Branch-Width, Parse Trees, and Monadic Second-Order Logic for Matroids*

Petr Hliněný

School of Mathematical and Computing Sciences,
Victoria University of Wellington, New Zealand
hlineny@member.ams.org

Current and mailing address:
Petr Hliněný
Department of Computer Science
VŠB – Technical University of Ostrava
17. listopadu 15, 708 33 Ostrava, Czech Republic
petr.hlineny@vsb.cz

July 7, 2005

Abstract. We introduce "matroid parse trees" which, using only a limited amount of information at each node, can build up the vector representations of matroids of bounded branch-width over a finite field. We prove that if $\mathcal{M}$ is a family of matroids described by a sentence in the monadic second-order logic of matroids, then there is a finite tree automaton accepting exactly those parse trees which build vector representations of the bounded-branch-width representable members of $\mathcal{M}$. Since the cycle matroids of graphs are representable over any field, our result directly extends the so called "MS_2-theorem" for graphs of bounded tree-width by Courcelle, and others. Moreover, applications and relations in areas other than matroid theory can be found, like for rank-width of graphs, or in the coding theory.

Keywords: matroid representation, branch-width, monadic second-order logic, tree automaton, fixed-parameter complexity.

2000 Math Subjects Classification: primary 05B35, secondary 68R05, 03D05.

* An extended abstract published in: STACS 2003, Lecture Notes in Computer Science 2607, Springer Verlag, 2003, 319–330.

1 Introduction

We assume that the reader is familiar with basic concepts of graph theory. In the past decade, the notion of a *tree-width* of graphs [24, 3] attracted plenty of attention, both from graph-theoretical and computational points of view. This attention followed the pioneer work of Robertson and Seymour on the Graph Minor Project [23], and results of various researchers using tree-width in designing algorithms and in the theory of parametrized complexity.

The notion of a *branch-width* is closely related to that of a tree-width [24], but a branch-decomposition does not refer to vertices and so branch-width directly generalizes from graphs to matroids. (We note a new result [21] showing that it is possible to define tree-decompositions without referring to graph vertices, and so also to extend the definition of tree-width to matroids.) We postpone formal definitions till next sections.

Branch-width has recently shown to be a very interesting structural matroid parameter, too. Besides other results, we mention the following interesting advances; well-quasi-ordering of matroids of bounded branch-width over finite fields [12], a size-bound on the excluded minors for matroids of fixed branch-width [13], or a so called "excluded-grid" theorem for matroids representable over finite fields [14]. Somehow surprisingly, binary matroids and their branch-width also have close relation with a clique-width of graphs, via the notion of a rank-width [25, 10].

We show in this and related papers that branch-width and branch-decompositions of representable matroids have interesting complexity-theoretical aspects. Namely we prove here a result analogous to so called "MS_2-theorem" by Courcelle [6] (see also in [2] or [5]) for matroids represented by matrices over a finite field $\mathbb{F}$ (Theorem 6.1): If $\mathcal{M}$ is a family of matroids described by a sentence in the monadic second-order logic, then the "parse trees" of bounded-branch-width $\mathbb{F}$-represented members of $\mathcal{M}$ are recognizable by a finite tree automaton. This result covers, among other applications, the cycle matroids of graphs.

Our proof follows the main ideas of Abrahamson–Fellows' [1] combinatorial approach to Courcelle's theorem; but, by using branch-width instead of tree-width, we manage to avoid some technical difficulties of their proof even in our more general setting. In the language of parametrized complexity [11], we prove that matroid properties expressible in MSO logic are fixed-parameter tractable for $\mathbb{F}$-represented matroids of bounded branch-width. See brief overviews in Sections 5,7.

We formulate our results in the language of matroid theory since it is natural and convenient, and since it shows the close relations of this research to well-known graph structural and computational concepts. Our work could be, as well, viewed as a result about matrices, point configurations, or about linear codes over a finite field. The key to these results is the notion of parse trees for bounded-width $\mathbb{F}$-represented matroids, defined in Section 3. We propose the parse trees as a powerful tool for handling matroids of bounded branch-width in general.

Our research involves, besides structural matroid theory, also areas of logic and of theoretical computer science. In order to make the paper accessible to a wide audience of combinatorists and computer scientists, we provide sufficient introductory definitions for all of these areas.

This paper is structured as follows: Section 2 defines matroids, their connectivity and branch-decompositions, and specifically matroids represented over a (finite) field. Parse trees for represented matroids of bounded branch-width are introduced and discussed in Section 3. Section 4 presents an analogue of the classical Myhill-Nerode theorem for our parse trees. The MSO theories of graphs and matroids are defined and discussed in Section 5. The previous concepts are then used to prove our main Theorem 6.1 in Section 6. Algorithmic and logic aspects or consequences of our results are reviewed in Section 7. Finally, Section 8 presents some supplementary results, and discusses relations and limits of our research.

