Tree-representation of set families and applications to combinatorial decompositions ⋆

Binh-Minh Bui-Xuan [a,∗] Michel Habib [b] Michaël Rao [c]

[a] *Department of Informatics, University of Bergen, Norway.* `buixuan@ii.uib.no`

[b] *CNRS – LIAFA – Université Paris 7, France.* `habib@liafa.jussieu.fr`

[c] *CNRS – LABRI, France.* `rao@labri.fr`

Abstract

The number of families over ground set V is $2^{2^{|V|}}$ and by this fact it is not possible to represent such a family using a number of bits polynomial in $|V|$. However, under some simple conditions, this becomes possible, like in the cases of a symmetric crossing family and a weakly partitive family, both representable using $\Theta(|V|)$ space.

We give a general framework for representing any set family by a tree. It extends in a natural way the one used for symmetric crossing families in [Cunningham and Edmonds, *Canadian Journal of Mathematics*, 1980]. We show that it also captures the one used for weakly partitive families in [Chein, Habib, and Maurer, *Discrete Mathematics*, 1981]. We introduce two new classes of families: weakly partitive crossing families are those closed under the union, the intersection, and the difference of their crossing members, and union-difference families those closed under the union and the difference of their overlapping members. Each of the two cases encompasses symmetric crossing families and weakly partitive families. Applying our framework, we obtain a linear $\Theta(|V|)$ and a quadratic $O(|V|^2)$ space representation based on a tree for them. We introduce the notion of a sesquimodule – one module and half – in a digraph and in a generalization of digraphs called 2-structure. From our results on set families, we show for any digraph, resp. 2-structure, a unique decomposition tree using its sesquimodules. These decompositions generalize strictly the clan decomposition of a digraph and that of a 2-structure. We give polynomial time algorithms computing the decomposition tree for both cases of sesquimodular decomposition.

Key words: graph decomposition, cross-free family, modular decomposition, symmetric submodular function minimizer, sesquimodular decomposition

⋆ This work encompasses the work reported in [4–6]. It was done while the first author was a Ph. D. student of the Université Montpellier II. For financial support, the authors are grateful to the Norwegian Research Council, project PARALGO, and the French National Research Agency, project GRAAL.

∗ Corresponding author. Tel: (+47) 55 58 42 83. Fax: (+47) 55 58 41 99.

1 Introduction

Many combinatorial decompositions rely on their connections to well-known families. An important key for this practice is the seminal bijection between the so-called cross-free families and unrooted trees [17]. It leads to decomposition framework not only in graph theory [13,15,22,32,36] and in combinatorial optimization [14,23,41], but also in phylogeny [42]. Therein, the basic idea is to study the distance of a given family $\mathcal{F} \subseteq 2^V$ from a tree structure, namely to find for $\mathcal{F}$ a representation by a (possibly labelled) tree in such a way that we can enumerate all members of $\mathcal{F}$ in $O(|\mathcal{F}|)$ time. From this perspective, we address more particularly the space complexity of a family as that of a tree representing it. Accordingly, at the first level we find simple hierarchies, a.k.a. laminar families, and cross-free families, both cases in $\Theta(|V|)$ complexity. Extending these we find symmetric crossing and weakly partitive families: symmetric crossing families are closed under the complementation of any member and under the intersection of their crossing members; weakly partitive families are closed under the union, the intersection, and the difference of their overlapping members. While the latter families are fundamental for the modular – or clan – decomposition of several discrete structures (see [18] for an extensive survey), the former families arise in a number of studies related to symmetric submodular functions (see [41]). We will come back to these two cases with a more specific discussion below. For now it matters that symmetric crossing families are $\Theta(|V|)$ tree representable [12,15,16], and the same holds for weakly partitive families [10]. All the classes of families we mentioned so far are included in the class of crossing families, also known as one of the largest classes of polynomial space complexity: a crossing family is closed under the union and the intersection of its crossing members. These families arise in the study of directed network flows (see [41]), and for them only an $O(|V|^2)$ representing tree is known [25]. However, this asymptotic bound is tight [1]: they are $\Theta(|V|^2)$ tree representable. Contrasting these cases, no polynomial tree is known for representing a family forming a binary vector space (namely closed under the symmetric difference), although its complexity is $\Theta(|V|^2)$: upper bound obtained by looking at a vectorial basis; lower bound by counting such bases [29].

The study of tree representations of set families can be beneficial to graph decompositions from other perspectives too. The minimization of submodular functions is fundamental in combinatorial optimization (see [41] for an extensive survey). When the function is also symmetric, *e.g.*, for undirected network flows or robber-and-cops graph searching, the minimization problem has a faster and very simple solution [39]. The symmetric case draws a particular attention from research in graph decompositions, as the properties of a symmetric submodular function are exploited in many ways for proving deep results (computability, well-quasi ordering, etc) related to the branch decomposition of a connectivity function, the branch-width decomposition of a graph, and the rank-width decomposition of a graph (see [28,38,40] and the bibliography therein). Beside this, modular graph decomposition is a classical and fundamental topic in graph theory. It has a rich history that goes back to the late 40s [46]. Nowadays, one of its most important properties is that the corresponding decomposition tree can be computed in linear time [44]. To this aim, considerable research effort has been put in developing a long list of efficient and clever algorithmic techniques (see [30] for a recent survey). Although classical and very well studied, techniques for modular decomposition and for submodular function minimization have been developed separately. The situation is similar between modular decomposition and branch decomposition (resp. its restriction to the branch-width or rank-width decompositions of graphs). From the perspective of bringing new insights from each case to the other, it could be desirable to ask for a common ground where these topics intersect.

We answer to such a question positively by looking at the structure of fundamental set families therein. It is well-known that the family of (non-trivial) minimizers of a symmetric submodular function is symmetric crossing, while the family of modules of a graph is (weakly) partitive. In Section 2 we give a framework for representing any set family by a tree. It captures and actually is a natural extension of the work presented in [15] on symmetric crossing families. On the other hand, its connection to weakly partitive families is not obvious. In Section 3 we show how, when restricted to them, our framework is strictly equivalent to the one used in [10], giving the first result generalizing both the modular decomposition of a graph and the structural behaviour of the minimizers of a symmetric submodular function. Deepening the question on representing set families by a tree, we address in Section 4 two natural generalizations of symmetric crossing families and weakly partitive families: a family is weakly partitive crossing if it is closed under the union, the intersection, and the difference of its crossing members; a family is a union-difference family if it is closed under the union and the difference of its overlapping members. Straight from definition, the class of weakly partitive crossing families encompasses weakly partitive families and symmetric crossing families. Also from definition, union-difference families encompass weakly partitive families. Curiously, they also encompass symmetric crossing families. Using the framework developed in Section 2 we show for each of the two new cases a canonical tree representation, resulting in that the complexity of weakly partitive crossing families is in $\Theta(|V|)$, and that of union-difference families in $O(|V|^2)$. The class of union-difference families is incomparable with the class of crossing families and in this sense our result extends the class of families having a polynomial space complexity. Furthermore, in Section 5 we show that each of the two new classes of families implies a new combinatorial decomposition. For this we introduce the notion of a sesquimodule – one module and half – in a 2-structure. The formalism includes the case when the 2-structure is a digraph. Sesquimodules are proper generalizations of the notion of a clan [19] (an excellent introduction to this topic is [18]). Based on sesquimodules and applying our framework for representing set families by an one-to-one correspondence with trees, we show uniqueness decomposition theorems for digraphs and 2-structures. In Sections 6 and 7 we describe polynomial algorithms to deal with these decompositions. We close the paper with some open questions and perspectives related to the use of combinatorics of set families in algorithmic graph decompositions.

Family name	Definition
crossing	closed under ∪ and ∩ of crossing members
intersecting	closed under ∪ and ∩ of overlapping members
weakly partitive crossing	crossing & closed under \ of crossing members
partitive crossing	weakly partitive crossing & closed under Δ of crossing members
symmetric crossing	crossing & closed under complementation
bipartitive	symmetric crossing & closed under Δ of crossing members
weakly partitive	intersecting & closed under \ of overlapping members
partitive	weakly partitive & closed under Δ of overlapping members
cross-free	no two members cross
overlap-free (or laminar)	no two members overlap
binary vector space	closed under Δ
union-difference	closed under ∪ and \ of overlapping members

Fig. 1. Glossary of families appearing in this paper. We denote by ∪, ∩, \ and Δ the set operations union, intersection, difference, and symmetric difference, respectively.

Family name	c-∪	c-∩	c-\	c-Δ	o-∪	o-∩	o-\	o-Δ	*symm.*	Δ
crossing	yes	yes								
intersecting	yes	yes			yes	yes				
weakly partitive crossing	yes	yes	yes							
partitive crossing	yes	yes	yes	yes						
symmetric crossing	yes	yes	yes						yes	
bipartitive	yes	yes	yes	yes					yes	
weakly partitive	yes	yes	yes		yes	yes	yes			
partitive	yes	yes	yes	yes	yes	yes	yes	yes		
cross-free	yes	yes	yes	yes						
overlap-free (or laminar)	yes	yes	yes	yes	yes	yes	yes	yes		
binary vector space										yes
union-difference	yes		yes		yes		yes			

Fig. 2. Closure properties. A set operation preceded by letter "c", resp. "o", means the closure applies on crossing, resp. overlapping, members of the family. By "*symm.*" we mean the closure under the complementation of any member of the family.

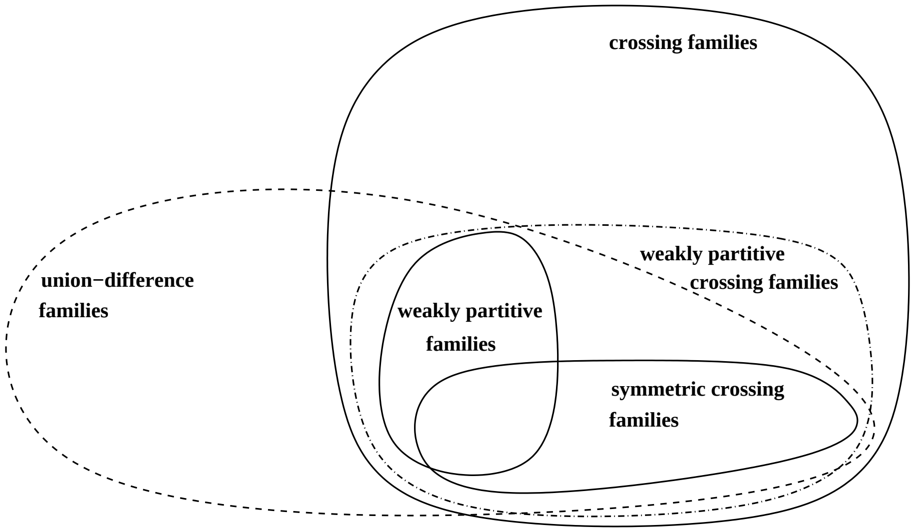


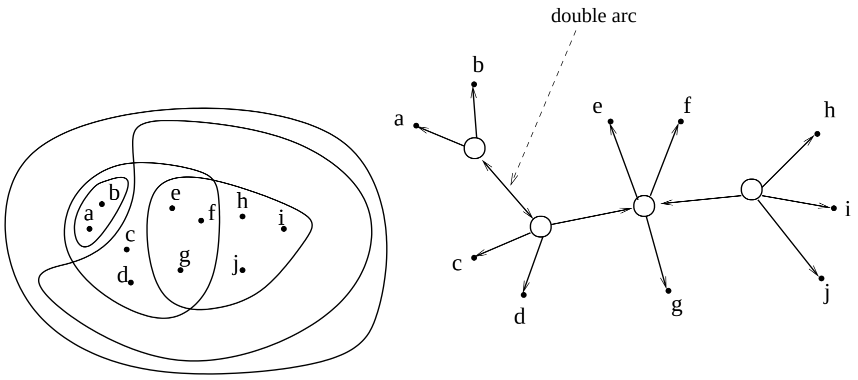
Fig. 3. Some classes of families. When comparable, the inclusion is strict. The space complexity of weakly partitive, symmetric crossing and weakly partitive crossing families are all in $\Theta(|V|)$ while the complexity of crossing families is in $\Theta(|V|^2)$. The complexity of union-difference families is in $O(|V|^2)$, but this bound is not known to be tight.

2 A general framework for representing set families

We make the convention that every family $\mathcal{F} \subseteq 2^V$ in this paper satisfies $|V| \geq 3$, $\emptyset \notin \mathcal{F}$, $V \in \mathcal{F}$, and $\{v\} \in \mathcal{F}$ (for all $v \in V$), unless we explicitly state otherwise. Further on in the paper we will need to address graphs and digraphs: graphs will refer to loopless simple undirected graphs whereas digraphs are pairs of the type (V, E) where $E \subseteq V \times V \setminus \{(v, v) : v \in V\}$. In particular, digraphs can have double-arcs (directed cycles over two vertices).

Let V be an n-element set. Two subsets $A \subseteq V$ and $B \subseteq V$ *overlap*, noted $A \between B$, if their intersection is not empty and their differences are not empty: $A \cap B \neq \emptyset$, $A \setminus B \neq \emptyset$, and $B \setminus A \neq \emptyset$. Two subsets $A \subseteq V$ and $B \subseteq V$ *cross* if we have both $A \between B$ and $\overline{A} \between \overline{B}$, where the complement of a subset $A \subseteq V$ is denoted by $\overline{A} = V \setminus A$. We call a family an *overlap-free family* (resp. *cross-free family*) if no two members of the family overlap (resp. cross). Overlap-free families are also known as *laminar families* [41].

Ordering the members of an overlap-free family by inclusion will result in a rooted tree, the *hierarchy* of the family. A similar result holds for every cross-free family $\mathcal{C} \subseteq 2^V$: for $x \in V$ consider the family containing the members of $\mathcal{C}$ excluding x and the complements of the members of $\mathcal{C}$ containing x. Remove the empty set, and obtain an overlap-free family over $V \setminus \{x\}$. Add $\{x\}$ to the children of the root of its hierarchy, unroot the resulting tree, and obtain T. For every edge uv in T, denote by $\{S_u, S_v\}$ the 2-partition of V induced by the leaves of the two trees we get by removing uv from T. Clearly, $S_u \in \mathcal{C}$ or $S_v \in \mathcal{C}$, or both facts hold. This can be represented by using an orientation over edge uv in T (double-arcs are allowed).



i. **ii.**

Fig. 4. i. A cross-free family ii. Its representation by the Edmonds-Giles tree.

