our assumption that  $\beta_v$  avoids  $\mathcal{F}$ . Hence  $[\beta_v]$  is an  $\mathcal{F}$ -tangle of  $S$ .  $\square$

<sup>10</sup>in this orientation of  $S$

<sup>11</sup>If  $[\beta_v]$  contains both orientations of some  $s \in S$ , delete one of them to obtain a consistent orientation of  $S$  as required in the definition of ‘rich’ for  $\mathcal{F}$ .

**Theorem 4.6.** *Let  $\vec{S}$  be a separation system equipped with an order function on  $S$ , and let  $\mathcal{F}$  be a set that is rich and standard for  $\vec{S}$ . Then there exists a thoroughly ordered  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$ .*

*Proof.* Start with the one-node tree  $T_0 = \{r\}$  and let  $\beta^0 = \emptyset$ . We iteratively build trees  $T_0 \subsetneq T_1 \subsetneq \dots$  with maps  $\beta^i$  so that each  $(T_i, r, \beta^i)$  is consistent and thoroughly ordered, and none of their sets  $\beta_v^i$  with  $v$  a non-leaf node will have a subset in  $\mathcal{F}$ . Note that  $T_0$  has all these properties. The last of those trees will be our desired  $\mathcal{F}$ -tangle structure tree.

If for some  $n$  the tree  $(T_n, r, \beta^n)$  is already an  $\mathcal{F}$ -tangle structure tree, we are done. Otherwise  $T_n$  has a leaf  $v$  which is neither an  $\mathcal{F}$ -tangle leaf nor forbidden. By Lemma 4.5,  $\lfloor \beta_v^n \rfloor$  does not orient all of  $S$ ; let  $s \in S$  be a separation of minimum order not oriented by  $\lfloor \beta_v^n \rfloor$ .

Form  $T_{n+1}$  by adding two children  $v_1, v_2$  at  $v$ . Let  $\beta^{n+1}$  agree with  $\beta^n$  on the edges of  $T_n$ , and pick orientations  $\vec{s} =: \beta^{n+1}(vv_1)$  and  $\bar{s} =: \beta^{n+1}(vv_2)$  of  $s$ ; then  $s = s_v$  in  $T_{n+1}$ . By construction,  $(T_{n+1}, r, \beta^{n+1})$  is a thoroughly ordered separation tree. It is consistent, because  $T_n$  was and neither orientation of  $s$  lies in  $\lfloor \beta_v^{n+1} \rfloor$ . Its only non-leaf node  $v$  that was not already a non-leaf node in  $T_n$  is  $v$ . Since  $v$  was not forbidden as a leaf of  $T_n$ , the set  $\beta_v^{n+1} = \beta_v^n$  has no subset in  $\mathcal{F}$ .

This process strictly increases  $|V(T_n)|$  at each step, but by Corollary 3.5 there is an upper bound on the size of any consistent separation tree in terms of  $|S|$ . Hence the process terminates after finitely many steps with a thoroughly ordered  $\mathcal{F}$ -tangle structure tree.  $\square$

If our order function is injective, Theorem 4.6 has a converse, which shows that our requirement of richness for  $\mathcal{F}$  is weakest possible to ensure the existence of an  $\mathcal{F}$ -tangle structure tree. This is established by our next lemma:

**Lemma 4.7.** *Let  $\vec{S}$  be a separation system with an injective order function on  $S$ , and let  $\mathcal{F}$  be any set. If there exists a thoroughly ordered  $\mathcal{F}$ -tangle structure tree  $T$  of  $\vec{S}$ , then  $\mathcal{F}$  is rich for  $\vec{S}$ .*

*Proof.* Let  $\tau$  be any consistent orientation of  $\vec{S}$  that has a subset in  $\mathcal{F}$ . We shall find an efficient subset  $\sigma \in \mathcal{F}$  of  $\tau$ , which will even be strongly efficient since our order function is injective. By Lemma 3.2,  $T$  has a leaf  $\ell$  with  $\beta_\ell \subseteq \tau$ . In particular,  $\beta_\ell$  is consistent, so  $\lfloor \beta_\ell \rfloor \subseteq \tau$  by Lemma 2.1.

Since  $T$  is an  $\mathcal{F}$ -tangle structure tree,  $\ell$  is either an  $\mathcal{F}$ -tangle leaf or forbidden. If  $\ell$  is an  $\mathcal{F}$ -tangle leaf then  $\lfloor \beta_\ell \rfloor$  is an  $\mathcal{F}$ -tangle of all of  $S$ . But then  $\lfloor \beta_\ell \rfloor = \tau$ , which contradicts the fact that  $\tau$  has a subset in  $\mathcal{F}$ .

Thus  $\ell$  is forbidden, so  $\beta_\ell$  has a subset  $\sigma$  in  $\mathcal{F}$ . By Lemma 4.2 (ii) this  $\sigma$  is efficient in  $\tau$ , as desired.  $\square$

Note that if our order function on  $S$  is not injective, the proof of Lemma 4.7 still goes through as stated except for one aspect: the efficient subset  $\sigma \in \mathcal{F}$  of  $\tau$  it finds may not be strongly efficient in  $\tau$  (as our definition of ‘rich’ requires).

If our order function is injective, the sets  $\mathcal{F}$  for which  $\vec{S}$  admits a thoroughly ordered  $\mathcal{F}$ -tangle structure tree are thus precisely the rich ones:

**Theorem 4.8.** *Let  $\vec{S}$  be a separation system with an injective order function on  $S$ , and let  $\mathcal{F}$  be a set that is standard for  $\vec{S}$ . There exists a thoroughly ordered  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$  if and only if  $\mathcal{F}$  is rich for  $\vec{S}$ .*

*Proof.* Immediate from [Theorem 4.6](#) and [Lemma 4.7](#).  $\square$

In [Section 5](#) we shall prove that our  $\mathcal{F}$ -tangle structure trees can be improved further, without any additional requirements on  $\mathcal{F}$ , by contracting ‘inessential’ edges: edges whose label  $\vec{s}$  is neither needed in any sets  $\sigma \in \mathcal{F}$  witnessing that a leaf is forbidden, nor needed to determine any tangles.

Unfortunately, this contraction process can cause our structure trees to lose their property of being thoroughly ordered. We shall therefore extract from this property its essence, a slightly weaker property to be called ‘efficiency’, which is still strong enough to make the branches  $rT\ell$  and their label sets  $\beta_\ell$  as efficient for displaying the tangles  $[\beta_\ell]$  as they are in thoroughly ordered structure trees, but weak enough to survive the contraction process we envisage for [Section 5](#).

Given a separation system  $\vec{S}$  with an order function on  $S$ , we call a separation tree  $(T, r, \beta)$  *efficient* if for every leaf  $\ell$  the set  $\beta_\ell$  is efficient in  $[\beta_\ell]$ .

**Lemma 4.9.** *Let  $\vec{S}$  be a separation system with an order function on  $S$ . Then every thoroughly ordered separation tree  $(T, r_0, \beta)$  on  $\vec{S}$  is efficient.*

*Proof.* Consider any leaf  $\ell$  of  $T$ . If  $\beta_\ell$  is not efficient in  $[\beta_\ell]$  as claimed, then some  $\vec{s} \in \beta_\ell$  is eclipsed by some  $\vec{r} \in [\beta_\ell]$ . Let  $v, w < \ell$  be such that  $\vec{s} = \vec{s}_v$  and  $\vec{r} \geq \vec{s}_w \in \beta_\ell$ .<sup>12</sup>

As  $\vec{r}$  eclipses  $\vec{s}$ , we have  $\vec{r} < \vec{s}$  and  $|r| < |s|$ ; in particular,  $r \neq s = s_v$ . The fact that  $T$  is thoroughly ordered thus implies that  $r$  is already oriented by  $[\beta_v]$ . We cannot have  $\vec{r} \in [\beta_v]$ , since this would place  $\vec{s}_v = \vec{s} > \vec{r}$  or  $\vec{s}_v$  in  $[\beta_v]$  too,<sup>13</sup> which would contradict the fact that  $T$  is thoroughly ordered. Thus,  $\vec{r} \in [\beta_v]$ ; let  $u < v$  be such that  $\vec{r} \geq \vec{s}_u \in \beta_v$ .