On the other hand, for every member $A \in \mathcal{C}$ with $A \neq V$, there exists an edge uv of T such that $A = S_u$ or $A = S_v$. This is therefore a representation of $\mathcal{C}$ by an unrooted tree (with edge orientations), the *Edmonds-Giles tree* of $\mathcal{C}$. An example is drawn in Figure 4. We have revised the proof [41, Theorem 13.21 proof] of a classical result:

Theorem 1 (Edmonds-Giles [17]) *A family is cross-free if and only if it has an unrooted tree representation, and overlap-free if and only if it has a rooted tree representation.*

Let $\mathcal{F} \subseteq 2^V$ be a family. We call $A \in \mathcal{F}$ an *overlap-free member* (resp. *cross-free member*) if there is no $B \in \mathcal{F}$ such that A and B overlap (resp. cross). The *overlap-free sub-family,* or *laminar sub-family*, $\mathcal{L}$ of $\mathcal{F}$ is the family containing all overlap-free members of $\mathcal{F}$. Likewise, we define the *cross-free sub-family* $\mathcal{C}$ of $\mathcal{F}$ as the family containing all cross-free members of $\mathcal{F}$. Clearly, $\mathcal{L}$ is an overlap-free family, $\mathcal{C}$ is a cross-free family, and $\mathcal{L} \subseteq \mathcal{C}$. In $\mathcal{F}$ we define two types of structural hierarchy, that we call *decomposition trees* of the family:

Definition 1 (Decomposition trees) Any family $\mathcal{F} \subseteq 2^V$ can be associated with the Edmonds-Giles tree of its cross-free sub-family, that we call the *cross-free decomposition tree* $\mathcal{T}_\mathcal{C}$ of $\mathcal{F}$. Likewise, $\mathcal{F}$ can be associated with the hierarchy of its overlap-free sub-family, that we call the *overlap-free decomposition tree,* or *laminar decomposition tree*, $\mathcal{T}_\mathcal{L}$ of $\mathcal{F}$.

Remark 1 *Suppose that we are given a family that is already overlap-free (and hence cross-free as well). Then, either its cross-free decomposition tree has one and only one source, or there is in that tree a unique double-arc and subdividing the double-arc by setting a source in between will result in a tree with one and only one source. In both cases, inverting the orientations defines a rooted tree which turns out to be isomorphic to the overlap-free decomposition tree of the initial family.*

As previously mentioned, an overlap-free member of a family is also a cross-free member of the family, however, note that the converse in not always true. We say that $\mathcal{T}_\mathcal{C}$ is a proper "refinement" of $\mathcal{T}_\mathcal{L}$. Basically, a decomposition tree is a (possibly trivial) sub-family of the initial family. We will now divide the rest into another type of sub-families that we call *quotients*.

Definition 2 (Quotient family) Let $\mathcal{F} \subseteq 2^V$ be a set family, and $\mathcal{T}_\mathcal{C}$, $\mathcal{T}_\mathcal{L}$ the cross-free and overlap-free decomposition trees of $\mathcal{F}$, respectively. Let u be an internal node of $\mathcal{T}_\mathcal{C}$ having neighbours $a_1, a_2, \ldots, a_k$. Note that $k \geq 3$. Let $V_1, V_2, \ldots, V_k$ be the subsets of V induced by the leaves of the trees containing $a_1, a_2, \ldots, a_k$ when removing node u from $\mathcal{T}_\mathcal{C}$, respectively. We consider $W = \{V_1, V_2, \ldots, V_k\}$ as a ground set and define the *quotient family* $\mathcal{Q}(u) \subseteq 2^W$ of node u in $\mathcal{T}_\mathcal{C}$ as

$$\mathcal{Q}(u) = \left\{ Q:\ \exists I \subseteq [\![1,k]\!],\ Q = \{V_i:\ i \in I\} \ \wedge \ \left(|I| = 1 \ \vee \ \bigcup_{i \in I} V_i \in \mathcal{F} \right) \right\}.$$

Let u be an internal node of $\mathcal{T}_\mathcal{L}$ having children $a_1, a_2, \ldots, a_k$. Let $V_1, V_2, \ldots, V_k$ be the subsets of V induced by the leaves of the subtrees of $\mathcal{T}_\mathcal{L}$ rooted at $a_1, a_2, \ldots, a_k$, respectively. We consider $W = \{V_1, V_2, \ldots, V_k\}$ as a ground set (here we deviate from the convention of this paper, when $k = 2$). The *quotient family* $\mathcal{Q}(u) \subseteq 2^W$ of node u in $\mathcal{T}_\mathcal{L}$ is defined as

$$\mathcal{Q}(u) = \left\{ Q:\ \exists I \subseteq [\![1,k]\!],\ Q = \{V_i:\ i \in I\} \ \wedge \ \bigcup_{i \in I} V_i \in \mathcal{F} \right\}.$$

For representation purposes, the main idea is that $\mathcal{F}$ can be obtained by, roughly, taking the union of all quotient families in $\mathcal{T}_\mathcal{C}$ (resp. $\mathcal{T}_\mathcal{L}$). We will now state and prove this formally for $\mathcal{T}_\mathcal{C}$: the argument for $\mathcal{T}_\mathcal{L}$ is analogous. Let f be the function mapping every set family to its cross-free decomposition tree. Then, given an arbitrary set family $\mathcal{F}$, the function $g_\mathcal{F}$ over the domain $\{u:\ u \text{ is an internal node of } f(\mathcal{F})\}$ which maps every node u of $f(\mathcal{F})$ to its quotient family $g_\mathcal{F}(u)$, is well-defined. Therefore, the function h mapping a set family $\mathcal{F}$ to the pair $h(\mathcal{F}) = (f(\mathcal{F}), g_\mathcal{F})$ is well-defined. We want to prove that h is injective. This way, h will be a bijection mapping the set of all set families to the image of h, in other words, $h(\mathcal{F})$ will be a representation for every set family $\mathcal{F}$. The following property appears frequently in literature related to the representation of set families, and is relatively well-known.

Proposition 1 *Let A be a member of a family $\mathcal{F} \subseteq 2^V$ that is not cross-free. There exists one and only one node x in the cross-free decomposition tree $\mathcal{T}_\mathcal{C}$ of $\mathcal{F}$ such that A corresponds to some member Q of the quotient family $\mathcal{Q}(x)$ of node x, namely $A = \bigcup_{V_i \in Q} V_i$.*

Proof: There are many proofs for this property. A graphical one, extending the ideas of [37, proof of Lemma 1], is as follows. With respect to A, we will define a special orientation over every edge uv of $\mathcal{T}_\mathcal{C}$. Let S_u and S_v be the 2-partition of V induced by the leaves of the two trees we get by removing uv from$\mathcal{T}_\mathcal{C}$. Note that if A crosses one among S_u and S_v, then A crosses both. Since at least one among S_u and S_v is a cross-free member of $\mathcal{F}$, A does not cross any of them. Besides, $A \neq S_u$ and $A \neq S_v$ since otherwise A would be a cross-free member of $\mathcal{F}$. Hence, there are only two cases, which are self-exclusive:

- A is strictly included in either S_u or S_v (the or is exclusive): w.l.o.g. suppose it was S_u, then the special orientation is defined to be from v to u.
- either S_u or S_v (the or is exclusive) is strictly included in A: w.l.o.g. suppose it was S_u, then the special orientation is defined to be from u to v.

It is a straightforward exercise to check that the special orientation has one and only one sink: for any edge uv of $\mathcal{T}_\mathcal{C}$ where the special orientation is from u to v, for every edge st belonging to the connected component which contains u when we remove uv from $\mathcal{T}_\mathcal{C}$, it suffices to prove – *e.g.* by case analysis – that the special orientation is from the farther node s (w.r.t. u) to the nearer node t (w.r.t. u).

Let x be the (one and only one) sink defined by the special orientation of $\mathcal{T}_\mathcal{C}$. Let $\{V_1, V_2, \ldots, V_k\}$ be the ground set of the quotient family $\mathcal{Q}(x)$ of node x. Here, A cannot be included in some V_i since otherwise the special orientation from the component containing V_i to the sink x would be reversed. By elimination, every V_i is either included in A or included in the complement of A. We can then conclude by applying Definition 2 on $\mathcal{Q}(x)$. □

Corollary 1 *The above defined function h is injective.*

Proof: Let $\mathcal{F}$ and $\mathcal{G}$ be such that $h(\mathcal{F}) = h(\mathcal{G})$. By symmetry it suffices to prove that $\mathcal{F} \subseteq \mathcal{G}$. Let $A \in \mathcal{F}$. If A is a cross-free member of $\mathcal{F}$ then we can conclude by using $f(\mathcal{F}) = f(\mathcal{G})$. Otherwise, from Proposition 1, there exists an internal node u of $f(\mathcal{F}) = f(\mathcal{G})$ such that there exists a member $Q \in g_\mathcal{F}(u)$ such that $A = \bigcup_{V_i \in Q} V_i$ and $|Q| \neq 1$. Now, from $h(\mathcal{F}) = h(\mathcal{G})$ we also have $g_\mathcal{F} = g_\mathcal{G}$. Hence: u is an internal node of $f(\mathcal{G})$, $Q \in g_\mathcal{G}(u)$, $A = \bigcup_{V_i \in Q} V_i$, and $|Q| \neq 1$. This, by Definition 2, implies $A \in \mathcal{G}$. □

A stronger claim than that of Corollary 1 is as follows. Its proof is straightforward.

Remark 2 *If one can represent exactly the quotient family $\mathcal{Q}(u)$ of every node u of the cross-free decomposition tree $\mathcal{T}_\mathcal{C}$ of $\mathcal{F}$, then one can represent exactly $\mathcal{F}$.*

Similar propositions can be made for the overlap-free decomposition tree $\mathcal{T}_\mathcal{L}$ with minor modifications (replacing "cross" with "overlap"). Accordingly, we say that $\mathcal{T}_\mathcal{C}$ and $\mathcal{T}_\mathcal{L}$, when equipped with their quotient families, are the *cross-free representation* and the *overlap-free representation* of $\mathcal{F}$, respectively. Further details can be found in [3]. Cross-free and overlap-free representations can benefit from data-structures called PC and PQ trees with minor modifications, respectively [2,43].

Definition 3 (Cross-free and overlap-free representations) For any set family we say that the cross-free decomposition tree, together with the quotient families of its internal nodes, is the *cross-free representation* of the family. The *overlap-free representation* of the family is its overlap-free decomposition tree, together with the quotient families of the internal nodes.

3 Revisiting the representation of a weakly partitive family and unifying it with the representation of a symmetric crossing family

This section focuses on two classical results in the topic of representing set families, about the so-called weakly partitive families and symmetric crossing families. The formal definitions will follow next, but let us first specify how the formalism given in the previous section will help in giving a link between modular decomposition and submodular function minimization, and what exactly we are trying to prove.

A set function $f : 2^V \to \mathbb{R}$ is submodular if $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$ for every $A \subseteq V$ and every $B \subseteq V$. It is symmetric if $f(A) = f(\overline{A})$ for every $A \subseteq V$. The family of non-trivial minimizers – those different from $\emptyset$ and V – of a symmetric submodular function is *symmetric crossing*, namely it is closed under the complementation of any member (except for V) and under the intersection of its crossing members [1] . The formal definition of a graph module will be given in Section 5. For now it matters that the family of modules of a graph is *weakly partitive*, namely it is closed under the union, the intersection, and the difference of its overlapping members. Any symmetric crossing family $\mathcal{F} \subseteq 2^V$ has a $\Theta(|V|)$ space representation [12,15,16], and the same holds for weakly partitive families [10]. The representation given in [15] is built on the Edmonds-Giles tree of the cross-free decomposition tree of the input symmetric crossing family. In this sense, we say that the cross-free representation of an arbitrary set family (as in Definition 3) is a natural extension of the work presented in [15]. Likewise, the overlap-free representation is a natural extension of the work presented in [10]. A particular case of [10] addresses the modular decomposition tree of a graph, which is nothing more than the overlap-free decomposition tree (as in Definition 1) of the family of modules of the graph, plus a classification of the internal nodes into three categories (via quotient families and something else). The connection between the two representations in Definition 3 is not obvious, even when restricted to weakly partitive families. In this section, we will give such a connection by showing that the cross-free representation of a weakly partitive family turns out to be exactly its overlap-free representation. The cross-free representation will then be a framework that unifies both [15] and [10], hence generalizes both the modular decomposition of a graph and the structural behaviour of the minimizers of a symmetric submodular function.

A *partitive* family is a weakly partitive family that is also closed under the symmetric difference of its overlapping members.

Theorem 2 (Chein, Habib, and Maurer [10]) *There is a $\Theta(|V|)$ space representation of any weakly partitive family over V. The representation can be based on the overlap-free decomposition tree of the family: a quotient family $\mathcal{Q} \subseteq 2^W$ in the overlap-free decomposition tree $\mathcal{T}_{\mathcal{L}}$ of a weakly partitive family $\mathcal{F} \subseteq 2^V$ can be of only three types: $\mathcal{Q} = \{W\} \cup \{\{w\} : w \in W\}$; or $\mathcal{Q} = 2^W \setminus \{\emptyset\}$; or there is an*

[1] By convention, we still assume for a symmetric crossing family $\mathcal{F}$ that $\emptyset \notin \mathcal{F}$ and $V \in \mathcal{F}$, and assume that the closure under complementation does not apply on V.

ordering of the elements of W such that a subset of W belongs to $\mathcal{Q}$ if and only if the subset is also an interval of the ordering. Moreover, if the family is partitive then the last case can not occur.

We will denote these three types by *prime*, *complete*, and *linear*, respectively. Such a classification leads directly to a representation in $O(|V|)$ space of $\mathcal{T_L}$ and its quotient families. Therefore, it also gives a representation in $\Theta(|V|)$ of any weakly partitive family over V.

As previously said, a family is *symmetric crossing* if it is closed under the complementation of any member (except for V) and under the intersection of its crossing members. Note that such a family is also closed under the union and the difference of its crossing members. The overlap-free decomposition tree $\mathcal{T_L}$ of $\mathcal{F}$ is trivial, namely it is a star having $\mathcal{F}$ as unique quotient family. A *bipartitive* family is a symmetric crossing family that is also closed under the symmetric difference of its crossing members.

Theorem 3 (Cunningham and Edmonds [15]) *There is a $\Theta(|V|)$ space representation of any symmetric crossing family over V. The representation can be based on the cross-free decomposition tree of the family: a quotient family $\mathcal{Q} \subseteq 2^W$ in the cross-free decomposition tree $\mathcal{T_C}$ of a symmetric crossing family $\mathcal{F} \subseteq 2^V$ can be of only three types: $\mathcal{Q} = \{W\} \cup \{\{w\} : \ w \in W\} \cup \{W \setminus \{w\} : \ w \in W\}$; or $\mathcal{Q} = 2^W \setminus \{\emptyset\}$; or there is a circular ordering of the elements of W such that a subset of W belongs to $\mathcal{Q}$ if and only if the subset is also a circular interval of the ordering. Moreover, if the family is bipartitive then the last case can not occur.*

We will denote these three types by *prime*, *complete*, and *circular*, respectively. Surprisingly, the easiest way to retrieve this result is to make a detour using the overlap-free paradigm (of a sub-family of $\mathcal{F}$): roughly, if $|V| = 3$ then the family is complete, otherwise pick $x \in V$, consider the family $\mathcal{F}'$ of members of $\mathcal{F}$ excluding x, check that $\mathcal{F}'$ is weakly partitive, apply Theorem 2, add x to the root of the resulting decomposition tree, unroot the tree, and classify the quotient families accordingly.

As mentioned before, while the cross-free representation can be seen as a natural extension of Theorem 3, the overlap-free representation can be seen as a natural extension of Theorem 2. We now unify these two viewpoints for weakly partitive families. Let $\mathcal{F} \subseteq 2^V$ be a weakly partitive family, and $\mathcal{T_L}$ (resp. $\mathcal{T_C}$) its overlap-free (resp. cross-free) decomposition tree. We have remarked in the previous section that in general $\mathcal{T_L}$ can be obtained from $\mathcal{T_C}$ by contracting some edges, but the converse is not always true. However, we claim that the converse is true when $\mathcal{F}$ is weakly partitive. Actually, we show this for a larger class of families:

Lemma 1 *If a family $\mathcal{F} \subseteq 2^V$ is closed under the intersection and the difference of its overlapping members, then every cross-free member $A \in \mathcal{F}$ is either an overlap-free member or the complement of an overlap-free member.*

Proof: Assume that A is not overlap-free, and let $B \in \mathcal{F}$ be such that $A \between B$.

Since they do not cross we have $A \cup B = V$, and hence $B \setminus A = \overline{A}$. By the difference closure, $\overline{A} \in \mathcal{F}$. If $|\overline{A}| = 1$ then $\overline{A}$ is clearly overlap-free and we can conclude. Otherwise, suppose that $\overline{A}$ is not overlap-free, and let $C \in \mathcal{F}$ be such that $\overline{A} \between C$. But they do not cross (otherwise A is not cross-free), and therefore $\overline{A} \cup C = V$. Then, B and C overlap. Hence, $B \cap C \in \mathcal{F}$ by the intersection closure. We now deduce that A and $B \cap C$ cross. Contradiction. □

Therefore, $\mathcal{T}_{\mathcal{L}}$ and $\mathcal{T}_{\mathcal{C}}$ have the same underlying graph (upto isomorphism). But retrieving the root of $\mathcal{T}_{\mathcal{L}}$ is tricky. Luckily, and somewhat unexpectedly, a notion arising from [4], then only developed for classifying some quotient families, will help here. We say that $A \in \mathcal{F}$ is *quasi-trivial* if $|A| = |V| - 1$.

Definition 4 (Simply-linked family and guard) A family $\mathcal{F} \subseteq 2^V$ is *simply-linked* if no quasi-trivial members of $\mathcal{F}$ are overlap-free members of $\mathcal{F}$. If $\mathcal{F}$ is not simply-linked, then there is one and only one element $v \in V$, so-called the *guard* of $\mathcal{F}$, such that $\mathcal{G} = \mathcal{F} \setminus \{V, \{v\}\}$ is a family over $W = V \setminus \{v\}$.

Definition 5 Let $\mathcal{Q}(u) \subseteq 2^W$ be the quotient family of a node u of the cross-free decomposition tree $\mathcal{T}_{\mathcal{C}}$ of a family $\mathcal{F} \subseteq 2^V$. Suppose that $\mathcal{Q}(u)$ is not simply-linked. Then, the guard of $\mathcal{Q}(u)$ corresponds to a unique neighbour v of u in $\mathcal{T}_{\mathcal{C}}$. We say that v is the *guarding parent* of u in $\mathcal{T}_{\mathcal{C}}$.

Lemma 2 *Let $\mathcal{F} \subseteq 2^V$ be a weakly partitive family. There is at most one simply-linked quotient family in the cross-free decomposition tree $\mathcal{T}_{\mathcal{C}}$ of $\mathcal{F}$.*

Proof: (by contradiction). Suppose there are two internal nodes u, v in $\mathcal{T}_{\mathcal{C}}$ such that the quotient family of each node is simply-linked. For convenience, we make use of abuse of terminology, and also denote the underlying graph of $\mathcal{T}_{\mathcal{C}}$ by $\mathcal{T}_{\mathcal{C}}$. Let $(u = u_1, u_2, \ldots, u_p = v)$ be the path linking u to v in $\mathcal{T}_{\mathcal{C}}$. Let A be the leaf set of the connected component containing u that we get when removing the edge between u_{p-1} and v in $\mathcal{T}_{\mathcal{C}}$. Let B be the leaf set of the connected component containing v that we get when removing the edge between u and u_2 in $\mathcal{T}_{\mathcal{C}}$.

We first prove that both A and B are members of $\mathcal{F}$ in two steps.

- If u and v are neighbours in $\mathcal{T}_{\mathcal{C}}$, then A and B are complementary. From the definition of a cross-free decomposition tree, either A or its complement B is a member of $\mathcal{F}$. W.l.o.g. suppose that A is a member of $\mathcal{F}$. Let us examine the quotient family $\mathcal{Q}(u)$ of u. There, B corresponds to an element of the ground set of $\mathcal{Q}(u)$, and the complement of the singleton $\{B\}$ belongs to $\mathcal{Q}(u)$ since $A = V \setminus B$ belongs to $\mathcal{F}$. Since $\mathcal{Q}(u)$ is not simply-linked, $\mathcal{Q}(u)$ can only be complete or linear according to Lemma 3 below. In both cases, there exists an element C of the ground set of $\mathcal{Q}(u)$ holding both $C \subsetneq V \setminus B$ and $\{B, C\}$ belongs to $\mathcal{Q}(u)$ ($C \neq V \setminus B$ because the ground set of $\mathcal{Q}(u)$ has at least 3 elements). One consequence is that $B \cup C$ is a member of $\mathcal{F}$. This member overlaps A, and the difference closure of $\mathcal{F}$ implies $B = (B \cup C) \setminus A$ is a member of $\mathcal{F}$.
- If u and v are not neighbours in $\mathcal{T}_{\mathcal{C}}$, then A and B overlap. From the definition of a cross-free decomposition tree, either B or its complement $V \setminus B$ is a member

of $\mathcal{F}$. Let us prove that B is always a member of $\mathcal{F}$. For this, it suffices to prove that if $V \setminus B$ is a member of $\mathcal{F}$, then so is B. We proceed exactly as before: B corresponds to an element of the ground set of $\mathcal{Q}(u)$, and the complement of the singleton $\{B\}$ belongs to $\mathcal{Q}(u)$ since $V \setminus B$ belongs to $\mathcal{F}$. Since $\mathcal{Q}(u)$ is not simply-linked, $\mathcal{Q}(u)$ can only be complete or linear according to Lemma 3 below. In both cases, there exists an element C of the ground set of $\mathcal{Q}(u)$ holding both $C \subsetneq V \setminus B$ and $\{B, C\}$ belongs to $\mathcal{Q}(u)$. One consequence is that $B \cup C$ is a member of $\mathcal{F}$. Since $B \cup C$ and $V \setminus B$ overlap, their difference $(B \cup C) \setminus (V \setminus B) = B$ is a member of $\mathcal{F}$ by the difference closure of $\mathcal{F}$. By a similar argument, we can also prove that A is a member of $\mathcal{F}$.

Hence, both A and B are members of $\mathcal{F}$. But then, we can also obtain that their respective complements are members of $\mathcal{F}$: if u and v are neighbours, then it is trivial; otherwise, just notice that $A \between B$ and $V \setminus B = A \setminus B$, then use the difference closure. Here again, since $V \setminus B$ is a member of $\mathcal{F}$, all arguments in the last paragraph applies: there exists $C \subsetneq V \setminus B$ such that $B \cup C$ is a member of $\mathcal{F}$. By a similar argument on A and the quotient family of v, we can also obtain that there exists $D \subsetneq V \setminus A$ such that $A \cup D$ is a member of $\mathcal{F}$. But then $C \cap D$ is a member of $\mathcal{F}$ which crosses both A and $V \setminus A$ (it also crosses both B and its complement). This would mean that the edge between u_{p-1} and v cannot be an edge of the cross-free decomposition tree $\mathcal{T}_{\mathcal{C}}$. Contradiction. □

Theorem 4 *Let $\mathcal{F} \subseteq 2^V$ be a weakly partitive family and $\mathcal{T}_{\mathcal{C}}$ its cross-free decomposition tree. Let $\overrightarrow{\mathcal{T}}$ be the oriented tree having the same underlying graph as $\mathcal{T}_{\mathcal{C}}$, but the orientation of $\overrightarrow{\mathcal{T}}$ is defined by the guarding parents as in Definition 5. Then, either $\overrightarrow{\mathcal{T}}$ has one and only one sink: it is a rooted tree, noted $\widehat{\mathcal{T}}$; or $\overrightarrow{\mathcal{T}}$ has one and only one double-arc uv, and subdividing uv by adding a new guarding parent of both u and v will result in a rooted tree, noted $\widehat{\mathcal{T}}$. Moreover, $\widehat{\mathcal{T}}$ in both cases is isomorphic to the overlap-free decomposition tree $\mathcal{T}_{\mathcal{L}}$ of $\mathcal{F}$.*

Proof: The fact that, in both cases, $\widehat{\mathcal{T}}$ and $\mathcal{T}_{\mathcal{L}}$ are isomorphic is straightforward from Lemma 1 and Definition 5. We only need to prove the other claims of the theorem. From Lemma 2, we have two pairwise exclusive configurations. Before continuing we highlight that one of the important facts in the following is that the ground set of a quotient family in $\mathcal{T}_{\mathcal{C}}$ always has at least 3 elements.

In one configuration of Lemma 2, $\mathcal{T}_{\mathcal{C}}$ has exactly one internal node u s.t. the quotient family $\mathcal{Q}(u)$ of u is simply-linked. Let us prove that u is the unique sink of $\overrightarrow{\mathcal{T}}$. Here it suffices to prove that, for every pair of neighbours s, t in $\mathcal{T}_{\mathcal{C}}$, the nearer node t (w.r.t. u) is the guarding parent of the farther node s (w.r.t. u). We will proceed by contradiction. Suppose that there exists a neighbour w of s such that $w \neq t$ and w is the guarding parent of s. Notice that t might coincide with u but $s \neq u$ and $w \neq u$. Let $(s = u_1, t = u_2, u_3, \ldots, u_p = u)$ be the path linking s to u in $\mathcal{T}_{\mathcal{C}}$. Let Z be the leaf set of the connected component containing u_{p-1} we get when removing the edge between u_{p-1} and u in $\mathcal{T}_{\mathcal{C}}$. We claim that Z is a member of $\mathcal{F}$. Indeed, by definition of the cross-free decomposition tree either Z or $\overline{Z}$ is a member of $\mathcal{F}$.

Moreover, if $\overline{Z}$ is a member of $\mathcal{F}$ then we can proceed by a similar argument as in the proof of Lemma 2 as follows. Z corresponds to an element of the ground set of $\mathcal{Q}(u)$. Since $\overline{Z}$ is a member of $\mathcal{F}$, by Lemma 3 (below) $\mathcal{Q}(u)$ is not prime. Besides, all cases lead to the existence of another element Y of the ground set of $\mathcal{Q}(u)$ s.t. $Y \subsetneq \overline{Z}$ and $\{Y, Z\}$ belongs to $\mathcal{Q}(u)$. But then $Y \cup Z$ is a member of $\mathcal{F}$ which overlaps the member $\overline{Z}$ of $\mathcal{F}$. By difference closure, $Z = (Y \cup Z) \setminus \overline{Z}$ is a member of $\mathcal{F}$. We will now use the membership of Z in order to exhibit a contradiction. Let A be the leaf set of the connected component containing w we get when removing node s from $\mathcal{T}_\mathcal{C}$. By definition of a guarding parent, A corresponds to the guard of the quotient family $\mathcal{Q}(s)$ of node s. Let B be the leaf set of the connected component containing s we get when removing the edge between s and t in $\mathcal{T}_\mathcal{C}$. Note that $\overline{B}$ corresponds to an element of the ground set of $\mathcal{Q}(s)$. Clearly B and $\overline{A}$ overlap. We claim that B is a member of $\mathcal{F}$. Indeed, if $t = u$ then $B = Z$ and we have just proved that Z is a member of $\mathcal{F}$. Otherwise, by definition of the cross-free decomposition tree either B or $\overline{B}$ is a member of $\mathcal{F}$. Moreover, if $\overline{B}$ is a member of $\mathcal{F}$ then it overlaps the member Z of $\mathcal{F}$, and $B = Z \setminus \overline{B}$ will also be a member of $\mathcal{F}$. But then in $\mathcal{Q}(s)$, the membership of B in $\mathcal{F}$ would lead to the existence of a (quasi-trivial) member of $\mathcal{Q}(s)$ which overlaps the complement of the guard A. This contradicts with the definition of a guard.

In the other configuration of Lemma 2, every internal node u of $\mathcal{T}_\mathcal{C}$ is such that the quotient family of u is not simply-linked. The crucial point is the following. Let u and v be two nodes of $\mathcal{T}_\mathcal{C}$ such that v is the guarding parent of u. Let st be an edge of the connected component containing u when removing the edge uv from $\mathcal{T}_\mathcal{C}$, with t being the nearer node (w.r.t. u) and s being the farther node (w.r.t. u). Then, we claim that t is the guarding parent of s. Indeed, suppose by contradiction that there exists a neighbour w of s such that $w \neq t$ and w is the guarding parent of s. Notice that t might coincide with u but $s \neq u$ and $w \neq u$. Let A be the leaf set of the connected component containing w we get when removing the edge between s and w in $\mathcal{T}_\mathcal{C}$. By definition of a guarding parent, A corresponds to the guard of the quotient family $\mathcal{Q}(s)$ of node s. Let Z be the leaf set of the connected component containing s we get when removing the edge between s and t in $\mathcal{T}_\mathcal{C}$. Note that $\overline{Z}$ corresponds to an element of the ground set of $\mathcal{Q}(s)$. Clearly Z and A overlap. Then Z cannot be a member of $\mathcal{F}$ since this would contradict the fact A is the guard of $\mathcal{Q}(s)$. However, either Z or its complement is a member of $\mathcal{F}$. Besides, let B be the leaf set of the connected component containing u we get when removing the edge between u and v in $\mathcal{T}_\mathcal{C}$. Clearly, B is a member of $\mathcal{F}$ for v is the guarding parent of u. Now, if $\overline{Z}$ is a member of $\mathcal{F}$ then it would overlap B and $Z = B \setminus \overline{Z}$ would also be a member of $\mathcal{F}$. Contradiction. We conclude that either $\overrightarrow{T}$ has one sink then the sink is unique, or there exist in $\overrightarrow{T}$ one and only one double-arc between u and v and then it is straightforward to conclude. $\square$

We now focus on the quotient families in $\mathcal{T}_\mathcal{L}$ and $\mathcal{T}_\mathcal{C}$. The proof of Chein-Habib-Maurer Theorem 2 can be adapted in a straightforward manner in order to obtain the following property (more details can be found in [3, Theorems 2.2 and 2.3]).

Lemma 3 *Let $\mathcal{F} \subseteq 2^V$ be a weakly partitive family. Let $\mathcal{Q} \subseteq 2^W$ be a quotient family in the cross-free decomposition tree $\mathcal{T}_\mathcal{C}$ of $\mathcal{F}$. If $\mathcal{Q}$ is simply-linked, then, either $\mathcal{Q}$ is prime, namely $\mathcal{Q} = \{W\} \cup \{\{w\} : w \in W\}$; or $\mathcal{Q}$ is complete, namely $\mathcal{Q} = 2^W \setminus \{\emptyset\}$; or $\mathcal{Q}$ is linear, namely there is an ordering of the elements of W such that a subset of W belongs to $\mathcal{Q}$ if and only if the subset is also an interval of the ordering. If $\mathcal{Q}$ is not simply-linked, let $w \in W$ be the guard of $\mathcal{Q}$, let $\mathcal{P} = \mathcal{Q} \setminus \{W, \{w\}\}$ and $X = W \setminus \{w\}$. Then, either $\mathcal{P}$ is prime, namely $\mathcal{P} = \{X\} \cup \{\{x\} : x \in X\}$; or $\mathcal{P}$ is complete, namely $\mathcal{P} = 2^X \setminus \{\emptyset\}$; or $\mathcal{P}$ is linear, namely there is an ordering of the elements of X such that a subset of X belongs to $\mathcal{P}$ if and only if the subset is also an interval of the ordering.*

After this, a straightforward case analysis shows that the corresponding labelled $\mathcal{T}_\mathcal{L}$ and $\mathcal{T}_\mathcal{C}$ (with prime, complete, and linear labels) are isomorphic.