If  $u = w$ , then  $\vec{s}_u = \vec{s}_w \leq \vec{r} < \vec{s}$  and hence  $\vec{s}_v = \vec{s} \in [\beta_v]$  or  $\vec{s}_v \in \beta_v$ ,<sup>13</sup> both of which contradict the fact that  $T$  is thoroughly ordered.

If  $u < w$ , then  $\vec{s}_w \geq \vec{r} \geq \vec{s}_u \in \beta_w$ . Then  $\vec{s}_w \in [\beta_w]$  unless  $\vec{s}_w \in \beta_w$ ,<sup>13</sup> which both contradict the fact that  $T$  is thoroughly ordered.

If  $w < u$ , finally, then  $\vec{s}_u \geq \vec{r} \geq \vec{s}_w \in \beta_u$ , so  $\vec{s}_u \in [\beta_u]$  unless  $\vec{s}_u \in \beta_u$ ,<sup>13</sup> which both contradict the fact that  $T$  is thoroughly ordered.  $\square$

## 5. IRREDUCIBLE AND EFFICIENT TANGLE STRUCTURE TREES

Let  $(T, r, \beta)$  be a separation tree on a separation system  $\vec{S}$ . Let  $v$  be a node of  $T$  with a child  $w$ . Define  $(T, r, \beta)_{w \rightarrow v} := (T_{w \rightarrow v}, r_{w \rightarrow v}, \beta_{w \rightarrow v})$ , where  $T_{w \rightarrow v}$  is obtained from  $T$  by contracting the edge  $vw$  of  $T$  and deleting any other child of  $v$  together with its subtree.

We continue to use ‘ $w$ ’ for the new node constructed from the edge  $vw$ , and thus think of  $V(T_{w \rightarrow v})$  as a subset of  $V(T)$ . Similarly, we think of  $E(T_{w \rightarrow v})$  as a subset of  $E(T)$  and continue to use ‘ $\beta$ ’ to denote its labelling  $\beta_{w \rightarrow v}$ . If  $v = r$  we set  $r_{w \rightarrow v} := w$ ; otherwise we keep  $r_{w \rightarrow v} := r$ .

Note that  $T_{w \rightarrow v}$  inherits the sets  $E_u$  from  $T$  for its nodes  $u$ . In particular, the separations  $s_u$  remain unchanged for all  $u \in V(T_{w \rightarrow v}) \subseteq V(T)$ . Thus,  $T_{w \rightarrow v}$  is still

<sup>12</sup>Recall that our arrow notation for oriented separations is never fixed. We are thus free to use forward arrows to denote the orientations of  $s_v$  and  $s_w$  that lie in  $\beta_\ell$ .

<sup>13</sup>By the definition of  $[\cdot]$ , the assumption of  $\vec{s} > \vec{r} \in [\beta_v]$  implies  $\vec{s} \in [\beta_v]$  only if  $\vec{s} \notin \beta_v$ .

ordered if  $T$  was. However,  $T_{w \rightarrow v}$  may no longer be thoroughly ordered in  $\vec{S}$  even if  $T$  was.

The sets  $\beta_u$  remain unchanged for all  $u \not\geq w$ , while for all  $u \geq w$  the sets  $\beta_u$  in  $T_{w \rightarrow u}$  arise from  $\beta_u$  in  $T$  by deleting  $\beta(vw)$  from it. (Recall that  $\beta$  was injective on  $E(rTu)$ , by [Definition 3.1](#).) In particular,  $T_{w \rightarrow v}$  is consistent if  $T$  was.

We say that  $\vec{s} \in \vec{S}$  is *necessary* for an  $\mathcal{F}$ -tangle leaf  $\ell$  if it is a  $\leq$ -minimal element of  $\beta_\ell$ . We say that  $\vec{s} \in \vec{S}$  is *necessary* for a forbidden leaf  $\ell$  if every subset of  $\beta_\ell$  in  $\mathcal{F}$  contains  $\vec{s}$ .

**Lemma 5.1.** *Let  $\vec{S}$  be a separation system, and let  $\mathcal{F}$  be any set. Let  $(T, r_0, \beta)$  be an  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$ , and let  $\ell$  be a leaf of  $T$ . A separation  $\vec{s} \in \vec{S}$  is necessary for  $\ell$  if and only if  $\beta_\ell \setminus \{\vec{s}\}$  has no subset in  $\mathcal{F}$  and  $[\beta_\ell \setminus \{\vec{s}\}]$  is not an  $\mathcal{F}$ -tangle of  $S$ .*

*Proof.* Suppose first that  $\vec{s}$  is necessary for  $\ell$ , and that  $\ell$  is an  $\mathcal{F}$ -tangle leaf. Then  $[\beta_\ell]$  is an  $\mathcal{F}$ -tangle of  $S$ , so  $\beta_\ell \setminus \{\vec{s}\}$  has no subset in  $\mathcal{F}$ .

For a proof that  $[\beta_\ell \setminus \{\vec{s}\}]$  is not an  $\mathcal{F}$ -tangle of  $S$ , we show that it contains neither  $\vec{s}$  nor  $\vec{s}$ . To see  $\vec{s} \notin [\beta_\ell \setminus \{\vec{s}\}]$ , note that any  $\vec{r} \in \beta_\ell \setminus \{\vec{s}\}$  with  $\vec{r} \leq \vec{s}$  would satisfy  $\vec{r} < \vec{s}$ , since also  $\vec{s} \in \beta_\ell$  and hence  $r \neq s$  by [Definition 3.1](#). Then  $\vec{s}$  would not be minimal in  $\beta_\ell$ , which it is by assumption. And  $\vec{s} \notin [\beta_\ell \setminus \{\vec{s}\}]$ , since  $[\beta_\ell]$  is a tangle of  $S$  containing  $\vec{s}$ . Thus,  $[\beta_\ell \setminus \{\vec{s}\}]$  does not orient  $s$ , so it is not an  $\mathcal{F}$ -tangle of  $S$ .

Suppose, second, that  $\vec{s}$  is necessary for  $\ell$ , and that  $\ell$  is a forbidden leaf. Then  $\beta_\ell \setminus \{\vec{s}\}$  has no subset in  $\mathcal{F}$ . Suppose  $[\beta_\ell \setminus \{\vec{s}\}]$  is an  $\mathcal{F}$ -tangle of  $S$ . This tangle cannot contain  $\vec{s}$ : it would then contain the entire set  $\beta_\ell$ , but this has a subset in  $\mathcal{F}$  since  $\ell$  is a forbidden leaf. But neither can  $[\beta_\ell \setminus \{\vec{s}\}]$  contain  $\vec{s}$ : then  $\beta_\ell \setminus \{\vec{s}\}$  would contain some  $\vec{r} \leq \vec{s}$ , which would contradict the consistency of  $\beta_\ell$ , since  $r \neq s$  by [Definition 3.1](#), and so  $\vec{s} \in \beta_\ell$  would point away from  $\vec{r} \in \beta_\ell$ .

Suppose, third, that  $\vec{s}$  is not necessary for  $\ell$  and that  $\ell$  is an  $\mathcal{F}$ -tangle leaf. Then  $[\beta_\ell]$  is an  $\mathcal{F}$ -tangle and  $\vec{s}$  is not a minimal element of  $\beta_\ell$ . We complete our proof in this case by showing that  $[\beta_\ell \setminus \{\vec{s}\}] = [\beta_\ell]$ . If  $\vec{s} \notin \beta_\ell$  this holds trivially. If  $\vec{s}$  lies in  $\beta_\ell$  but is not minimal in it, then there exists some  $\vec{r} \in \beta_\ell$  with  $\vec{r} < \vec{s}$ . This implies  $\vec{s} \in [\beta_\ell \setminus \{\vec{s}\}]$ , and hence

$$[\beta_\ell] = [\beta_\ell \setminus \{\vec{s}\}] \cup \{[\vec{s}]\} \subseteq [[\beta_\ell \setminus \{\vec{s}\}]] = [\beta_\ell \setminus \{\vec{s}\}]$$

as desired, unless  $\vec{s} \in [\beta_\ell \setminus \{\vec{s}\}]$ . But in that case the tangle  $[\beta_\ell]$  contains both  $\vec{s}$  and  $\vec{s}$ , which it cannot.