Remark 3 *The isomorphism in Theorem 4 is label-preserving when using quotient families as labels for the two kinds of decomposition trees, in the way described in Theorem 2 for the overlap-free representation and Lemma 3 for the cross-free representation.*

4 Weakly partitive crossing families and union-difference families

The family of minimizers over $2^V \setminus \{\emptyset, V\}$ of a submodular function over ground set V is a crossing family, meaning that it is closed under the union and the intersection of its crossing members. For crossing families only a representation tree in $O(|V|^2)$ space is known [25], and moreover this asymptotic bound is essentially tight [1]: they have $\Theta(|V|^2)$ complexity. As previously mentioned, the family of (non-trivial) minimizers of a symmetric submodular function is a symmetric crossing family, a.k.a. a crossing family that is also symmetric. We have seen that symmetric crossing families have $\Theta(|V|)$ complexity. In this section we focus on two generalizations of symmetric crossing families: weakly partitive crossing families and union-difference families. One of them is a particular case of crossing families while the other is incomparable with crossing families (see Figure 3). A natural question then is to decide the complexity of the two new classes. We in fact show that weakly partitive crossing families have $\Theta(|V|)$ space complexity, witnessed by a representation built on a tree: the cross-free representation as in Definition 3. On the other hand, we show that the cross-free representation of union-difference families has $O(|V|^2)$ space complexity, but leave open the question whether this is tight.

A family $\mathcal{F} \subseteq 2^V$ is *weakly partitive crossing* if it is closed under the union, the intersection, and the difference of its crossing members. A *partitive crossing* family is a weakly partitive crossing family that is also closed under the symmetric difference of its crossing members. It is clear from definition that weakly partitive crossing families contain at the same time symmetric crossing families and weakly partitive families. Also, since the class of symmetric crossing families and the class of weakly partitive families are incomparable, the previous inclusion is strict. Finally, it is straightforward to build weakly partitive crossing families that are neither symmetric crossing nor weakly partitive.

A family $\mathcal{F} \subseteq 2^V$ is a *union-difference* family if it is closed under the union and the difference of its overlapping members. It is clear from definition that union-difference families contain weakly partitive families. It is less obvious, but an easy exercise to check that they also contain symmetric crossing families. Here again, the inclusions are strict (by the same argument as before). Also, it is straightforward to build union-difference families that are neither symmetric crossing nor weakly partitive. Finally, note that if a union-difference family is also closed under the symmetric difference of its overlapping members, then it is also closed under the intersection of its overlapping members, or, in other words, the family becomes a partitive family.

From Remark 2, in order to obtain a representation theorem for weakly partitive crossing, resp. union-difference, families, it suffices to prove there are a constant number of simple types of quotient families in a cross-free decomposition tree of a weakly partitive crossing, resp. union-difference, family. For a family $\mathcal{F} \subseteq 2^V$ we say that $A \in \mathcal{F}$ is *trivial* if either $|A| = 1$ or $A = V$, and recall that $A \in \mathcal{F}$ is quasi-trivial if $|A| = |V| - 1$. For a member that is neither trivial nor quasi-trivial, we say that it is a *regular member*.

Remark 4 *Let $\mathcal{Q}$ be a quotient family in the cross-free decomposition tree of a set family $\mathcal{F}$. Let Q be a quasi-trivial member of $\mathcal{Q}$. Then, the member of $\mathcal{F}$ which corresponds to Q, namely $A = \bigcup_{V_i \in Q} V_i$, is a cross-free member of $\mathcal{F}$. A consequence is that the fact that A belongs to $\mathcal{F}$ can be known by reading the edge orientations in the cross-free decomposition tree of $\mathcal{F}$.*

Basically, the crucial point is to have control over the regular members. For instance, we say that a family is *prime* if it has no regular members, and from Remark 4, if $\mathcal{Q}$ is prime, there is nothing to do in order to represent the members of $\mathcal{F}$ in correspondence with $\mathcal{Q}$. Otherwise let us inspect the regular members of $\mathcal{Q}$, and argue that, if exist, cross-free members of $\mathcal{Q}$ are nasty cases. (In fact we will see that $\mathcal{Q}$ does not contain any such member.) Roughly, if there is some regular member X of $\mathcal{Q}$ that is also a cross-free member of $\mathcal{Q}$, then X would divide the family into four fractions: those included in X, those included in $\overline{X}$, and the complement of these two fractions. Such a division would make the family, in some sense, less compact and hard to describe. Let alone if the family has more cross-free members among its regular members. Typically, we would rather end up in a situation similar to the one in Theorem 3: either the family consists of very isolated individuals (prime), or it is quite close-knit (complete and circular). A naive way to get rid of the cross-free members consists in removing them from the family. However, such a practice could destroy some important structural property of the family, such as some closure axioms (if A and B cross and $A \cap B$ is cross-free, then removing $A \cap B$ results in a family that is not closed under the intersection of its crossing members). In fact the main purpose of Definition 2 is to eliminate the cross-free members among the regular members without much noise: we stay with the same type of family.

Remark 5 *In a cross-free decomposition tree of any set family $\mathcal{F}$, no regular member of a quotient family $\mathcal{Q}$ can be a cross-free member of $\mathcal{Q}$. Moreover, if $\mathcal{F}$ is a weakly partitive crossing family, resp. union-difference family, then so is $\mathcal{Q}$.*

Typically, for weakly partitive crossing families we would like to prove Lemma 5 below. However, we begin by showing a useful tool. A family $\mathcal{F} \subseteq 2^V$ can also be seen as an undirected hypergraph over vertex set V. Let us define the 2*-graph* of $\mathcal{F}$ as its restriction to size 2 hyperedges: $G_{\mathcal{F}} = (V, E)$ with $E = \{A \in \mathcal{F} \text{ and } |A| = 2\}$.

Lemma 4 *Let $\mathcal{F}$ be a weakly partitive crossing family and let $\mathcal{Q}$ be a quotient family in the cross-free decomposition tree of $\mathcal{F}$. Let X be a regular member of $\mathcal{Q}$. Then, every $x \in X$ has a neighbour in the 2-graph $G_{\mathcal{Q}}$ of $\mathcal{Q}$.*

Proof: Among the regular members of $\mathcal{Q}$ which contain x (for instance X is one such), let A be one with minimum size. If $|A| = 2$, then A is an edge in $G_{\mathcal{Q}}$, and there is nothing to show. Otherwise we will exhibit a contradiction. From Remark 5, there is a member of $\mathcal{Q}$ crossing A (which besides has to be a regular member).

Suppose there is a member of $\mathcal{Q}$ which crosses A and contains x. Let $B \in \mathcal{Q}$ be such a member with minimum size. By the intersection closure $A \cap B \in \mathcal{Q}$. Then, by minimality of A, we have $A \cap B = \{x\}$. By the difference closure, $B \setminus A$ is a member of $\mathcal{Q}$. Moreover, $B \setminus A$ is a regular member of $\mathcal{Q}$ since $|B| \geq |A| > 2$. Therefore, there is a member $C \in \mathcal{Q}$ which crosses $B \setminus A$. If $A \cap C = \emptyset$, then B and C cross, and hence $B \setminus C \in \mathcal{Q}$. Moreover, $B \setminus C$ and A still cross, $B \setminus C$ still contains x, yet $B \setminus C$ is of size strictly smaller than B: this is a contradiction to the minimality of B. Hence, $A \cap C \neq \emptyset$. If $x \notin C$, then $B \setminus C$ belongs to $\mathcal{Q}$, contains x, and crosses A: contradiction by minimality of B. If $x \in C$ then $B \cap C$ belongs to $\mathcal{Q}$, contains x, and crosses A: contradiction by minimality of B.

Therefore, every $B \in \mathcal{Q}$ crossing A will exclude x. Moreover $A \setminus B = \{x\}$ by the difference closure and the minimality of A. It is easy to check that $A \cap B$ is a regular member of $\mathcal{Q}$, and hence there exists $C \in \mathcal{Q}$ crossing $A \cap B$ (then C has to be regular). If $C \subseteq A$ then $x \in C$, contradicting the minimality of A. If $C \not\subseteq A$ and $A \cup C \neq V$, then A and C cross, and hence $A \setminus C$ is a regular member of $\mathcal{Q}$ containing x, contradicting the minimality of A. Finally, $C \not\subseteq A$ and $A \cup C = V$. But then $C \setminus B \neq \emptyset$, and in particular B and C cross. Hence $B \cap C \in \mathcal{Q}$, but this member $B \cap C$ crosses A, and contradicts the minimality of B. □

We denote the complete graph over n vertices by K_n, the path over n vertices by P_n, and the cycle over n vertices by C_n. For any graph G over vertex set $V(G)$, and for any new vertex $v \notin V(G)$, we denote by $G + v$ the disjoint union of G and the one vertex graph made by v.

Lemma 5 *Let $\mathcal{F}$ be a weakly partitive crossing family and let $\mathcal{Q}$ be a quotient family in the cross-free decomposition tree of $\mathcal{F}$. Suppose that $\mathcal{Q}$ is not prime and that the ground set W of $\mathcal{Q}$ has at least 5 elements. Then, $G_{\mathcal{Q}}$ is either K_n, $K_{n-1} + v$, P_n, $P_{n-1} + v$, or C_n. Moreover, if $\mathcal{F}$ is a partitive crossing family, then the three last cases (P_n, $P_{n-1} + v$, and C_n) can not occur.*

Proof: Since $\mathcal{Q}$ is not prime, we deduce from Lemma 4 that $G_{\mathcal{Q}}$ has an edge. Let X be a connected component of $G_{\mathcal{Q}}$ of size at least 2. We first prove that $|X| \geq |W|-1$. Assume this is not the case. Clearly, we have $X \in \mathcal{Q}$ (union closure) and moreover it is a regular member by the assumption. Let A be a member of $\mathcal{Q}$ crossing X, and minimum by size (exists by Remark 5). We would like to inspect $A \setminus X$, which is a member of $\mathcal{Q}$ by the difference closure. Clearly, $|A| > 2$ (maximality of X). Let uv be an edge in $G_{\mathcal{Q}}[X]$ with $u \notin A$ and $v \in A$ (exists by connectivity of $G_{\mathcal{Q}}[X]$ and the fact A and X cross). Then, $\{u,v\}$ and A are crossing members of $\mathcal{Q}$ and hence $A \setminus \{v\}$ is a member of $\mathcal{Q}$ (difference closure). By maximality of A, we deduce that $A \cap X = \{v\}$. With a size check we deduce that $A \setminus X$ is a regular member of $\mathcal{Q}$, and so there exists $B \in \mathcal{Q}$ crossing $A \setminus X$. If moreover B crosses A then both $A \setminus B$ and $A \cap B$ belong to $\mathcal{Q}$, and one of them would contradict the minimality of A. Therefore, $A \cup B = W$. We claim that $v \notin B$ since otherwise we would have $(A \setminus X) \cup B = (A \setminus X) \cup (A \cap X) \cup B = A \cup B = W$ and that would mean B does not cross $A \setminus X$. But now we can check that $\{u,v\}$ and B are crossing members of $\mathcal{Q}$, and hence $B \setminus \{u\}$ is a member of $\mathcal{Q}$ crossing A, and finally $A \setminus (B \setminus \{u\})$ is a member of $\mathcal{Q}$ contradicting the minimality of A. Whence, $|X| \geq |W| - 1$.

To conclude we use a similar technique as the one used in [18, Proof of Lemma 5.4]. Suppose there is in $G_{\mathcal{Q}}$ a vertex v adjacent to at least 3 vertices a, b and c. Clearly, $\{a,v,b\}$ and $\{v,c\}$ are members of $\mathcal{Q}$ (union closure). Since $|W| \geq 5$, these two cross, and hence $\{a,b\}$ is a member of $\mathcal{Q}$ (difference closure). By symmetry, $\{v,a,b,c\}$ induces a clique in $G_{\mathcal{Q}}$, and besides, it is included in X (size argument). Let K be a maximal clique in $G_{\mathcal{Q}}[X]$ and suppose that $K \neq X$. In particular $|K| \geq 4$ because of the clique $\{v,a,b,c\}$. By connectivity there exists a vertex $x \in X \setminus K$ adjacent to a vertex $y \in K$. This y has at least two neighbours (*e.g.*, in K) and so with a similar argument as before, we can prove that x is adjacent to every neighbour of y, contradicting the maximality of K.

Hence X either induces a complete graph, or a graph with max degree 2, *i.e.* a path or a cycle. Now suppose that $|X| = |W| - 1 \geq 4$ and that X induce a cycle. In $G_{\mathcal{Q}}$ let s be the isolated vertex and let x be adjacent to y and z. Then $\{x,y\}$ and $\{x,z\}$ are crossing members of $\mathcal{Q}$ (also because s exists), and hence $\{x,y,z\}$ is a member of $\mathcal{Q}$. But $X \setminus \{x\}$ is also a member of $\mathcal{Q}$ (by union closure on the edges of $G_{\mathcal{Q}}$ that are neither xy nor xz). Since $\{x,y,z\}$ and $X \setminus \{x\}$ cross, $\{y,z\}$ is a member of $\mathcal{Q}$, contradicting the fact $G_{\mathcal{Q}}[X]$ is a cycle of length at least 4.

Finally, suppose that $\mathcal{Q}$ is also partitive crossing, and there is in $G_{\mathcal{Q}}$ a vertex x adjacent to y and z. Then yz is an edge of $G_{\mathcal{Q}}$ (symmetric difference closure). $\square$

Recall from Remark 4 that quasi-trivial members of $\mathcal{Q}$ in fact correspond to cross-free members of $\mathcal{F}$ and are already encoded by the cross-free decomposition tree of $\mathcal{F}$: the crucial point is to represent the regular members. Let W denote the ground set of $\mathcal{Q}$. By the union closure, $G_{\mathcal{Q}}$ is K_n if and only if $\mathcal{Q} = 2^W \setminus \{\emptyset\}$. If $G_{\mathcal{Q}}$ is $K_{n-1} + v$ then $\mathcal{Q}$ and $2^{W \setminus \{v\}} \cup \{v\}$ have exactly the same regular members, since from Lemma 4 no regular member of $\mathcal{Q}$ can contain the isolated vertex v. If $G_{\mathcal{Q}}$ is a path $(v_1, \dots v_n)$ of size $n \geq 5$, then all intervals $\{v_i, \dots, v_j\}$ $(1 \leq i \leq j \leq n)$ of the path belong to $\mathcal{Q}$ by the union closure. On the other hand, it is straightforward to check that $\mathcal{Q}$ has no other members than these intervals. Similarly, if $G_{\mathcal{Q}}$ is the disjoint union of a path and an isolated vertex, then the regular members of $\mathcal{Q}$ are exactly the intervals of the path having the right size. Finally, if $G_{\mathcal{Q}}$ is a cycle, then it is straightforward to check that $\mathcal{Q}$ is exactly the family of all circular intervals of the cycle. We have proved that

Theorem 5 *There is a $\Theta(|V|)$ space representation of any weakly partitive crossing family over V. The representation can be based on the cross-free decomposition tree of the family: a quotient family $\mathcal{Q} \subseteq 2^W$ in the cross-free decomposition tree $\mathcal{T}_{\mathcal{C}}$ of a weakly partitive crossing family $\mathcal{F} \subseteq 2^V$ satisfies one and only one of the following*

- basic*: $W \leq 4$;*
- prime*: $W \geq 5$ and $\mathcal{Q}$ has no regular members;*
- complete*: $W \geq 5$ and either $\mathcal{Q} = 2^W \setminus \{\emptyset\}$ or there is $v \in W$ such that $\mathcal{Q}$ and $2^{W \setminus \{v\}} \cup \{v\}$ have exactly the same regular members;*
- linear*: $W \geq 5$ and either there is a linear ordering of W such that the members of $\mathcal{Q}$ are exactly the intervals of the ordering, or there is $v \in W$ and a linear ordering of $W \setminus \{v\}$ such that $\mathcal{Q}$ and the family over W containing all intervals of the ordering have exactly the same regular members;*
- circular*: $W \geq 5$ and there is a circular ordering of W such that the members of $\mathcal{Q}$ are exactly the circular intervals of the ordering.*

Moreover, if $\mathcal{F}$ is a partitive crossing family, then $\mathcal{Q}$ can not be linear nor it can be circular.

We now focus on union-difference families. Ideally, we would like to prove similar statements as in Lemma 5: if a quotient family is not prime then its 2-graph has to be of simple types, such as a clique, a path, a cycle, etc. Unfortunately, the crucial Lemma 4 does not seem to hold for union-difference families, making such an approach *a priori* difficult. Instead, we will use another approach here. As mentioned before, the main purpose of Definition 4 (first appeared in [4]) was not to study weakly partitive families as what was done in Section 3. Rather than that, simply-linkedness was introduced to help classifying the quotient families in the cross-free decomposition tree of a union-difference family:

Theorem 6 *There is a $O(|V|^2)$ space representation of any union-difference family over V. The representation can be based on the cross-free decomposition tree of the family: a quotient family $\mathcal{Q} \subseteq 2^W$ in the cross-free decomposition tree $\mathcal{T}_{\mathcal{C}}$ of a union-difference family $\mathcal{F} \subseteq 2^V$ satisfies one and only one of the following*

- prime*: $\mathcal{Q}$ is simply-linked and has no regular members;*
- complete*: $\mathcal{Q} = 2^W \setminus \{\emptyset\}$;*
- linear*: there is a linear ordering of W such that the members of $\mathcal{Q}$ are exactly the intervals of the ordering;*
- circular*: there is a circular ordering of W such that the members of $\mathcal{Q}$ are exactly the circular intervals of the ordering;*

- recursive*: $\mathcal{Q}$ is not simply-linked and for w being the guard of $\mathcal{Q}$ (cf. Definition 4) we have that $\mathcal{Q}' = \mathcal{Q} \setminus \{W, \{w\}\}$ is a union-difference family over $W \setminus \{w\}$; a consequence is that the regular members of $\mathcal{Q}$ can be represented by the cross-free decomposition of $\mathcal{Q}'$.*

Proof: If $\mathcal{Q}$ is not simply-linked then it is straightforward to check the properties of the last item (the recursive case). Otherwise $\mathcal{Q}$ could be prime and there is nothing to show. If $\mathcal{Q}$ is simply-linked but it is not prime, then we will prove in Lemma 6 and Lemma 7 (both below) that $G_{\mathcal{Q}}$ is a connected graph. Next, we complete the classification by proving in Lemma 8 (below) that if $G_{\mathcal{Q}}$ is connected, then it is either a clique, a path, or a cycle. Finally, we prove in Lemma 9 (below) the $O(|V|^2)$ space complexity of the encoding. □

The remaining of the section is to prove Theorem 6. A *chain* of length k in a family $\mathcal{F}$ is a sequence $(A_1, \ldots, A_k)$ of members of $\mathcal{F}$ such that $A_i \between A_{i+1}$ for all i, and $A_i \cap A_j = \emptyset$ for all $|i - j| > 1$. The chain is *covering* if $A_1 \cup \cdots \cup A_k = V$, and *irreducible* if $|A_i| = 2$ for all $1 \leq i \leq k$. An irreducible and covering chain of $\mathcal{F}$ can also be seen as a Hamiltonian path in the 2-graph $G_{\mathcal{F}}$, which would imply its connectivity, and enable the use of Lemma 8.

Lemma 6 *Let $\mathcal{F}$ be a union-difference family and let $\mathcal{Q}$ be a quotient family in the cross-free decomposition tree of $\mathcal{F}$. Suppose that $\mathcal{Q}$ is simply-linked but not prime. Then $\mathcal{Q}$ has a length 3 covering chain.*

Proof: Since $\mathcal{Q}$ is not prime, let A be a regular member of $\mathcal{Q}$, maximum by size. Since $\mathcal{Q}$ is a quotient family there exists $B \in \mathcal{Q}$ such that A and B cross (cf. Remark 5). Then, $A \cup B \in \mathcal{Q}$ (union closure), and moreover $A \cup B$ is not a regular member by maximality of A. However, $A \cup B$ cannot be trivial since A and B cross. Since $\mathcal{Q}$ is simply-linked, $A \cup B$ overlaps some $C \in \mathcal{Q}$. Here, all cases lead to either $D = C \cup B \setminus A$ or $E = C \cup A \setminus B$ is a member of $\mathcal{Q}$. Then, either (A, B, D) or (B, A, E) is a covering chain of length 3. □

Lemma 7 *Let $\mathcal{F}$ be a union-difference family and let $\mathcal{Q}$ be a quotient family in the cross-free decomposition tree of $\mathcal{F}$. Suppose that $\mathcal{Q}$ is simply-linked but not prime. Suppose moreover that $\mathcal{Q}$ has a covering chain of length at least 3. Then $\mathcal{Q}$ has an irreducible covering chain (and hence $G_{\mathcal{Q}}$ is connected).*

Proof: Let $\mathcal{A} = (A_1, \ldots, A_k)$ be a covering chain of $\mathcal{Q}$ with $k \geq 3$. We take k maximum. Assume for some $1 < i < k$ that $A_i \setminus (A_{i-1} \cup A_{i+1}) \neq \emptyset$. In this case $B = A_i \setminus A_{i+1}$ and $C = A_i \setminus A_{i-1}$ are overlapping members of $\mathcal{Q}$ (see Figure 5(a)). Then, replacing $\mathcal{A}$ with $(A_1, \ldots, A_{i-1}, B, C, A_{i+1}, \ldots, A_k)$ would improve k. Hence, $A_i \setminus (A_{i-1} \cup A_{i+1}) = \emptyset$ for all $1 < i < k$.

We now assume that $|A_i| > 2$ for some $1 < i < k$. Then at least one among $B = A_i \setminus A_{i+1}$ and $C = A_i \setminus A_{i-1}$ is a regular member of $\mathcal{Q}$, and hence not cross-free because $\mathcal{Q}$ is a quotient family. By symmetry we suppose it was B. Let $D \in \mathcal{Q}$ cross B. We show in all cases a contradiction as follows (see also Figure 5(b)).

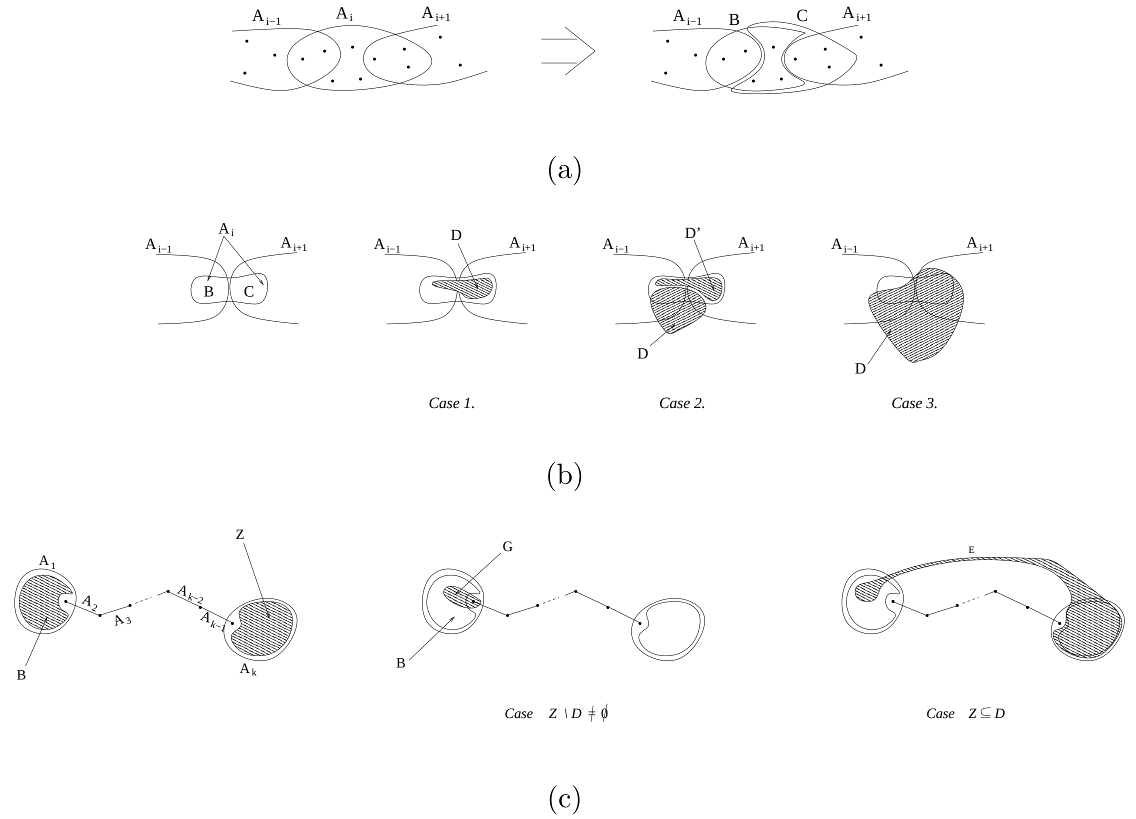


Fig. 5. Illustration for the proof of Lemma 7.

- *Case 1:* $D \subseteq A_i$. In particular, D and A_{i-1} overlap. Let $E = A_{i-1} \setminus D$, we can improve k by replacing $\mathcal{A}$ with $(A_1, \ldots, A_{i-2}, E, B, D, A_{i+1}, \ldots, A_k)$.
- *Case 2:* $D \setminus A_i \neq \emptyset$ and $C \setminus D \neq \emptyset$. Then, we are conducted to *Case 1* by replacing D with $D' = A_i \setminus D$.
- *Case 3:* $D \setminus A_i \neq \emptyset$ and $C \subseteq D$. We define the left and right as $L = A_1 \cup \cdots \cup A_{i-2}$ and $R = A_{i+1} \cup \cdots \cup A_k$. Notice that $L \cup R = \overline{B}$. Since D and B cross, there is some element in either L or R that does not belong to D. If it was L, replacing D with $A_i \setminus (A_{i-1} \setminus (D \setminus L))$ leads back to *Case 1*. If it was R, the same can be done with $A_i \setminus (D \setminus R)$.

Hence, $|A_i| = 2$ for all $1 < i < k$. Now, assume that $|A_1| > 2$, and let $D \in \mathcal{Q}$ cross $B = A_1 \setminus A_2$. Let $Z = A_k \setminus A_{k-1}$. We will examine whether $Z \setminus D \neq \emptyset$ or $Z \subseteq D$ (see Figure 5(c)). In the first case, let $E = A_3 \cup \cdots \cup A_k$ and $F = D \cup A_2 \cup \cdots \cup A_{k-1}$: they overlap. Then, $G = F \setminus E$ is a member of $\mathcal{Q}$, and replacing $\mathcal{A}$ with $(B, G, A_2, \ldots, A_k)$ would improve k. In the second case, since D and B cross, there is some element in $A_2 \cup \cdots \cup A_{k-1}$ that does not belong to D. In other words, $A_2 \cup \cdots \cup A_{k-1}$ and D overlap. Then, $E = D \setminus (A_2 \cup \cdots \cup A_{k-1})$ is a member of $\mathcal{Q}$ which contains Z. The fact that $E \in \mathcal{Q}$ implies $(E, A_1, \ldots, A_{k-1})$ is a chain of $\mathcal{Q}$. The fact that E contains Z implies the chain is covering. Moreover, it is of length k, *i.e.* of maximum length. However, from the last paragraph, this chain cannot have A_1 with more than two elements. Therefore, $|A_1| = 2$. Then, by symmetry we obtain $|A_k| = 2$, and $\mathcal{A}$ is an irreducible covering chain. $\square$

To conclude, we use a beautiful technique which was discovered for weakly partitive families (see, *e.g.*, [18, Proof of Lemma 5.4]) but which only required the union and difference closures (the almost same technique was also used as part of the proof of Lemma 5).