Suppose finally that  $\vec{s}$  is not necessary for  $\ell$  and that  $\ell$  is a forbidden leaf. Then  $\beta_\ell \setminus \{\vec{s}\}$  has a subset in  $\mathcal{F}$ , as desired.  $\square$

A node  $v$  is called *necessary* in  $T$  if for every child  $w$  of  $v$  there exists a leaf  $\ell \geq w$  such that  $\beta(vw)$  is necessary for  $\ell$ .

**Lemma 5.2.** *Let  $\vec{S}$  be a separation system, and let  $\mathcal{F}$  be any set. Let  $(T, r, \beta)$  be an  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$ , and let  $v$  be a node of  $T$ . Then  $v$  has a child  $w$  such that  $(T, r, \beta)_{w \rightarrow v}$  is an  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$  if and only if  $v$  is not necessary in  $T$ .*

*Proof.* Suppose first that  $v$  is necessary in  $T$ . If  $v$  has no children, the intended implication holds. If  $v$  does have a child,  $w$  say, there is a leaf  $\ell \geq w$  such that  $\vec{s} := \beta(vw)$  is necessary for  $\ell$ . Then  $\ell$  is a leaf both of  $T$  and of  $T_{w \rightarrow v}$ . By [Lemma 5.1](#), however,  $\ell$  is neither an  $\mathcal{F}$ -tangle leaf nor a forbidden leaf of  $T_{w \rightarrow v}$ . Hence  $T_{w \rightarrow v}$  is not an  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$ .

Conversely, suppose  $v$  is not necessary in  $T$ . Then it has a child  $w$  such that  $\beta(vw)$  is not necessary for any leaf  $\ell \geq w$  in  $T$ . By [Lemma 5.1](#),  $T_{w \rightarrow v}$  is an  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$ .  $\square$

A tangle structure tree  $(T, r, \beta)$  is called *irreducible* if every node of  $T$  is necessary.

**Theorem 5.3.** *Let  $\vec{S}$  be a separation system, and let  $\mathcal{F}$  be any set. For every  $\mathcal{F}$ -tangle structure tree  $T$  of  $\vec{S}$  there exists an irreducible  $\mathcal{F}$ -tangle structure tree  $T'$  of  $\vec{S}$  obtained from  $T$  by a sequence  $T = T^1, \dots, T^n = T'$  with  $T^{i+1} = T_{w \rightarrow v}^i$  for suitable nodes  $v, w$  of  $T^i$ .*

*The tree  $T'$  imposes the same partial ordering on its node as  $T$  does. In particular,  $T'$  is ordered if  $T$  is. Its leaves  $\ell$  are also leaves of  $T$ , and their incident edges in  $T$  are still edges of  $T'$ . Every leaf  $\ell$  of  $T'$  is forbidden in  $T'$  if and only if it is forbidden in  $T$ .*

*Proof.* If  $T^i$  is not irreducible, pick a node  $v \in T^i$  that is not necessary. By [Lemma 5.2](#) it has a child  $w$  for which  $T_{w \rightarrow v}^i$  is an  $\mathcal{F}$ -tangle structure tree. Let  $T^{i+1} := T_{w \rightarrow v}^i$ . For the second paragraph of the theorem, note that separations of the form  $\beta(v\ell)$ , with  $\ell$  a leaf, are necessary in every  $\mathcal{F}$ -tangle structure tree: by [Corollary 3.4](#) if  $\ell$  is an  $\mathcal{F}$ -tangle leaf, and by [Definition 3.6](#) if  $\ell$  is forbidden.  $\square$

If  $\mathcal{F}$  is standard, the  $\mathcal{F}$ -tangle structure trees we considered in [Section 4](#) were all efficient, because they were thoroughly ordered ([Lemma 4.9](#)). The irreducible structure trees we obtained from them in [Theorem 5.3](#) may no longer be thoroughly ordered: as noted before, this property can get lost in the contraction process. But the efficiency of the original structure trees is maintained:

**Lemma 5.4.** *If  $T'$  is obtained from an efficient  $\mathcal{F}$ -tangle structure tree  $T$  as in [Theorem 5.3](#), then  $T'$  too is efficient.*

*Proof.* Consider  $T^{i+1} = T_{w \rightarrow v}^i$  as in [Theorem 5.3](#) for  $i = 1, \dots, n-1$  in turn. Assuming that  $T^i$  is efficient, we have to show that so is  $T^{i+1}$ . Write  $\beta^{i+1} := \beta_{w \rightarrow v}^i$ , where  $\beta^i$  is the edge labelling of  $T^i$ . Given any leaf  $\ell$  of  $T^{i+1}$ , we have to show that  $\beta_\ell^{i+1}$  is efficient in  $[\beta_\ell^{i+1}]$ . This follows from the fact that  $\ell$  is also a leaf of  $T^i$ , the assumed efficiency of  $T^i$ , and the fact that  $\beta_\ell^{i+1} \subseteq \beta_\ell^i$ .  $\square$

## 6. THE $\mathcal{F}$ -TANGLE STRUCTURE THEOREMS

In this section we summarize what we have shown so far, in a few, concise, statements that we think of as the main results of this paper.

The first of these tells us when  $\mathcal{F}$ -tangle structure trees exist:

**Theorem 6.1.** *Let  $\vec{S}$  be a separation system with an order function on  $S$ . If a set  $\mathcal{F}$  is standard and rich for  $\vec{S}$ , then  $\vec{S}$  has an efficient and irreducible ordered  $\mathcal{F}$ -tangle structure tree.*

*Proof.* By [Theorem 4.6](#),  $\vec{S}$  has a thoroughly ordered  $\mathcal{F}$ -tangle structure tree  $(T, r, \beta)$ . By [Lemma 4.9](#) it is efficient. The desired structure tree can be obtained from  $(T, r, \beta)$  by [Theorem 5.3](#); it is efficient by [Lemma 5.4](#).  $\square$

The  $\mathcal{F}$ -tangle structure trees constructed for the proof of [Theorem 6.1](#) have a host of properties designed to help with the tangle analysis of a given separation system. It might have been natural to list these properties as part of [Theorem 6.1](#).

However, since the latter is an existence theorem, this would have allowed us to establish these properties just for the structure tree we constructed in its proof.

We would like to stress the fact that those properties are common to all efficient and irreducible ordered  $\mathcal{F}$ -tangle structure trees: these terms were designed to encode precisely those properties. Let us show this next:

**Theorem 6.2.** *Let  $\vec{S}$  be a separation system with an order function. Every efficient ordered  $\mathcal{F}$ -tangle structure tree  $(T, r, \beta)$  for  $\vec{S}$  has the following properties.*

- (i) *For every leaf  $\ell$  of  $T$ , the set  $\beta_\ell$  is consistent in  $S$ . If  $\beta_\ell$  has no subset in  $\mathcal{F}$ , then  $[\beta_\ell]$  is an  $\mathcal{F}$ -tangle of  $S$ , and every  $\mathcal{F}$ -tangle of  $S$  arises in this way.*
- (ii) *For every non-leaf node  $v$  of  $T$ , the set  $\beta_v$  has no subset in  $\mathcal{F}$ , and  $s_v$  is not oriented by  $[\beta_v]$ . Both its orientations  $\vec{s} \in \{\vec{s}_v, \vec{s}'_v\}$  are minimal in  $\beta_v \cup \{\vec{s}\}$  in the partial ordering of  $\vec{S}$ . Its order  $|s|$  is maximal in  $\{|s_u| : u \leq_r v\}$ .*
- (iii) *For every orientation  $\tau$  of  $S$  there is a unique leaf  $\ell$  of  $T$  such that  $\beta_\ell \subseteq \tau$ . If  $\tau$  is consistent, then  $[\beta_\ell] \subseteq \tau$  and  $\beta_\ell$  is efficient in  $\tau$ .*
- (iv) *For every  $\mathcal{F}$ -tangle  $\tau$  of  $S$  there is a unique leaf  $\ell$  of  $T$  such that  $\beta_\ell \subseteq \tau$ . Then  $\tau = [\beta_\ell]$ , and  $\beta_\ell$  is efficient in the tangle  $\tau$ .*
- (v) *For every consistent orientation  $\tau$  of  $S$  that is not an  $\mathcal{F}$ -tangle, the set  $\beta_\ell$  in (iii) has a subset  $\sigma$  in  $\mathcal{F}$ . Every such  $\sigma$  is efficient in  $\tau$ .*

*Proof.* (i) The first two statements are immediate from [Definition 3.6](#):  $\mathcal{F}$ -tangle structure trees are consistent, and leaves that are not forbidden are tangle leaves. The third assertion is part of (iv) and will be proved there.