Lemma 8 *(cf. [18] with partitive families) Let $\mathcal{F}$ be a union-difference family. If its 2-graph $G_{\mathcal{F}}$ is connected then $G_{\mathcal{F}}$ is either a clique, a path, or a cycle.*

Proof: The proof given in [18] is as follows. Suppose that $G_{\mathcal{F}}$ has a vertex x with degree at least 3, and let y, z, t be three distinct neighbours of x. In other words, $\{x, y\}$ and $\{x, z\}$ are members of $\mathcal{F}$, and so is $\{x, y, z\}$ by the union closure. But $\{x, t\}$ is also a member of $\mathcal{F}$. By the difference closure we deduce that $\{y, z\}$ is an edge of $G_{\mathcal{F}}$. Likewise, we can deduce that $\{x, y, z, t\}$ induces a clique in $G_{\mathcal{F}}$. Now, let v be a vertex that is connected to the previous clique at some point, say t. Then, by a similar argument on the fact that t is of degree at least 3, we can show that v is connected to all other vertices of the clique. Thus the previous clique plus vertex v form a bigger clique, and so on. The connectivity of $G_{\mathcal{F}}$ then can be used to conclude that the whole graph $G_{\mathcal{F}}$ is a clique. Finally, the only connected graphs of degree at most 2 are paths and cycles. $\square$

Lemma 9 *Let $\mathcal{F} \subseteq 2^V$ be a union-difference family. Labelling the internal nodes of the cross-free decomposition tree of $\mathcal{F}$ according to Theorem 6 will result in an $O(|V|^2)$ global space encoding.*

Proof: By induction on $n = |V|$. Let $f(n)$ be the maximum size of all such decomposition trees of n leaves. Obviously, $f(1)$ and $f(2)$ are non null constants. Let $f(k) \leq \alpha \times k^2$ hold for all $k < n$. We suppose without loss of generality that α is greater than any other constant in this proof. Let us consider a decomposition tree of n leaves and let N be the set of its internal nodes. For each $i \in N$, let n_i be its degree. The label of i is either of constant size (*cf.* prime and complete nodes), of linear size on n_i (*cf.* linear and circular nodes), or of size bounded by $f(n_i - 1) + \beta$ (*cf.* recursive nodes). In all cases, it is bounded by $\alpha \times (n_i - 1)^2 + \alpha$ since $n_i \geq 3$ and $\alpha \geq \beta$. The total size of leaves, edges, and orientations is linear in n, hence bounded by $\alpha \times n$. We deduce that

$$f(n) \leq \alpha \times \left(\sum_{i \in N} ((n_i - 1)^2 + 1) + n \right) \leq \alpha \times \left(\sum_{i \in N} (n_i - 1)^2 + n' + n \right),$$

where $n' = |N|$. Notice that $\sum_{i \in N} n_i = n + 2 \times (n' - 1)$ (the n pendant edges are counted once while other edges are counted twice). In other words, $\sum_{i \in N} (n_i - 1) = n + n' - 2$. Let $S = n + n' - 2$ The greatest value that $\sum_{i \in N} (n_i - 1)^2$ can reach happens when one among the n_i gets the greatest value possible. Since $n_i - 1 \geq 2$, we have $\sum_{i \in N} (n_i - 1)^2 \leq (n' - 1) \times 2^2 + (S - (n' - 1) \times 2)^2$. Then, $f(n) \leq \alpha \times (n^2 + n'^2 + 5n' + n(1 - 2n') - 4)$. Besides, that there are no degree 2 nodes in the tree provides us with $n \geq n' + 2$. Moreover, it is clear that $1 - 2n' \leq 0$. Hence, $n(1 - 2n') \leq (n' + 2)(1 - 2n')$, which is also $n(1 - 2n') \leq -2n'^2 - 3n' + 2$. Therefore, $f(n) \leq \alpha \times (n^2 - n'^2 + 2n' - 2) \leq \alpha \times (n^2 - (n' - 1)^2 - 1) \leq \alpha \times n^2$. $\square$

5 Sesquimodular decompositions

Modular decomposition has become a classical topic in graph theory [10,18,27,35], as well as some of its generalizations [13,15,32,33,36]. A module in an undirected graph is a vertex subset M such that $\forall x, y \in M$, $N(x) \setminus M = N(y) \setminus M$. As having been studied in other fields, this notion also appears under various names, including intervals, externally related sets, autonomous sets, partitive sets, clans, etc. Direct applications of modular decomposition include tractable constraint satisfaction problems [11], computational biology [26], graph clustering for network analysis, and graph drawing. This rich research field relies on the nice combinatorial properties of modules. To name but one, the family of modules of any graph is a partitive family, and therefore can be efficiently represented by a tree, the modular decomposition tree of the graph [10,18,35]. Now, in the area of social networks, several vertex partitioning have been introduced in order to formalize the idea of finding regularities [45]. Modular decomposition gives such a partitioning. But the notion of a module (of undirected graphs) seems to be too restrictive for real-life applications. On the other hand, although the concept of a role [20] seems promising, its computation is unfortunately NP-hard [21]. It is then a natural question to look for *relaxed*, but *tractable*, decomposition schemes related to modular decomposition.

We investigate the case of digraphs, and their generalization to 2-structures. Therein, the major tractable decomposition that has been addressed in the literature so far is the so-called clan decomposition [2] : a clan in a digraph is a vertex subset M such that $\forall x, y \in M$, $N^-(x) \setminus M = N^-(y) \setminus M$ and $N^+(x) \setminus M = N^+(y) \setminus M$. In order to further decompose, we propose a weakened definition. Fortunately enough, we still obtain a well-structured variation, thanks to weakly partitive crossing families and union-difference families.

Definition 6 (Digraph sesquimodule) In a digraph $G = (V, A)$ we say that $M \subseteq V$ is a *sesquimodule* if:

- $\forall x, y \in M$, $N^-(x) \setminus M = N^-(y) \setminus M$, and
- $\forall x, y \in M$, either $N^+(x) \setminus M = N^+(y) \setminus M$ or $N^+(x) \setminus M = \overline{N^+(y)} \setminus M$.

Roughly, the classical generalization of a graph module to digraph clans asks for two full conditions, one on in-neighbours and one on out-neighbours. Let us consider a clan as a double-module. Now, in the new definition, there is a full condition on in-neighbours, and a relaxed one on out-neighbours: the exterior still has to be partitioned w.r.t. the out-neighbourhood homogeneously, however, the order of the partition classes is irrelevant. We qualify the relaxed condition as a half-condition and this is the reason for the terminology of a module-and-half, namely a sesquimodule.

[2] A clan of a digraph is also called a module in [34]. However, for the sake of clarity, we will *not* use this terminology throughout the paper. Instead, we simply refer to the clans of a digraph, according to their introduction by Ehrenfeucht and Rozenberg [19] (an excellent introduction to this topic is [18]).

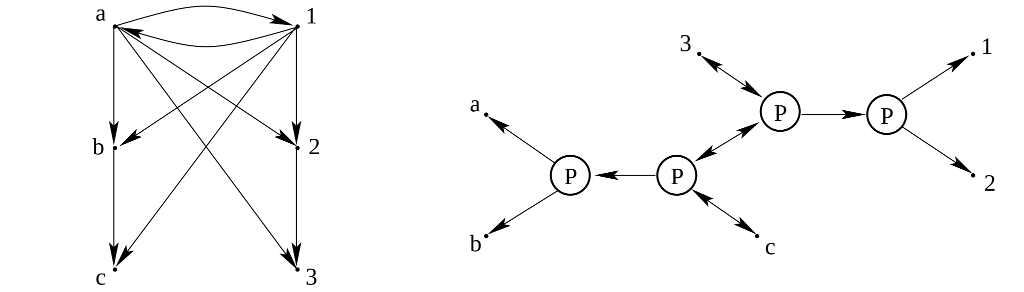
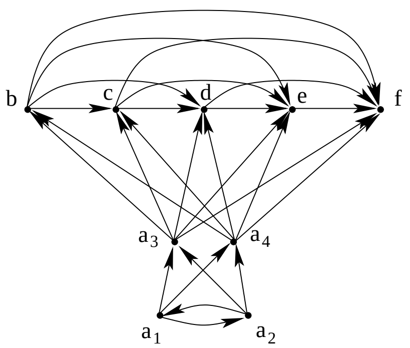
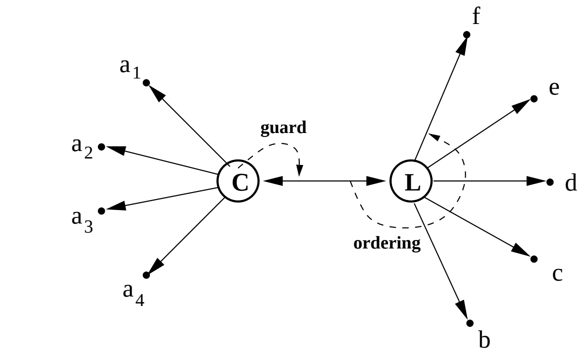


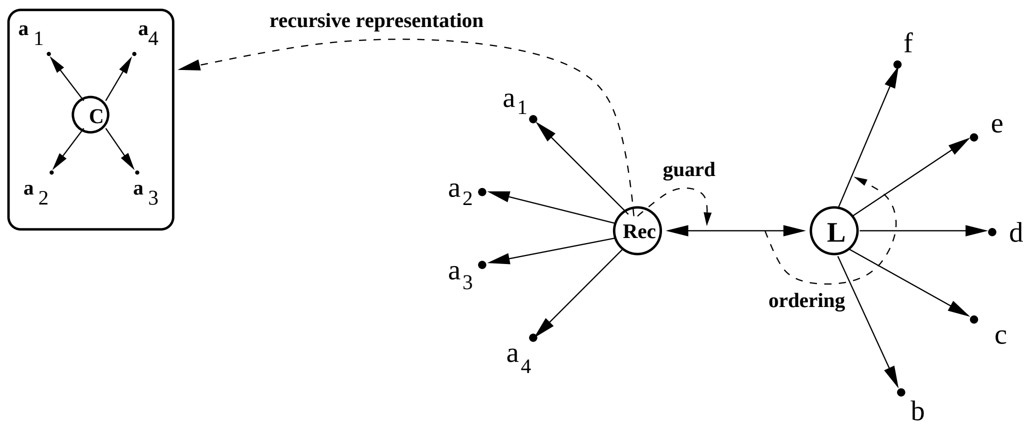
Fig. 6. A digraph with no non-trivial clans and its sesquimodular decomposition tree.



(a) A digraph.



(b) Weakly partitive crossing sesquimodular decomposition tree.



(c) Union-difference sesquimodular decomposition tree.

Fig. 7. Sesquimodular decomposition. Some sesquimodules of the digraph are: all subsets of $A = \{a_1, a_2, a_3, a_4\}$, $A \cup \{b\}$, $A \cup \{b, c\}$, $A \cup \{b, c, d\}$, $\{b, c, d\}$, $\{b, c, d, e\}$, $\{b, c, d, e, f\}$, $\{c, d, e\}$, ...

Note that when the digraph is an undirected graph, the three notions of a module, a sesquimodule, and a clan are equivalent. For digraphs however, we show in Figure 6 that the generalization from clans to sesquimodules is strict. In Figure 7 we give an example of two sesquimodular decomposition schemes, depending on whether we follow the statement of Theorem 7, or that of Theorem 8.

A $2-$structure is an edge-colored complete digraph: a pair (V, C) where V is a finite set and $C : V^2 \setminus \{(v, v) : v \in V\} \rightarrow \mathbb{N}$. Note that graphs, digraphs and tournaments are special cases of 2-structures with $C(u, v) = 1$ if $(u, v) \in A$, and $C(u, v) = 0$ otherwise. Digraph clans can be generalized to clans of a $2-$structure: M is a clan if $\forall x, y \in M$, $\forall c \in \mathbb{N}$, $N^c(x) \setminus M = N^c(y) \setminus M$, where $N^c(x) = \{z : C(x, z) = c\}$. Likewise, digraph sesquimodules can be generalized to the sesquimodules of a $2-$structure: M is a sesquimodule of a $2-$structure if it holds two following conditions. For all $x, y \in M$, for all $s \notin M$, the arcs (s, x) and (s, y) are of the same color. For all $x, y \in M$, for all $s, t \notin M$, (x, s) and (x, t) are of the same color if and only if (y, s) and (y, t) are of the same color. This has an equivalent formulation. We say that v is *uniform* to x and y if $C(v, x) = C(v, y)$, and this is denoted by $v|xy$. Otherwise, we say that v is a *splitter* of $\{x, y\}$, and denote this fact by $\overline{v|xy}$. For a partition $\mathcal{P}$ and $V \subseteq X$, let $\mathcal{P} \cap V = \{P \cap V : P \in \mathcal{P} \text{ and } P \cap V \neq \emptyset\}$, and let $\mathcal{P} \setminus V = \mathcal{P} \cap \overline{V}$.

Definition 7 (2-structure sesquimodule) Let $G = (V, C)$ be a 2-structure. For any $u \in V$, there is a unique partition $\text{Part}(u) = \{M_1, M_2, \ldots, M_k\}$ of $V \setminus \{u\}$ such that for all $v, w \in V \setminus \{u\}$, $C(u, v) = C(u, w)$ if and only if v and w belong to the same M_i. Then, $M \subseteq V$ is a *sesquimodule* if and only if we have both:

- $\forall x, y \in M$ and $v \in V \setminus M$, $C(v, x) = C(v, y)$
- $\forall x, y \in M$, $\text{Part}(x) \setminus M = \text{Part}(y) \setminus M$.

We will prove the following theorems.

Theorem 7 (2-structures uniqueness decomposition theorem) *There is a unique unrooted tree associated to a 2-structure $G = (V, C)$ such that: the leaves of the tree are in one-to-one correspondence with the vertices of G; the edges of the tree are oriented; the internal nodes of the tree are marked with at most 5 types of labels; and all sesquimodules of G can be generated from this tree without the knowledge of the 2-structure. The size of this tree and its labels is in $O(|V|^2)$.*

Note that a generating object of quadratic size is not an improvement in space in itself since the initial 2-structure is already such. However, the one given for sesquimodules in Theorem 7 follows a tree structure, and furthermore Lemma 10 below proves that the sesquimodules form a union-difference family. These are instructive structural properties (*cf.* also the open question whether the quadratic space complexity of union-difference families is tight or not). More importantly, when the 2-structure is a digraph, we have the following major improvement.

Theorem 8 (Digraphs uniqueness decomposition theorem) *There is a unique unrooted tree associated to a digraph $G = (V, A)$ such that: the leaves of the tree* *are in one-to-one correspondence with the vertices of G; the edges of the tree are oriented; the internal nodes of the tree are marked with at most 5 types of labels; and all sesquimodules of G can be generated from this tree without the knowledge of the graph. The size of this tree and its labels is in $O(|V|)$.*

Theorems 7 and 8 follow from the two simple facts that:

Lemma 10 *The sesquimodules of a 2-structure form a union-difference family. Furthermore there are no circular nodes in its decomposition tree.*

Proof: Let $G = (V, C)$ be a 2-structure. Clearly, the trivial vertex subsets are sesquimodules of G. Let X and Y be two overlapping sesquimodules of G. It follows straight from definition that $X \cup Y$ is a sesquimodule. We only need to prove that $Z = X \setminus Y$ is also a sesquimodule.

First suppose that there exist an exterior vertex $s \notin Z$ and two vertices $x, y \in Z$ s.t. s is a splitter for $\{x, y\}$. Since X is a sesquimodule s belongs to $X \cap Y$. Moreover, that X and Y overlap implies there is a vertex t belonging to $Y \setminus X$. Notice that $s, t \in Y$ and $x, y \notin Y$. Since Y is a sesquimodule, t is a splitter for $\{x, y\}$. But then X no more is a sesquimodule as $t \notin X$ and $x, y \in X$. Hence, for all $x, y \in Z$ and $s \notin Z$, $C(s, x) = C(s, y)$.

Now let $x, y \in Z$ and $s, t \notin Z$. We need to prove that $x|st \Leftrightarrow y|st$. If none of s and t belong to X, that X is a sesquimodule allows to conclude. If both s and t belong to Y, that Y is a sesquimodule allows to conclude. By symmetry, the only remaining case is when $s \in X \cap Y$ and $t \notin X \cup Y$. In this case, let $u \in Y \setminus X$. Since X is a sesquimodule, we already have $x|tu \Leftrightarrow y|tu$, but we would like the same property with vertex u replaced by vertex s. For this, notice that $x \notin Y$, but $s, u \in Y$, and Y is a sesquimodule. Therefore, $C(x, u) = C(x, s)$. Likewise, $C(y, u) = C(y, s)$. Then, combining the two latter facts and $x|tu \Leftrightarrow y|tu$ leads to the desired property.

Finally, a circular sesquimodule quotient node would be a complete one. □

Lemma 11 *The sesquimodules of a digraph form a weakly partitive crossing family. Furthermore there are no circular nodes in its decomposition tree.*

Proof: Let $G = (V, A)$ be a digraph. Let X and Y be two crossing sesquimodules of G. By Lemma 10, $X \cup Y$ and $X \setminus Y$ are sesquimodules of G. We only need to prove that $Z = X \cap Y$ is a sesquimodule.

It is straightforward that for every $u \notin Z$ and for every $x, y \in Z$, $(u, x) \in A \Leftrightarrow (u, y) \in A$. Now there is a partition $\{A, B\}$ of $\overline{X}$ and a partition $\{A', B'\}$ of $\overline{Y}$ such that for every $z \in Z$, $\mathrm{Part}(z) \setminus X = \{A, B\}$ and $\mathrm{Part}(z) \setminus Y = \{A', B'\}$. Since X and Y cross, there is a $v \in \overline{X \cup Y}$. Suppose w.l.o.g. that $v \in A \cap A'$. For every $z \in Z$ either $v \in N^+(z) \setminus Z$ or $v \notin N^+(z) \setminus Z$. Thus $N^+(z) \setminus Z$ is either $A \cup A'$ or $B \cup B'$, and so $\mathrm{Part}(z) \setminus Z = \{A \cup A', B \cup B'\}$.

Finally, a circular sesquimodule quotient node would be a complete one. □

6 Computing the sesquimodular decomposition tree of a 2-structure

This section describes a brute-force algorithm to compute in $O(n^7)$ time the sesquimodular decomposition tree of a given 2-structure $G = (V, C)$, where $n = |V|$. In Section 7 we will improve this downto $O(n^3)$ when the 2-structure is a digraph. Our algorithm in this section borrows ideas from [24]. Let us assume from now on that $C(u, v) \leq n^2$ for every $u, v \in V$. We will constantly need to test if a given subset is a sesquimodule.

Lemma 12 *One can test in $O(n^2)$ if a given vertex subset X is a sesquimodule of a given 2-structure $G = (V, C)$.*

Proof: We first check for every vertex $y \notin X$ and every vertex $x \in X$ if for every $x' \in X$, $C(y, x') = C(y, x)$. In a second step we check for $x \in X$ and $x' \in X$ if $\text{Part}(x) \setminus X = \text{Part}(x') \setminus X$. This can be done in $O(n^2)$ time if we suppose that $C(u, v) \leq n^2$ for every $u, v \in V$. □

A set X *separates* a set Y if $Y \cap X$ and $Y \setminus X$ are both non-empty. The following lemma and corollary show how one can find a regular member of the family of sesquimodules.

Lemma 13 *Let $\mathcal{P}$ be a partition of V and let $A, B \in \mathcal{P}$. One can compute in $O(n^3)$ time the unique maximal sesquimodule S such that $A \subsetneq S$, $B \subseteq \overline{S}$ and which does not separate any set in $\mathcal{P}$.*

Proof: Firstly note that there are no distinct maximal sesquimodules S and S' satisfying the required properties, otherwise S would cross S', and $S \cup S'$ would be a bigger sesquimodule satisfying these properties. We take an arbitrary $x \in A$ and start with $Y = B$. As long as there is a $x' \in V \setminus Y$ such that either $\text{Part}(x) \cap Y \neq \text{Part}(x') \cap Y$ or $C(x', y') \neq C(x, y)$ for a $y' \in Y$, we add the set in $\mathcal{P}$ containing x' to Y. When no such x' exists anymore, $V \setminus Y$ is the maximal sesquimodule with the property. About complexity issues, finding such x' can be done trivially in $O(n^2)$, and the algorithm repeats the loop at most $O(n)$ times. Thus the total running time is $O(n^3)$. □

Applying $|\mathcal{P}| - 2$ times the previous procedure on $(\mathcal{P} \setminus \{B, D\}) \cup \{B \cup D\}$, for every $D \in \mathcal{P} \setminus \{A, B\}$, we get the following corollary.

Corollary 2 *One can compute in $O(n^4)$ time a sesquimodule S such that $A \subsetneq S$, $B \subsetneq \overline{S}$ and S does not separate any $C \in \mathcal{P}$.*

Now we present an $O(n^5)$ algorithm which finds every cross-free sesquimodule of a 2-structure G with at most $O(n)$ calls to the procedure described in Corollary 2.

Lemma 14 *Given a 2-structure $G = (V, C)$ and a partition $\mathcal{P}$ of V, one can compute in $O(n^5)$ time a family $\mathcal{S}$ of sesquimodules none of which will cross any set in $\mathcal{P}$. Moreover, $\mathcal{S}$ has the following property: for every sesquimodule S which does not separate any set in $\mathcal{P}$, and such that $\{S, \overline{S}\} \cap \mathcal{P} = \emptyset$, either $S \in \mathcal{S}$ or* *there is $S' \in \mathcal{S}$ such that S' crosses S. In particular, $\mathcal{S}$ contains every cross-free sesquimodule which does not separate any set in $\mathcal{P}$, and such that $\{S, \overline{S}\} \cap \mathcal{P} = \emptyset$.*

Proof: While $|\mathcal{P}| > 3$, we take $A, B \in \mathcal{P}$. We test if there is a sesquimodule X which does not separate any $D \in \mathcal{P}$, and such that either $A \subseteq X$ and $B \subseteq \overline{X}$, or $B \subseteq X$ and $A \subseteq \overline{X}$. To do that, we call two times the sub-routine of Corollary 2. If such X exists, we add X into $\mathcal{S}$ and we recurse on the partitions $\mathcal{P}_1 = \{D \in \mathcal{P} : D \in X\} \cup \{\overline{X}\}$ and $\mathcal{P}_2 = \{D \in \mathcal{P} : D \in \overline{X}\} \cup \{X\}$. Otherwise we add the possible sesquimodules among the set $\{A \cup B, \overline{A \cup B}\}$ into $\mathcal{S}$ (using the procedure of Lemma 12), and we recurse on $\mathcal{P}' = (\mathcal{P} \setminus \{A, B\}) \cup \{A \cup B\}$.

We address time complexity first. There is at most $O(n)$ calls to the sub-routine described in Corollary 2 (in fact it follows from a straightforward induction that there are at most $2 \times (|\mathcal{P}| - 3)$ such calls). Thus the overall running time is $O(n^5)$.

About correctness of the algorithm, clearly $\mathcal{S}$ is a family of non crossing sets. We show the property by induction on $|\mathcal{P}|$. If $|\mathcal{P}| \leq 3$ then such S cannot exists, and the empty family has the desired property. Now suppose that $|\mathcal{P}| > 3$. Suppose that the sub-routine of Corollary 2 finds the sesquimodule X. If $X \in \{S, \overline{S}\}$ or X crosses S then the property is satisfied. Otherwise there is a unique $i \in \{1, 2\}$ such that S does not separate any $D \in \mathcal{P}_i$. By induction, the procedure will find S, $\overline{S}$, or a sesquimodule crossing S. Suppose now that no such X exists. Then either $A \cup B \subseteq S$ or $A \cup B \subseteq \overline{S}$. If $S = A \cup B$ or $\overline{S} = A \cup B$, then S will be immediately found by the sub-routine of Lemma 12. Otherwise, S does not cross any $D \in (\mathcal{P} \setminus \{A, B\}) \cup \{A \cup B\}$. Thus by induction, the procedure will find S or a sesquimodule crossing S. □

Using the procedure of the previous lemma on $\mathcal{P} = \{\{v\} : v \in V\}$, one can compute in $O(n^5)$ time a family $\mathcal{S}$ containing every regular cross-free sesquimodule of G (and potentially something else). Moreover, $\mathcal{S}$ is a cross-free family (it can actually be seen as a subdivision of the sesquimodular decomposition tree of G). Accordingly, $\mathcal{S}$ can be represented by an unrooted tree $T_{\mathcal{S}}$ such that the leaves of $T_{\mathcal{S}}$ are in bijection with V, and there is a bijection between $\mathcal{S}$ and internal edges in $T_{\mathcal{S}}$.

We briefly show how to find non-cross free members of $\mathcal{S}$. Suppose that S is a non cross-free sesquimodule. Then there is a sesquimodule S' which crosses S. Let $(a, b, c) \in (S \cap S') \times (S \cap \overline{S'}) \times (\overline{S} \cap \overline{S'})$. Using the algorithm described in Lemma 13, one can check in $O(n^3)$ if there is a sesquimodule S'' with $\{b, c\} \subseteq \overline{S''}$ and $a \in S''$ (and thus S'' crosses S). Now we fix $a \in S$, and for every pair $(b, c) \in S \times (V \setminus S)$, we check if there is either a sesquimodule S_1 with $\{b, c\} \subseteq \overline{S_1}$ and $a \in S_1$, or a sesquimodule S_2 with $\{a, c\} \subseteq \overline{S_2}$ and $b \in S_2$. If such a S_1 or S_2 exists, S is not cross-free. Otherwise S is cross-free by the previous observation.

Let us come back to $\mathcal{S}$ as being the output of the algorithm described in Lemma 14. We check for every $S \in \mathcal{S}$ if S is cross-free with the latter routine, and we compute the family $\mathcal{S}'$ of cross-free sesquimodules of G, and thus the sesquimodular decomposition tree of G. The overall running time is $O(n^6)$.

The only remaining thing is to type the nodes. The main difficulty is how to test for nodes that are not simply-linked. Actually, we avoid this test by elimination of cases. Let α be a internal node of the decomposition tree. We compute the 2-graph of the quotient w.r.t. node α (quadratic number of tests for membership). If this is a clique or a path we conclude accordingly, and stop. Now we compute all quasi-trivial members of the quotient (there are at most as many quasi-trivial members as incident edges of the node). If there are more than one or none of such, report a prime node, and stop. Else either the node is prime or it is not simply-linked with that unique quasi-trivial member which is overlap-free. Let $\{c\}$ be the complement of the unique quasi-trivial member. Assume that the node α is not simply-linked and recursively compute the decomposition tree of the quotient excluding $\{c\}$. If the latter tree is anything except a single prime node then node α effectively was not simply-linked, we conclude and stop. The latter tree is a single prime node. If there is some quasi-trivial member therein then node α effectively was not simply-linked, we conclude and stop. Otherwise node α was simply-linked. We report a prime node. Without recursive calls the process is in $O(n^6)$ time. Then, an inductive argument similar to the proof of Lemma 9 gives an $O(n^7)$ time bound. To sum up we have:

Theorem 9 *The sesquimodular decomposition tree of a given 2-structure $G = (V, C)$ can be computed in $O(|V|^7)$ time.*

7 Computing the sesquimodular decomposition tree of a digraph

We show in this section how the time complexity of the computation of the sesquimodular decomposition can be improved when the input is a digraph.

Lemma 15 *One can test in $O(m)$ if a given vertex subset X is a sesquimodule of a given digraph.*

Proof: For every vertex $v \notin X$ with $N^+(v) \cap X \neq \emptyset$, we check if $X \subseteq N^+(v)$. Moreover we take $x \in X$, and for every $y \in X \setminus \{x\}$ we check if either $N^+(y) \setminus X = N^+(x) \setminus X$ or $N^+(y) \setminus X = (V \setminus X) \setminus N^+(x)$. □

We adapt the procedure of Lemma 13 to digraphs. The algorithm computing a regular sesquimodule on digraphs works in two steps. It takes $C \in \mathcal{P} \setminus \{A, B\}$, and try in a first step to find a minimal sesquimodule (w.r.t. inclusion) containing A and C, and not B. If it fails, it tries to find a maximal sesquimodule which contains A, and not B and C. We say that (s, t) with $s, t \in V$ is a *violation for* (x, y) (with $x, y \in V$) if $\{x, s\}$ is not a sesquimodule in $G[\{x, y, s, t\}]$.

Lemma 16 *Let X be a vertex subset, $x \in X$ and $y \notin X$. Then X is a sesquimodule if and only if there is no $s \in X$ and $t \notin X$ such that (s, t) is a violation for (x, y).*

Proof: If (s, t) is a violation then trivially X is not a sesquimodule. Suppose now that X is not a sesquimodule. Either there are $t \notin X$ and $s, s' \in X$ such that $\overline{t|ss'}$, or there are $s, s' \in X$ such that $\mathrm{Part}(s) \setminus X \neq \mathrm{Part}(s') \setminus X$. In the first case we have either $\overline{t|xs}$ or $\overline{t|xs'}$, thus either (s,t) or (s',t) is a violation for (x,y). (Note that we can have $t = y$.) In the second case, either $\text{Part}(x) \setminus X \neq \text{Part}(s) \setminus X$ or $\text{Part}(x) \setminus X \neq \text{Part}(s') \setminus X$. W.l.o.g. we suppose that $\text{Part}(x) \setminus X \neq \text{Part}(s) \setminus X$. Note that $\text{Part}(x) \setminus X$ and $\text{Part}(s) \setminus X$ are two partitions of $\overline{X}$ of size at most two. Thus either $\text{Part}(x) \setminus X$ or $\text{Part}(s) \setminus X$ is of size two. If $\text{Part}(s) \setminus X$ has size one, then (s,t) is a violation, where $t \in W$ and $W \in \text{Part}(s) \setminus X$ with $y \notin W$. If $\text{Part}(x) \setminus X$ has size one, we conclude similarly. Thus $\text{Part}(x) \setminus X$ and $\text{Part}(s) \setminus X$ are two partitions of size two. Let $W \in \text{Part}(x) \setminus X$ and $W' \in \text{Part}(s) \setminus X$ such that $y \in W \cap W'$. $W \Delta W'$ is non empty since the two partitions differ. Then (s,t) is a violation, where $t \in W \Delta W'$. □

Lemma 17 *Let $\mathcal{P}$ be a partition of V and let $A, B \in \mathcal{P}$. One can compute in $O(n^2)$ time a minimal (w.r.t inclusion) sesquimodule S of G such that $A \subseteq S$, $B \subsetneq \overline{S}$ and which does not separate any set in $\mathcal{P}$ (if such a sesquimodule exists).*

Proof: Let $x \in A$ and $y \in B$. We start with $X = A$. As long as there is a violation (s,t) for (x,y) such that $s \in X$ and $t \in V \setminus X$, we add the component of $\mathcal{P}$ containing t into X. When X cannot be augmented anymore, X is a sesquimodule by Lemma 16, and by construction X is minimal. This can be done in $O(n^2)$, since there is at most a quadratic number of couples to test, and each violation test take constant time. □

The proof of the following lemma is similar to the proof of Lemma 17, and is omitted.

Lemma 18 *Let $\mathcal{P}$ be a partition of V and let $A, B \in \mathcal{P}$. One can compute in $O(n^2)$ time a maximal (w.r.t inclusion) sesquimodule S of G such that $A \subsetneq S$, $B \subseteq \overline{S}$ and which does not separate any set in $\mathcal{P}$ (if such a sesquimodule exists).*

Lemma 19 *Let $\mathcal{P}$ be a partition of V and let $A, B \in \mathcal{P}$. One can compute in $O(n^2)$ time a sesquimodule S of G such that $A \subsetneq S$, $B \subsetneq \overline{S}$, and which does not separate any set in $\mathcal{P}$ (if such a sesquimodule exists).*

Proof: Let $C \in \mathcal{P} \setminus \{A, B\}$. If there is a sesquimodule S with $A \subsetneq S$ and $B \subsetneq \overline{S}$, then either $C \subseteq S$ or $C \subseteq \overline{S}$. In the first case, Lemma 17 says that one can compute a sesquimodule S with $A \cup C \subseteq S$, $B \subsetneq \overline{S}$ in $O(n^2)$time. Similarly by Lemma 18, in the second case one can compute sesquimodule with $A \subsetneq S$, $B \cup C \subseteq \overline{S}$ in the same time. □

The procedure of Lemma 14 remains unchanged. Thus we get the following.