(ii) The first assertion is again part of [Definition 3.6](#). The second is [Corollary 3.4](#). The minimality assertion is [Lemma 3.3](#). The maximality assertion holds, because  $(T, r, \beta)$  is ordered.

(iii) By [Lemma 3.2](#),  $T$  has a unique leaf  $\ell$  such that  $\beta_\ell \subseteq \tau$ . If  $\tau$  is consistent, then  $[\beta_\ell] \subseteq \tau$  by [Lemma 2.1](#). The claim now follows by [Lemma 4.2 \(ii\)](#).

(iv) We have to show that the leaf  $\ell$  from (iii) satisfies the inclusion  $[\beta_\ell] \subseteq \tau$  provided there with equality. As  $\tau$  has no subset in  $\mathcal{F}$ , the leaf  $\ell$  is not forbidden. It is therefore a tangle leaf, which means that  $[\beta_\ell]$  is a tangle of  $S$ . Since distinct tangles of  $S$  cannot contain one another, this tangle can only be  $\tau$ .

(v) Consider the leaf  $\ell$  provided for  $\tau$  by (iii). If  $\tau$  is not an  $\mathcal{F}$ -tangle, then neither is  $[\beta_\ell] \subseteq \tau$ : if it was, then  $[\beta_\ell]$  and  $\tau$  would both be orientations of  $S$ , implying  $[\beta_\ell] = \tau$  with a contradiction. Hence  $\beta_\ell$  has a subset in  $\mathcal{F}$ , by (i). The efficiency claim follows from (iii), since all subsets of  $\beta_\ell$  are efficient in  $\tau$  if  $\beta_\ell$  is.  $\square$

Let us briefly address the question of how efficient our structure trees are in displaying the  $\mathcal{F}$ -tangles of  $S$ , as well as certificates from  $\mathcal{F}$  for consistent orientations of  $S$  that are not  $\mathcal{F}$ -tangles.

There is a unique way to encode a single tangle  $\tau$  efficiently: we have to know its minimal elements, but only these, since the orientations of all the other separations in  $S$  are then determined by consistency of  $\tau$ .<sup>14</sup> The structure tree from [Theorem 6.1](#) does not achieve this for all the tangles of  $S$  simultaneously. Indeed, one only needs two separations in  $S$  to see that no structure tree can in general display only the minimal elements of all its  $\mathcal{F}$ -tangles.<sup>15</sup>

<sup>14</sup>Apply [Lemma 2.2](#) to the set  $\sigma$  of its minimal elements to reobtain the entire tangle as  $\tau = [\sigma]$ .

<sup>15</sup>For example, let  $S$  consist of just two nested separations,  $r$  and  $s$ . Let  $\mathcal{F} = \emptyset$ , so that all three consistent orientations of  $S$  are  $\mathcal{F}$ -tangles. The orientations  $\vec{r}$  of  $r$  and  $\vec{s}$  of  $s$  that point

However, since our structure tree  $T$  in [Theorem 6.1](#) is irreducible, it does the next best thing. By [Theorem 6.2](#)(iv), the minimal elements of every  $\mathcal{F}$ -tangle  $\tau$  of  $S$  are displayed as labels along the path  $rT\ell$ , where  $\ell$  is the tangle leaf with  $\tau = \lfloor \beta_\ell \rfloor$ . Conversely, every label  $\vec{s}_v$  of an edge of  $T$  is minimal in *some*  $\mathcal{F}$ -tangle of  $S$  (with a tangle leaf  $\ell > v$ ), or indispensable as an element of the subsets  $\sigma \in \mathcal{F}$  of  $\beta_\ell$  that certify that  $\lfloor \beta_\ell \rfloor$  is not an  $\mathcal{F}$ -tangle, for some forbidden leaf  $\ell > v$ .

[Theorem 6.1](#) is best possible in another sense too. Its characterizing condition for the existence of  $\mathcal{F}$ -tangle structure trees is, essentially, that  $\mathcal{F}$  must be rich. This cannot be weakened: if our order function is injective, which in real-world applications is the rule rather than the exception, then by [Lemma 4.7](#) the existence of any ordered  $\mathcal{F}$ -tangle structure tree (efficient and irreducible or not) implies that  $\mathcal{F}$  is rich. The richness condition on  $\mathcal{F}$ , thus, captures precisely what we need for an  $\mathcal{F}$ -tangle structure tree to exist.

If  $\mathcal{F}$  is rich not only for  $\vec{S}$  but also for some (or all)  $\vec{S}_k$ , [Theorem 6.1](#) applied in  $\vec{S}_k$  says that there exist efficient and irreducible ordered  $\mathcal{F}$ -tangle structures trees of all these  $\vec{S}_k$ . But we do not have to obtain these by independent applications of [Theorem 6.1](#), separately for each  $k$ : we can find them all in the original  $\mathcal{F}$ -tangle structure tree  $T$  for  $\vec{S}$  provided by [Theorem 4.6](#), before we applied [Theorem 5.3](#) to ‘reduce’ it. Let us now see how.

Given an ordered separation tree  $(T, r, \beta)$  on  $\vec{S}$  and  $k \in \mathbb{R}$ , let

$$(T, r, \beta)|_k = (T|_k, r, \beta|_k)$$

be defined as follows. Since  $T$  is ordered, its edges  $e$  with  $|\beta(e)| < k$  form a subtree of  $T$  rooted at  $r$ , which we take as  $T|_k$ . As its edge labelling  $\beta|_k$  we take the restriction of  $\beta$  to these edges.

**Theorem 6.3.** *If  $(T, r, \beta)$  is an efficient and thoroughly ordered  $\mathcal{F}$ -tangle structure tree of  $\vec{S}$  and  $\mathcal{F}$  is rich for  $\vec{S}_k$ , then  $(T, r, \beta)|_k$  is an efficient and thoroughly ordered  $\mathcal{F}$ -tangle structure tree of  $\vec{S}_k$ .*

*Proof.* It is clear from the definition of  $T|_k$  that it inherits from  $T$  all the properties it needs for being an efficient, thoroughly ordered, and consistent separation tree on  $\vec{S}_k$ ; in particular, if  $\beta_\ell$  is efficient in its closure  $\lfloor \beta_\ell \rfloor$  taken in  $\vec{S}$ , then its subsets  $\beta_v = \lfloor \beta|_k \rfloor_v$  with  $v < \ell$  a leaf of  $T|_k$  are efficient in their (smaller) closure  $\lfloor \beta_v \rfloor_k$  taken in  $\vec{S}_k$ . Moreover, non-leaf nodes  $v$  of  $T|_k$  are non-leaf nodes also of  $T$ , so  $\lfloor \beta_v \rfloor_k \subseteq \lfloor \beta_v \rfloor$  has no subset in  $\mathcal{F}$ .

For a proof that  $T|_k$  is an  $\mathcal{F}$ -tangle structure tree it remains to show that every leaf  $v$  of  $T|_k$  is either a tangle leaf or a forbidden leaf of  $T|_k$ . If it is neither, then by [Lemma 4.5](#) applied in  $\vec{S}_k$  the set  $\lfloor \beta_v \rfloor_k$  does not orient all of  $S_k$ .