Corollary 3 *One can compute in $O(n^3)$ time a family $\mathcal{S}$ of sesquimodules of a digraph G such that: for every non cross-free sesquimodule S then either $S \in \mathcal{S}$ or there is a $S' \in \mathcal{S}$ such that S' which crosses S. In particular, $\mathcal{S}$ contains every cross-free sesquimodule.*

Let $\mathcal{S}$ be the family of the Corollary 3 and let $T_{\mathcal{S}}$ be its representative tree. We show now how to find non cross-free members of $\mathcal{S}$. Let T be the sesquimodular decomposition tree of G. Let α be a node in $T_{\mathcal{S}}$, let k be its degree, and $\{V_1, \dots, V_k\}$

be the partition of V induced by α. We know that for every i, V_i or $\overline{V_i}$ is a sesquimodule. Moreover by Lemma 14, for every $I \subsetneq \{1,\ldots,k\}$ such that $1 < |I| < k$, neither $W = \cup_{i\in I} V_i$ nor $\overline{W}$ is a sesquimodule. Thus we know that if $k > 3$, the node α is prime. Accordingly, α is marked prime for later use. Let $\{\alpha, \beta\}$ be an internal edge of $T_{\mathcal{S}}$ corresponding to a partition $\{A, B\}$ of V. Suppose w.l.o.g. that A is a sesquimodule. If α or β is prime, then A is cross-free since a sesquimodule X which crosses A will be a regular member of the quotient family corresponding to α and β. Thus α and β have degree 3, and separate V into 4 sets $\{A_1, A_2, B_1, B_2\}$.

Lemma 20 *If a sesquimodule X crosses A, then there is a $(i,j) \in \{1,2\} \times \{1,2\}$ such that $A_i \cup B_j$ is a sesquimodule.*

Proof: If X crosses A_1 and A_1 is a sesquimodule, then $X \setminus A_1$ and $X \cup A_1$ are both sesquimodules. If X crosses A_1 and $\overline{A_1}$ is a sesquimodule, then $\overline{A_1} \cap X = X \setminus A_1$ and $\overline{A} \setminus X = \overline{X \cup A_1}$ are both sesquimodules. In all cases, there is a sesquimodule which crosses A and not A_1. With the same argument on A_2, B_1 and B_2, there is a sesquimodule which crosses A and does not cross A_1, A_2, B_1 or B_2. □

By Lemma 20, one can check in $O(m)$ time if A is cross-free since there is at most 4 sets to check. The family of cross-free sesquimodules and the sesquimodular decomposition tree T of G can be computed from $\mathcal{S}$ in time $O(nm)$.

The only remaining thing is to get the type of the internal nodes. Let α be a node in a sesquimodular decomposition tree, and let $\{V_1,\ldots,V_k\}$ be the partition of V induced by α. Let $\mathcal{F}$ be the quotient family corresponding to α. Each vertex in $G_{\mathcal{F}}$ corresponds to a set in $\{V_1,\ldots,V_k\}$. For a subset $I \subseteq \{1,\ldots k\}$, let $W_I = \cup_{i\in I} V_i$. If α has already been marked prime in the previous step, then it is prime. If α has degree at most 4, then we check for every $I \subseteq \{1,\ldots k\}$ if W_I is a sesquimodule, thus we can deduce the type of α. Otherwise α is either complete or linear. Clearly, α is of type K_k if for every $I \subseteq \{1,2,3,4\}$, W_I and $\overline{W_I}$ are sesquimodules, and α is of type $K_{k-1}+v$ if for every $I \subseteq \{1,2,3,4\}$, exactly one of W_I and $\overline{W_I}$ is a sesquimodule. Otherwise, α is linear. We know that $\mathcal{S}$ contains exactly $k-3$ non cross-free sesquimodules of the family corresponding to α, since α comes from the contraction of edges of a connected subgraph of $T_{\mathcal{S}}$, and each node of this subgraph has degree 3. Each non cross-free sesquimodule is a W_I where I is consecutive in the ordering of α. Thus we can deduce an ordering $(v_1,\ldots,v_k)$ of α in $O(n^2)$. Now α is $P_{k-1}+v$ if there is a i such that $W_{\{i,i+1\}}$ and $W_{\{i,i-1\}}$ are not sesquimodules, otherwise α is P_k. We have proved that

Theorem 10 *The sesquimodular decomposition tree of a given digraph $G = (V,A)$ can be computed in $O(|V|^3)$ time.*

8 Conclusion and Perspectives

Two new classes of set families, namely weakly partitive crossing families and union-difference families have been studied in the paper. It is shown that both cases can be represented via a unique tree. This result is also applied on two new combinatorial decompositions, both proper generalizations of clan decomposition. Polynomial algorithms are given for computing the corresponding decomposition trees. But of course the runtimes of these algorithms have to be improved for a practical use.

In a digraph $G = (V, A)$, a split-module is a vertex subset $M \subseteq V$ such that $\forall x, y \in M$ we have at the same time that $N^+(x) \setminus M = N^+(y) \setminus M$ and that $(N^-(x) \setminus M \neq \emptyset \wedge N^-(y) \setminus M \neq \emptyset) \Rightarrow (N^-(x) \setminus M = N^-(y) \setminus M)$. The family of split-modules arise in the study of directed splits of G and in fact this family is at the same time a weakly partitive crossing family and a union-difference family, but it is not necessarily a partitive crossing family (hence not symmetric crossing nor weakly partitive) [31]. In this sense, representing set families satisfying a number of closure operations is an important question, and we are convinced that other combinatorial decompositions can be expressed in this framework.

For the family of sesquimodules of a 2-structure we have seen that it is a union-difference family. In fact it is not necessarily a weakly partitive crossing family (the most general family known to have a sub-quadratic space complexity). In this sense the open question whether union-difference families have a sub-quadratic space complexity could be of special interest.

In Section 3 we gave a connection between modular decomposition and submodular functions by analyzing the structures of the underlying set families that can be defined therein. This could turn out to be useful. On the one hand, there are still on-going large research efforts, *e.g.*, [44], for simplifying the impracticality of linear time modular decomposition algorithms. We hope that the connection to the powerful framework of submodular function minimization would help in this direction. On the other hand, the main algorithmic advantage of modular decomposition consists in the tricky techniques developed on the way to obtaining the linear runtime of modular decomposition algorithms. Then, our connection hints applications of these techniques into the study of graph parameters based on symmetric submodular functions such as branch-width and rank-width. In this sense, techniques as partition refinements (a.k.a. vertex splitting) have recently contributed in preliminary steps for fast FPT algorithms parameterized by clique-width [8], rank-width [9], and the recently introduced boolean-width [7]. However, modular decomposition has stronger computational properties, whose application, if possible, would be very interesting.

Acknowledgements: The first author is grateful to S. Thomassé for helpful discussions and pointers.

References

[1] A. Bernáth. *private communication.*

[2] K. Booth and G. Lueker. Testing for the Consecutive Ones Property, Interval Graphs, and Graph Planarity Using PQ-Tree Algorithms. *Journal of Computer and System Sciences*, 13(3):335–379, 1976.

[3] B.-M. Bui-Xuan. *Tree-representation of set families in graph decompositions and efficient algorithms.* PhD thesis, Université Montpellier II, 2008.

[4] B.-M. Bui-Xuan and M. Habib. A Representation Theorem for Union-Difference Families and Application. In *8th Latin American Theoretical Informatics (LATIN'08)*, volume 4957 of *LNCS*, pages 492–503, 2008.

[5] B.-M. Bui-Xuan and M. Habib. Unifying the representation of symmetric crossing families and weakly partitive families. In *European Conference on Combinatorics, Graph Theory and Applications (EuroComb'09)*, volume 34 of *ENDM*, pages 329–333, 2009.

[6] B.-M. Bui-Xuan, M. Habib, and M. Rao. Representation theorems for two set families and applications to combinatorial decompositions. In *International Conference on Relations, Orders and Graphs: Interaction with Computer Science (ROGICS'08)*, pages 532–546. Nouha editions, 2008.

[7] B.-M. Bui-Xuan, J. A. Telle, and M. Vatshelle. Boolean-width of graphs. In *4th International Workshop on Parameterized and Exact Computation (IWPEC'09)*, volume 5917 of *LNCS*, pages 61–74, 2009.

[8] B.-M. Bui-Xuan, J. A. Telle, and M. Vatshelle. Feedback vertex set on graphs of low cliquewidth. In *20th International Workshop on Combinatorial Algorithms (IWOCA'09)*, volume 5874 of *LNCS*, pages 113–124, 2009.

[9] B.-M. Bui-Xuan, J. A. Telle, and M. Vatshelle. H-join decomposable graphs and algorithms with runtime single exponential in rankwidth. *Discrete Applied Mathematics*, 158(7):809–819, 2010.

[10] M. Chein, M. Habib, and M.C. Maurer. Partitive hypergraphs. *Discrete Mathematics*, 37(1):35–50, 1981.

[11] D. A. Cohen, M. C. Cooper, and P. G. Jeavons. Generalising submodularity and Horn clauses: Tractable optimization problems defined by tournament pair multimorphisms. Technical Report CS-RR-06-06, Oxford University, 2006.

[12] W. Cunningham. *A combinatorial decomposition theory.* PhD thesis, University of Waterloo, 1973.

[13] W. Cunningham. Decomposition of directed graphs. *SIAM Journal on Algebraic and Discrete Methods*, 3:214–228, 1982.

[14] W. Cunningham. Decomposition of submodular functions. *Combinatorica*, 3(1):53–68, 1983.

[15] W. Cunningham and J. Edmonds. A combinatorial decomposition theory. *Canadian Journal of Mathematics*, 32:734–765, 1980.

[16] E. Dinitz, A. Karzanov, and M. Lomonosov. On the structure of a family of minimal weighted cuts in a graph. In *A. Pridman (Ed.), Studies in Discrete Optimization, Nauka, Moscow*, pages 290–306, 1976.

[17] J. Edmonds and R. Giles. A min-max relation for submodular functions on graphs. *Annals of Discrete Mathematics*, 1:185–204, 1977.

[18] A. Ehrenfeucht, T. Harju, and G. Rozenberg. *The Theory of 2-Structures- A Framework for Decomposition and Transformation of Graphs*. World Scientific, 1999.

[19] A. Ehrenfeucht and G. Rozenberg. Theory of 2-structures. *Theoretical Computer Science*, 3(70):277–342, 1990.

[20] M. Everett and S. Borgatti. Regular Equivalence: General Theory. *Journal of Mathematical Sociology*, 18:29–52, 1994.

[21] J. Fiala and D. Paulusma. A complete complexity classification of the role assignment problem. *Theoretical Computer Science*, 349(1):67–81, 2005.

[22] J.-L. Fouquet, M. Habib, F. de Montgolfier, and J.-M. Vanherpe. Bimodular decomposition of bipartite graphs. In *30th International Workshop on Graph-Theoretic Concepts in Computer Science (WG'04)*, volume 3353 of *LNCS*, pages 117–128, 2004.

[23] S. Fujishige. Canonical decompositions of symmetric submodular systems. In *Graph Theory and Algorithms*, volume 108 of *LNCS*, pages 53–64, 1981.

[24] C. Gabor, K. Supowit, and W.-L. Hsu. Recognizing circle graphs in polynomial time. *Journal of the ACM*, 36(3):435–473, 1989.

[25] H. Gabow. Centroids, Representations, and Submodular Flows. *Journal of Algorithms*, 18(3):586–628, 1995.

[26] J. Gagneur, R. Krause, T. Bouwmeester, and G. Casari. Modular decomposition of protein-protein interaction networks. *Genome Biology*, 5(8), 2004.

[27] Tibor Gallai. Transitiv orientierbare Graphen. *Acta Mathematica Academiae Scientiarum Hungaricae*, 18:25–66, 1967.

[28] J. Geelen, A. Gerards, and G. Whittle. Branch-width and well-quasi-ordering in matroids and graphs. *Journal of Combinatorial Theory, Series B*, 84(2):270–290, 2002.

[29] J. Goldman and G.-C. Rota. The number of subspaces of a vector space. *Recent Progress in Combinatorics*, pages 75–83, 1969.

[30] M. Habib and C. Paul. A survey on the algorithmic aspects of modular decomposition. *Computer Science Review*, 4(1):41–59, 2010.

[31] P. Heggernes, R. McConnell, D. Meister, and A. Proskurowski. *private communication.*

[32] W.-L. Hsu and R. M. McConnell. PC-trees and circular-ones arrangements. *Theoretical Computer Science*, 296:99–116, 2003.

[33] J.-M. Lanlignel. *Autour de la décomposition en coupes.* PhD thesis, Université Montpellier II, 2001.

[34] R. McConnell and F. de Montgolfier. Algebraic operations on PQ trees and modular decomposition trees. In *31st International Workshop on Graph-Theoretic Concepts in Computer Science (WG'05)*, volume 3787 of *LNCS*, pages 421–432, 2005.

[35] R.H. Möhring and F.J. Radermacher. Substitution decomposition for discrete structures and connections with combinatorial optimization. *Annals of Discrete Mathematics*, 19:257–356, 1984.

[36] F. de Mongolfier and M. Rao. The bi-join decomposition. In *7th International Colloquium on Graph Theory (ICGT'05)*, 2005.

[37] F. de Montgolfier. *Décomposition modulaire des graphes. Théorie, extensions et algorithmes.* PhD thesis, Université Montpellier II, 2003.

[38] S. Oum and P. Seymour. Approximating clique-width and branch-width. *Journal of Combinatorial Theory, Series B*, 96(4):514–528, 2006.

[39] M. Queyranne. Minimizing symmetric submodular functions. *Mathematical Programming*, 82(1–2):3–12, 1998.

[40] N. Robertson and P. Seymour. Graph minors X. Obstructions to tree-decomposition. *Journal of Combinatorial Theory, Series B*, 52(2):153–190, 1991.

[41] A. Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency.* Springer-Verlag, 2003.

[42] C. Semple and M. Steel. *Phylogenetics.* Oxford Lecture Series in Mathematics and Its Applications, number 24, 2003.

[43] W.-K. Shih and W.-L. Hsu. A new planarity test. *Theoretical Computer Science*, 223(1–2):179–191, 1999.

[44] M. Tedder, D. Corneil, M. Habib, and C. Paul. Simple, Linear-Time Modular Decomposition. In *35th International Colloquium on Automata, Languages and Programming (ICALP'08)*, volume 5125 of *LNCS*, pages 634–645, 2008.

[45] D. White and K. Reitz. Graph and Semigroup Homomorphisms on Networks of Relations. *Social Networks*, 5:193–234, 1983.

[46] A. Zykov. On some properties of linear complexes. *Siberian Mathematical Journal*, 24:163–188, 1949.