Such a node  $v$  cannot be a leaf of  $T$ . Indeed, since  $\beta_v$  has no subset in  $\mathcal{F}$ , it would be a tangle leaf of  $T$ , so  $\lfloor \beta_v \rfloor$  would be an  $\mathcal{F}$ -tangle of  $S$ , and  $\lfloor \beta_v \rfloor \cap \vec{S}_k = \lfloor \beta_v \rfloor_k$  would be an  $\mathcal{F}$ -tangle of  $S_k$  (contrary to our assumption). So there exists an  $s_v \in S$  whose orientations label the edges  $e$  and  $e'$  from  $v$  to its children in  $T$ .

---

to each other then form a tangle in which they are both minimal, so they must both label an edge of  $T$ . Now if we flip either one of them, we obtain a tangle in which the other is still an element, but no longer minimal. If  $|r| < |s|$ , say, then  $\{\vec{r}, \vec{s}\}$  would be such a tangle which our structure tree displays by using both  $\vec{r}$  and  $\vec{s}$  as labels of a path from the root to a tangle leaf, although  $\vec{s}$  alone already determines this tangle. The tangle  $\{\vec{r}, \vec{s}\}$  would be displayed by the same tree more efficiently, with  $\vec{r}$  labelling an edge from the root to a leaf  $\ell$ , with  $\beta_\ell = \{\vec{r}\}$  and  $\lfloor \beta_\ell \rfloor = \{\vec{r}, \vec{s}\}$ .

Since  $T$  is thoroughly ordered and there exists a separation in  $S_k$  not oriented by  $\lfloor \beta_v \rfloor_k = \lfloor \beta_v \rfloor \cap \vec{S}_k$ , we know that  $|s_v| < k$ . But this means that  $e$  and  $e'$  are edges of  $T|_k$ , contradicting the fact that  $v$  is a leaf of  $T|_k$ .  $\square$

Note that the structure trees  $T|_k$  found inside  $T$  by [Theorem 6.3](#) are nested: for  $i < j$  we have  $T|_i \subseteq T|_j$ , indeed  $T|_i = (T|_j)|_i$ . This hierarchy of structure trees commutes with the following well-known hierarchy of tangles. For all integers  $i < j$ , every  $\mathcal{F}$ -tangle  $\tau_j$  of  $S_j$  induces an  $\mathcal{F}$ -tangle  $\tau_i$  of  $S_i$ , as  $\tau_i = \tau_j \cap \vec{S}_i$ . Now every  $\mathcal{F}$ -tangle in  $\vec{S}$ , say  $\tau_k$  of order  $k$ , corresponds to a tangle leaf  $\ell(\tau_k)$  of  $T|_k$  in that  $\tau_k = \lfloor \beta_{\ell(\tau_k)} \rfloor_k$ . All these  $\ell(\tau_k)$  are nodes of  $T$ .

Conversely, if our order function on  $S$  is injective, then *every* non-leaf node  $v$  of  $T$  is such a tangle leaf  $\ell(\tau_k)$  of  $T|_k$  for some  $\mathcal{F}$ -tangle  $\tau_k$  in  $\vec{S}$ , since  $\beta_v$  has no subset in  $\mathcal{F}$  by [Theorem 6.2 \(ii\)](#). And the positions of these nodes in  $T$  commute with the relationship of their tangles: if  $\tau_j$  induces  $\tau_i$  for  $i \leq j$ , then  $\ell(\tau_i) \leq_r \ell(\tau_j)$  in the tree-order of  $T$ .

In the proof of [Theorem 6.3](#) we made use of the assumption that  $T$  is thoroughly ordered. This holds for the  $\mathcal{F}$ -tangle structure trees provided by [Theorem 4.6](#), and these are already efficient by [Lemma 4.9](#). But while their efficiency is maintained when we apply [Theorem 5.3](#) to ‘reduce’ them, as in the proof of [Theorem 6.1](#), the property of being thoroughly ordered is lost. For the trees  $T|_k$  in [Theorem 6.3](#) to be  $\mathcal{F}$ -tangle structure trees, however, it is essential that  $T$  is thoroughly ordered.<sup>16</sup>

But the  $\mathcal{F}$ -tangle structure trees of  $\vec{S}_k$  that [Theorem 6.3](#) finds in  $T$  can be reduced afterwards: we simply apply [Theorem 5.3](#) to those  $T|_k$ . This yields irreducible, efficient, ordered  $\mathcal{F}$ -tangle structure trees for all the  $\vec{S}_k$ , the same that [Theorem 6.1](#) would give us if we applied it directly in  $\vec{S}_k$ .

Finally, let us apply [Theorem 6.1](#) to the special case that  $S$  has no  $\mathcal{F}$ -tangles. In this case it is desirable to be able to certify this efficiently.

In the special case that  $\mathcal{F}$  consists of stars (of oriented separations; see [Section 2](#)), such certificates are known: they are the *S-trees over  $\mathcal{F}$*  from [15], which generalize the branch-decompositions introduced by Robertson and Seymour [20] for graph tangles to  $\mathcal{F}$ -tangles in abstract separation systems.

Our structure trees from [Theorem 6.1](#) provide maximally efficient certificates for the non-existence of  $\mathcal{F}$ -tangles for general  $\mathcal{F}$ , not necessarily consisting of stars, when all their leaves are forbidden. Let us call such  $\mathcal{F}$ -tangle structure trees  *$\mathcal{F}$ -trees*.

**Theorem 6.4.** *Let  $\vec{S}$  be a separation system, and let  $\mathcal{F}$  be any set that is standard and rich for  $\vec{S}$ . Then exactly one of the following assertions holds:*

- (i) *there exists an  $\mathcal{F}$ -tangle of  $S$ ;*
- (ii) *there exists an  $\mathcal{F}$ -tree of  $\vec{S}$ .*

*In the case of (ii), the  $\mathcal{F}$ -tree can be chosen to be irreducible, efficient, and ordered with respect to any order function on  $S$ .*

*Proof.* By [Theorem 3.7 \(ii\)](#), any  $\mathcal{F}$ -tree of  $S$  precludes the existence of an  $\mathcal{F}$ -tangle of  $\vec{S}$ , so we cannot have both (i) and (ii).

<sup>16</sup>In the reduction process we contracted or deleted edges whose label was ‘unnecessary’ for  $T$  to display the  $\mathcal{F}$ -tangles of  $S$  and any certificates in  $\mathcal{F}$  of their non-existence. But such edges may have been necessary for  $T|_k$  to display the  $\mathcal{F}$ -tangles of  $S_k$ . As an extreme example, take  $\mathcal{F} = \emptyset$  and assume that all the oriented separations that are  $\leq$ -minimal in tangles of  $S$  have order  $k$ . While  $S$  may have elements of lower order that can be oriented to form lower-order tangles, those tangles will not be visible in  $T$ , since  $\beta(E(T)) \cap \vec{S}_k = \emptyset$ .

For a proof that (i) or (ii) holds, let any order function on  $S$  be given. By [Theorem 6.1](#),  $\vec{S}$  has an efficient and irreducible ordered  $\mathcal{F}$ -tangle structure tree  $T$ . If (i) fails, then  $T$  has no  $\mathcal{F}$ -tangle leaf. Then all its leaves are forbidden, so  $T$  is an  $\mathcal{F}$ -tree.  $\square$

## 7. APPLICATIONS: BLOCKS, PROFILES, AND CLUSTER TANGLES

Graph tangles can be expressed as  $\mathcal{F}$ -tangles in the universe of all separations of a graph  $G$ : take as  $\mathcal{F}$  the set  $\mathcal{T}$  of triples  $\{(A_i, B_i) \mid i = 1, 2, 3\}$  of oriented separations such that  $G[A_1] \cup G[A_2] \cup G[A_3] = G$ . These are not stars. But it is not hard to show [11] that the  $\mathcal{T}$ -tangles of a graph are precisely its  $\mathcal{T}^*$ -tangles, where  $\mathcal{T}^*$  consists of the sets in  $\mathcal{T}$  that are stars.

Expressed as  $\mathcal{F}$ -tangles in this way, graph tangles form a special case of certain  $\mathcal{T}$ -tangles in any abstract separation systems  $\vec{S}$  that lives in some universe  $\vec{U}$  of separations, where, unlike in the case of graphs, the triples in  $\mathcal{T}$  are described purely in terms of  $\vec{U}$  [7, Section 5]. If  $\vec{S}$  is submodular, has no degenerate elements, and  $\vec{U}$  is distributive, the  $\mathcal{T}$ -tangles of  $\vec{S}$  are again precisely its  $\mathcal{T}^*$ -tangles, where  $\mathcal{T}^*$  is the set of stars in  $\mathcal{T}$  [11, Lemma 17].

Due to this history,  $\mathcal{F}$ -tangles have so far been studied mostly when  $\mathcal{F}$  consisted of stars of separations. Such  $\mathcal{F}$ -tangles already allow for vast generalizations of the original tangles of graphs. However, there are contexts in which  $\mathcal{F}$ -tangles occur naturally and  $\mathcal{F}$  does not consist of stars. Three of these are:

- blocks in graphs, more precisely:  $k$ -blocks for all  $k \in \mathbb{N}$ ;
- profiles: the most comprehensive generalization of graph tangles that still admits a tree-of-tangles structure theorem;
- cluster tangles in large datasets.

We shall apply our results to all these in this section, giving a minimum of background for context in each case.

Common to all tangle contexts are two fundamental theorems about tangles in graphs, which both go back to the original paper of Robertson and Seymour [20] in which tangles were first introduced. These have become known [9] as

- the tree-of-tangles theorem; and
- the tangle-tree duality theorem.

The tree-of-tangles theorem, in its simplest form, says that all the maximal tangles in a graph can be distinguished<sup>17</sup> by a nested set  $T$  of separations, one that can be represented as the set of separations associated with a tree-decomposition of the graph. The separations in  $T$  can be chosen to distinguish the tangles *efficiently*, which means that they have minimum order for any separation that distinguishes a given pair of tangles, and *canonically*, which means that  $T$  is invariant under all the automorphisms of the graph [9]. These enhanced tree-of-tangle theorems have been further generalized to abstract separation systems [6, 11, 12].

The tangle-tree duality theorem, in its simplest form, says that if a graph has no tangle of some given order  $k$  then it has a nested set  $T$  of separations of order  $< k$  that witnesses the non-existence of those tangles: any orientation  $\tau$  of the graph's separations of order  $< k$  will in particular orient  $T$ , and thereby the edges of the

---

<sup>17</sup>A separation *distinguishes* two tangles if they orient it differently.

decomposition tree associated with  $T$ , and following those oriented edges will take us to a node whose incident edges correspond to a set in  $\mathcal{T}^*$ ; thus,  $\tau$  is not a tangle. The tangle-tree duality theorem, too, has been generalized to abstract separation systems [11, 15]. There are also unifications of both theorems into one [1, 2, 18].

The certificates for the non-existence of an  $\mathcal{F}$ -tangle offered by the tangle-tree duality theorems require that  $\mathcal{F}$  consist of stars. Indeed, since  $T$  is nested, the edges at any node of the decomposition tree that represents  $T$  will map to a nested set of separations too, and orienting them towards that node will yield a star of separations. Hence if we look for certificates for the non-existence of  $\mathcal{F}$ -tangles when  $\mathcal{F}$  does not consist of stars, we need a fresh start. Our [Theorem 6.4](#) offers such an alternative.

For our first application, let us now look at blocks. For any  $k \in \mathbb{N}$ , a  $k$ -block in a graph is any maximal set of at least  $k$  vertices no two of which can be separated by  $< k$  vertices. Every  $k$ -block  $X$  in a graph  $G$  defines an  $\mathcal{B}_k$ -tangle of  $S_k$  for

$$\mathcal{B}_k := \left\{ \sigma \subseteq \vec{S} : \left| \bigcap \{B \mid (A, B) \in \sigma\} \right| < k \right\},$$

where  $S$  is the set of all separations of  $G$ : just orient every separation in  $S_k$  towards  $X$ . Conversely, given any  $\mathcal{B}_k$ -tangle  $\tau$  of  $S_k$  in  $G$ , the set  $\bigcap \{B \mid (A, B) \in \tau\}$  is a  $k$ -block in  $G$ , which in turn defines  $\tau$  as indicated.

One of the earliest applications of what later became  $\mathcal{F}$ -tangles was that finite graphs have canonical tree-decompositions that distinguish all their  $k$ -blocks efficiently [5]. This extended a classical decomposition theorem of Tutte for 2-blocks. The proof in [5] was not based on the tangle techniques from [20], which could not deliver canonical tree-decompositions. As a consequence of this different approach, any tangle-tree duality theorem for blocks seemed beyond reach for a long time.

It was shown much later that blocks can in fact be captured by  $\mathcal{F}$ -tangles with  $\mathcal{F}$  consisting of stars [10, 11, 16]. But these  $\mathcal{F}$  have to be constructed recursively for each given graph, and are therefore highly artificial: tailor-made for the purpose, but not defining  $\mathcal{F}$ -tangles that can be understood from their definition. As a consequence, the certificates they yield in theory for the non-existence of blocks are of little use in practice. Our [Theorem 6.4](#), however, offers very tangible certificates:

**Theorem 7.1.** *Let  $S$  be the set of separations of a graph  $G$ , with the usual order function [9], and let  $k \in \mathbb{N}$ . Then exactly one of the following assertions holds:*

- (i)  *$G$  has a  $k$ -block;*
- (ii) *there exists an  $\mathcal{B}_k$ -tree of  $\vec{S}_k$ .*

*In the case of (ii), the  $\mathcal{B}_k$ -tree can be chosen to be irreducible, efficient, and ordered.*

*Proof.* We have already seen that  $k$ -blocks induce  $\mathcal{B}_k$ -tangles of  $S_k$  and vice versa. We can thus replace (i) with the assertion that  $S_k$  has a  $\mathcal{B}_k$ -tangle, and proceed to apply [Theorem 6.4](#).

Let us check the theorem's premise. The set  $\mathcal{B}_k$  is standard for  $\vec{S}_k$ , since the all co-trivial oriented separations in  $\vec{S}_k$  are of the form  $(V, A)$ , where  $V = V(G)$  and  $|A| < k$ .<sup>18</sup> Clearly, then,  $\{(V, A)\} \in \mathcal{B}_k$ .

For a proof that  $\mathcal{B}_k$  is rich for  $\vec{S}_k$ , it suffices by [Lemma 4.4](#) to show that  $\mathcal{B}_k$  is closed under minimization in  $\vec{S}$ . Given any  $\sigma \in \mathcal{F}$ , let  $\sigma'$  be obtained from  $\sigma$  by

<sup>18</sup>In addition, the set  $A$  has to lie in the separator of another separation of order  $< k$ .

replacing every  $(A, B) \in \sigma$  by some  $(A', B') \leq (A, B)$  from  $\vec{S}$ .<sup>19</sup> As

$$\bigcap \{ B' : (A', B') \in \sigma' \} \subseteq \bigcap \{ B : (A, B) \in \sigma \},$$

our assumption of  $\sigma \in \mathcal{B}_k$  clearly implies that  $\sigma' \in \mathcal{B}_k$ , as required.  $\square$

Next, profiles. Profiles of graphs were introduced by Hundertmark [19], who noticed that the above-mentioned tree-structure theorem for blocks [5] required in its proof much less information about blocks than is provided in their definition. All that is needed is that the orientation of  $S_k$  which a  $k$ -block induces<sup>20</sup> avoids the set  $\mathcal{F}$  of triples of the form  $\{ \vec{r}, \vec{s}, \vec{r} \vee \vec{s} \}$ . In the context of graphs, this supremum is taken in the universe of all the separations of the given graph.

More generally, given any separation system  $\vec{S}$  in some universe  $\vec{U}$  of separations, we let

$$\mathcal{P} = \{ \{ \vec{r}, \vec{s}, \vec{r} \vee \vec{s} \} : \vec{r}, \vec{s} \in \vec{U} \},$$

and refer to the  $\mathcal{P}$ -tangles of  $S$  as its *profiles*. A profile is *regular* if none of its elements is small.

Hundertmark's discovery was seminal: since graph tangles are profiles too, the canonical construction from [5], rewritten for profiles [12], offered canonical trees of tangles – something that Robertson and Seymour [20] had been unable to achieve directly by their tangle methods.

So what about tangle-tree duality for profiles? It was the hope to find such a theorem that motivated the original generalization of graph tangles to  $\mathcal{F}$ -tangles in abstract separation systems [15], since profiles had this form by definition. But the hope proved elusive. We did indeed find a duality theorem for  $\mathcal{F}$ -tangles in [15], but we needed crucially that  $\mathcal{F}$  consisted of stars. This was good enough for tangles, since  $\mathcal{T}$ -tangles coincide with  $\mathcal{T}^*$ -tangles, as well as for some further duality theorems in graphs and matroids [14], but there was no tangle-tree duality theorem for profiles among these.

As in the case of blocks it is possible to construct, recursively, sets  $\mathcal{F}$  of stars in submodular separation systems embedded in a universe so that its profiles are precisely its  $\mathcal{F}$ -tangles. One can then apply [15] to obtain a tangle-tree duality theorem for profiles, encoded as such  $\mathcal{F}$ -tangles [10].

Alternatively, we can now obtain certificates for the non-existence of profiles directly from [Theorem 6.4](#).

As preparation, we need a lemma. Given a separation system  $\vec{S}$  in some universe  $\vec{U}$ , we call a profile of  $S$  *strong* if it has no subset in

$$\mathcal{P}_s = \{ \{ \vec{r}, \vec{s}, \vec{t} \} : \vec{r}, \vec{s}, \vec{t} \in \vec{U} \text{ and } \vec{t} \leq \vec{r} \vee \vec{s} \}.$$

Note that strong profiles are regular: if  $\vec{r} \leq \vec{r}$  then no strong profile contains  $\vec{r}$ , since  $\vec{r} \leq \vec{r} \vee \vec{r}$  puts  $\{ \vec{r} \}$  in  $\mathcal{P}_s$ . If  $\vec{S}$  is submodular and  $\vec{U}$  is distributive, the converse holds too:

**Lemma 7.2.** *Every regular profile of a submodular separation system in a distributive universe of separations is strong.*

*Proof.* Let  $(\vec{S}, \leq)$  be any submodular separation system in a distributive lattice  $\vec{U}$ . Suppose  $S$  has a regular profile  $\tau$  that is not strong. Then there exist  $\vec{r}, \vec{s}, \vec{t}$  in  $\tau$  such that  $\vec{t} \leq \vec{r} \vee \vec{s}$ . Choose this triple with  $\vec{t}$  as large as possible under  $\leq$ .

<sup>19</sup>Recall that, for graph separations,  $(A', B') \leq (A, B)$  if and only if  $A' \supseteq A$  and  $B' \subseteq B$ .

<sup>20</sup>orient every separation towards the side that contains the given  $k$ -block

Our aim will be to show that both  $\vec{t} \vee \vec{r}$  and  $\vec{t} \vee \vec{s}$  lie in  $\vec{S}$ . For if they do, they must lie in  $\tau$  by the consistency of  $\tau$ , since  $\vec{t} \vee \vec{r} \geq \vec{r} \in \tau$  and  $\vec{t} \vee \vec{s} \geq \vec{s} \in \tau$  and  $\tau$  is regular.<sup>21</sup> Then, as

$$(\vec{t} \vee \vec{s})^* \vee (\vec{t} \vee \vec{r})^* = (\vec{t} \wedge \vec{s}) \vee (\vec{t} \wedge \vec{r}) = \vec{t} \wedge (\vec{s} \vee \vec{r}) = \vec{t}$$

by distributivity, we shall have  $\tau \supseteq \{\vec{t} \vee \vec{r}, \vec{t} \vee \vec{s}, \vec{t}\} \in \mathcal{P}$ , which contradicts our assumption that  $\tau$  is a profile.

So let us show that  $\vec{t} \vee \vec{r} \in \vec{S}$ ; the proof of  $\vec{t} \vee \vec{s} \in \vec{S}$  is analogous. By submodularity we have  $\vec{t} \vee \vec{r} \in \vec{S}$  as desired unless  $\vec{t} \wedge \vec{r} \in \vec{S}$ , so let us assume this. Then also  $\vec{S} \ni (\vec{t} \wedge \vec{r})^* = \vec{t} \vee \vec{r} \geq \vec{t} \in \tau$ , so  $\vec{t} \vee \vec{r} \in \tau$  by the consistency and regularity<sup>21</sup> of  $\tau$ .

If  $\vec{t} \vee \vec{r} = \vec{t}$ , then  $\vec{t} \geq \vec{r}$  and thus  $\vec{t} \leq \vec{r}$ , giving  $\vec{t} \vee \vec{r} = \vec{r} \in \vec{S}$  as desired. But otherwise  $\vec{t} \vee \vec{r} > \vec{t} \in \tau$ , which contradicts the maximality of  $\vec{t}$  in our choice of  $\vec{r}, \vec{s}, \vec{t} \in \tau$  with  $\vec{t} \leq \vec{r} \vee \vec{s}$ , since also  $\vec{t} \vee \vec{r} \leq \vec{r} \vee \vec{s}$  by definition of the supremum  $\vec{t} \vee \vec{r}$ .  $\square$

**Lemma 7.2** implies that the regular profiles of a submodular separation system in a distributive universe are precisely its  $\mathcal{P}_s$ -tangles. Indeed, being profiles, they are  $\mathcal{P}$ -tangles and therefore consistent, and they avoid  $\mathcal{P}_s$  since, by the lemma, they are strong, so they are  $\mathcal{P}_s$ -tangles. Conversely, every  $\mathcal{P}_s$ -tangle avoids  $\mathcal{P} \subseteq \mathcal{P}_s$  and is therefore a profile, and it is strong since  $\mathcal{P}_s$ -tangles cannot contain small separations, as noted earlier.

**Theorem 7.3.** *Let  $\vec{S}$  be any submodular separation system in a distributive universe of separations. Then exactly one of the following assertions holds:*

- (i) *there exists a regular profile of  $S$ ;*
- (ii) *there exists an  $\mathcal{P}_s$ -tree of  $\vec{S}$ .*

*In the case of (ii), the  $\mathcal{P}_s$ -tree can be chosen to be irreducible, efficient, and ordered with respect to any order function on  $S$ .*

*Proof.* Our aim is to apply **Theorem 6.4** to  $\vec{S}$  with  $\mathcal{F} = \mathcal{P}_s$ . As remarked after **Lemma 7.2**, the regular profiles of  $S$  are precisely its  $\mathcal{P}_s$ -tangles. It remains to check that  $\mathcal{P}_s$  is standard and rich for  $\vec{S}$ .

For a proof that  $\mathcal{P}_s$  is standard, we have to show that  $\{\vec{s}\} \in \mathcal{P}_s$  whenever  $\vec{s}$  is trivial in  $\vec{S}$ . Let  $r \in S$  witness that  $\vec{s}$  is trivial. Then  $\vec{r}, \vec{r} < \vec{s}$ , so  $\vec{s} < \vec{r} < \vec{s}$ . Hence  $\vec{s} \leq \vec{s} \vee \vec{s}$ , which puts  $\{\vec{s}\}$  in  $\mathcal{P}_s$ .

For a proof that  $\mathcal{P}_s$  is rich for  $\vec{S}$ , it suffices by **Lemma 4.4** to show that  $\mathcal{P}_s$  is closed under minimization in  $\vec{S}$ . Given any set  $\sigma = \{\vec{r}, \vec{s}, \vec{t}\} \in \mathcal{P}_s$ , consider any set  $\sigma' = \{\vec{r}', \vec{s}', \vec{t}'\}$  with  $\vec{r}' \leq \vec{r}$  and  $\vec{s}' \leq \vec{s}$  and  $\vec{t}' \leq \vec{t}$ . Then  $\sigma' \in \mathcal{P}_s$  as desired, since  $\vec{t}' \leq \vec{t} \leq \vec{r} \vee \vec{s} \leq \vec{r}' \vee \vec{s}'$ ; the middle inequality comes from  $\sigma \in \mathcal{P}_s$ .  $\square$

Finally, let us apply our results to tangles defined by clusters in large datasets. The basic setup is that  $\vec{S}$  is the set  $2^V$  of all subsets of our data set  $V$ , the involution  $*$  given by complementation in  $V$ . Then  $S$  is the set of bipartitions  $s = \{A, B\}$  of  $V$ , whose orientations are  $\vec{s} = A$  and  $\vec{s} = B$  (say). We then choose an order function on  $S$  that assigns high values to separations that divide many pairs of elements of  $V$

---

<sup>21</sup>Compare **Lemma 2.1** and the remark preceding it.

which we consider as ‘close’,<sup>22</sup> the idea being that separations of low order cannot divide many close pairs and therefore cannot cut through a cluster, only chip off a few points. This, then, implies that large clusters induce  $\mathcal{F}$ -tangles for  $\mathcal{F}$  defined as

$$\mathcal{C}_n := \{ \{ \vec{r}, \vec{s}, \vec{t} \} : |\vec{r} \cap \vec{s} \cap \vec{t}| < n \},$$

where the *agreement* value  $n$  is chosen to fit the context.

Thus,  $\mathcal{C}_n$ -tangles in datasets are a bit like  $\mathcal{B}_k$ -tangles in graphs, except that  $n$  is now a value separate from the order  $k$  of the tangles considered. An important difference is that, unlike  $k$ -blocks in graphs, the clusters captured by those tangles are not normally subsets  $X$  of  $V$  that we can, or even wish to, specify explicitly: they can be ‘fuzzy’, and do not have to lie entirely on the side of the separations chosen by the tangle they induce, only mostly. Quantitatively, while the intersection of any three elements of an  $\mathcal{C}_n$ -tangle contains at least  $n$  datapoints, the intersection of all of them will likely be empty – unlike in the case of  $k$ -blocks in graphs, which are equal to the intersection of *all* the sides of the graph’s separations chosen by the tangle.

In other contexts than generic clustering one sometimes looks for  $\mathcal{C}_n$ -tangles not of the set  $S$  of all the bipartitions of  $V$ , but of some specific, hand-designed set  $S$  of bipartitions of  $V$ . For example,  $S$  might be a questionnaire whose elements partition a population  $V$  of individuals that have answered it into those that answered yes and those that answered no. In such a context,  $\mathcal{C}_n$ -tangles can be interpreted as ‘typical’ ways to answer the questions in  $S$ , as *mindsets* found in the population  $V$  regarding the questions in  $S$ . See [8] for more.

In all these contexts we can now use [Theorem 6.1](#) to efficiently display the clusters or mindsets in  $V$  as determined by  $S$ , as long as  $\mathcal{C}_n$  is standard and rich for  $\vec{S}$ . If we are interested specifically in certifying that our dataset has no clusters of some desired density at all, we can similarly apply [Theorem 6.4](#).

The sets  $\mathcal{C}_n$  are standard for all  $\vec{S} \neq \emptyset$  and  $n > 0$ : then  $\{\emptyset\} \in \mathcal{C}_n$ , and only  $\emptyset$  can be co-trivial in  $\vec{S}$ . Also,  $\mathcal{C}_n$  is clearly closed under minimization in  $\vec{S}$ : replacing the elements  $\vec{s}$  of a triple  $\sigma$  with subsets  $\vec{s}' \leq \vec{s}$  leads to  $\bigcap \sigma' \subseteq \bigcap \sigma$ , so  $\sigma \in \mathcal{C}_n$  implies  $\sigma' \in \mathcal{C}_n$ . By [Lemma 4.4](#), therefore,  $\mathcal{C}_n$  is rich for  $\vec{S}$ .

[Theorems 6.1](#) and [6.4](#) thus have the following instances in our context:

**Corollary 7.4.** *If  $\vec{S}$  is any non-empty set of subsets of a set  $V$  and  $n > 0$ , then  $\vec{S}$  has an efficient and irreducible ordered  $\mathcal{C}_n$ -tangle structure tree.*

**Corollary 7.5.** *Let  $\vec{S}$  be any non-empty set of subsets of a set  $V$ , and let  $n > 0$ . Then exactly one of the following assertions holds:*

- (i) *there exists a  $\mathcal{C}_n$ -tangle of  $S$ ;*
- (ii) *there exists a  $\mathcal{C}_n$ -tree of  $\vec{S}$ .*

*In the case of (ii), the  $\mathcal{C}_n$ -tree can be chosen to be irreducible, efficient, and ordered with respect to any order function on  $S$ .*

## REFERENCES

[1] Sandra Albrechtsen, *Optimal trees of tangles: refining the essential parts*, 2023.

---

<sup>22</sup>Choosing this order function to fit the application context allows for considerable variety in using tangles for clustering. See [8, Chapter 9] for an overview of some standard order functions, or [21] for a concrete example.

- [2] ———, *Refining trees of tangles in abstract separation systems: inessential parts*, Combinatorial Theory **4**(1) (2024; arXiv:2302.01808).
- [3] J. Carmesin, R. Diestel, M. Hamann, and F. Hundertmark, *Canonical tree-decompositions of finite graphs I. Existence and algorithms*, J. Comb. Theory Ser. B **116** (2016), 1–24; arXiv:1305.4668.
- [4] ———, *Canonical tree-decompositions of finite graphs II. Essential parts*, J. Comb. Theory Ser. B **118** (2016), 268–283; arXiv:1305.4909.
- [5] J. Carmesin, R. Diestel, F. Hundertmark, and M. Stein, *Connectivity and tree structure in finite graphs*, Combinatorica **34** (2014), no. 1, 1–35; arXiv:1105.1611.
- [6] J. Carmesin and J. Kurkofka, *Entanglements*, J. Comb. Theory, Ser. B **164** (2024), 17–28; arXiv:2205.11488.
- [7] R. Diestel, *Abstract separation systems*, Order **35** (2018), 157–170; arXiv:1406.3797.
- [8] ———, *Tangles: A structural approach to artificial intelligence in the empirical sciences*, Cambridge University Press, 2024. e-book: [tangles-book.com/book/](http://tangles-book.com/book/).
- [9] ———, *Graph theory* (6th edition), Springer-Verlag, 2025.  
Electronic edition available at <http://diestel-graph-theory.com/>.
- [10] R. Diestel, Ph. Eberenz, and J. Erde, *Duality theorems for blocks and tangles in graphs*, SIAM J. Discrete Math. **31** (2017), no. 3, 1514–1528; arXiv:1605.09139.
- [11] R. Diestel, J. Erde, and D. Weißauer, *Structural submodularity and tangles in abstract separation systems*, J. Comb. Theory Ser. A **167C** (2019), 155–180; arXiv:1805.01439.
- [12] R. Diestel, F. Hundertmark, and S. Lemanczyk, *Profiles of separations: in graphs, matroids, and beyond*, Combinatorica **39** (2019), no. 1, 37–75; arXiv:1110.6207.
- [13] R. Diestel and H.v.Bergen, *Tangle structure tree applications*.
- [14] R. Diestel and S. Oum, *Tangle-tree duality in graphs, matroids and beyond*, Combinatorica **39** (2019), 879–910. arXiv:1701.02651.
- [15] ———, *Tangle-tree duality in abstract separation systems*, Advances in Mathematics **377** (2021), 107470; arXiv:1701.02509.
- [16] Ph. Eberenz, *Characteristics of profiles*, Master’s Thesis, Hamburg University, 2015.
- [17] C. Elbracht, J. Kneip, and M. Teegen, *Obtaining trees of tangles from tangle-tree duality*, J. Combinatorics **13** (2022), 251–287; arXiv:2011.09758.
- [18] Joshua Erde, *Refining a tree-decomposition which distinguishes tangles*, SIAM J. Discrete Math. **31** (2017), 1529–1551; arXiv:1512.02499.
- [19] F. Hundertmark, *Profiles. An algebraic approach to combinatorial connectivity*, 2011.
- [20] N. Robertson and P.D. Seymour, *Graph minors. X. Obstructions to tree-decomposition*, J. Comb. Theory Ser. B **52** (1991), 153–190.
- [21] H. von Bergen and R. Diestel, *Traits and tangles: an analysis of the Big Five paradigm by tangle-based clustering*, J. Math. Psychology **125 (102920)** (2025; arXiv:2411.18670).