2 Basics of Matroids

We refer to Oxley [22] for our matroid terminology. A *matroid* is a pair $M = (E, \mathcal{B})$ where $E = E(M)$ is the ground set of M (elements of M), and $\mathcal{B} \subseteq 2^E$ is a nonempty collection of *bases* of M. Matroid bases satisfy the "exchange axiom"; if $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 - B_2$, then there is $y \in B_2 - B_1$ such that $(B_1 - \{x\}) \cup \{y\} \in \mathcal{B}$. In particular, no two bases are in an inclusion, and all have the same cardinality – the *rank* of M. Subsets of bases are called *independent sets*, and the remaining sets are *dependent*. The *rank function* $\mathrm{r}_M(X)$ in M is the maximal cardinality of an independent subset of a set $X \subseteq E(M)$.

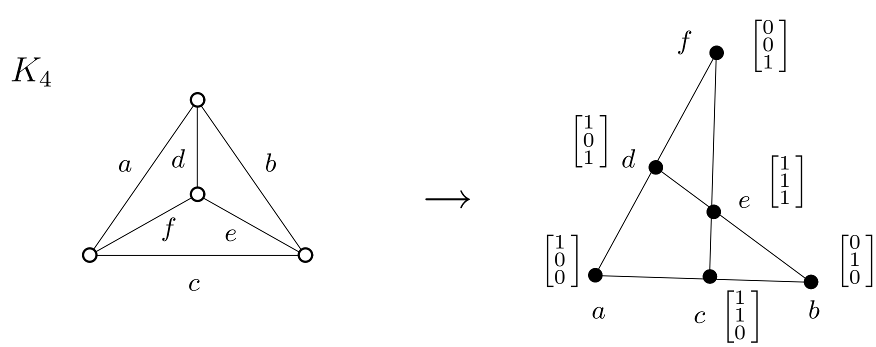

Fig. 1. An example of a vector representation of the cycle matroid $M(K_4)$. The matroid elements are depicted by dots, and their (linear) dependency is shown using lines.

If G is a graph, then its *cycle matroid* on the ground set $E(G)$ is denoted by $M(G)$. The bases of $M(G)$ are the spanning forests of G, and the minimal dependent sets – *circuits* of $M(G)$, are the cycles of G. Then the rank of a subset of edges $F \subseteq E(G)$ equals the number of vertices minus the number of components induced by F. In fact, a lot of matroid terminology is inherited from graphs. Another typical example of a matroid is a finite set of vectors with usual linear dependency. To illustrate these two basic examples, we show in Fig. 1 the cycle matroid of the complete graph K_4, together with its vector representation.

The *dual* matroid M^* of M is defined on the same ground set E, and the bases of M^* are the set-complements of the bases of M. An element e of M is called a *loop* (a *coloop*), if $\{e\}$ is dependent in M (in M^*). The matroid $M \setminus e$ obtained by *deleting* a non-coloop element e is defined as $(E - \{e\}, \mathcal{B}^-)$ where $\mathcal{B}^- = \{B : B \in \mathcal{B}, e \notin B\}$. The matroid M/e obtained by *contracting* a non-loop element e is defined using duality $M/e = (M^* \setminus e)^*$. (This corresponds to contracting an edge in a graph.) Contracting a loop means deleting it, and analogously for coloops.

A *minor* of a matroid is obtained by a sequence of deletions and contractions of elements. A matroid family $\mathcal{M}$ is *minor-closed* if $M \in \mathcal{M}$ implies that all minors of M are in $\mathcal{M}$. A matroid N is called an *excluded* minor for a minor-closed family $\mathcal{M}$ if N is a minimal matroid not in $\mathcal{M}$ in the minor order.

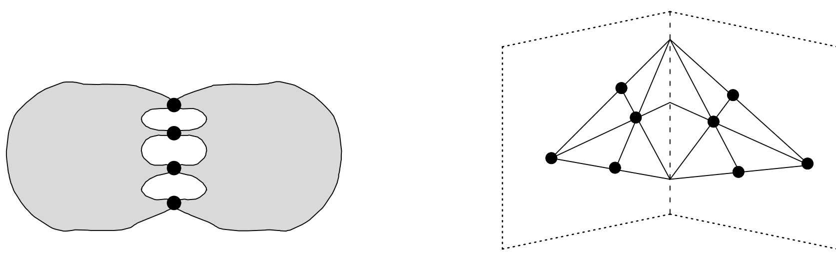

Fig. 2. An illustration to a 4-separation in a graph, and to a 3-separation in a matroid.

Another important concept is matroid connectivity, which is close, but somehow different, to traditional graph connectivity. The *connectivity function* λ_M of a matroid M is defined for all subsets $A \subseteq E$ by

$$\lambda_M(A) = \mathrm{r}_M(A) + \mathrm{r}_M(E - A) - \mathrm{r}(M) + 1\,.$$

Here $\mathrm{r}(M) = \mathrm{r}_M(E)$. Notice that the function is symmetric $\lambda_M(A) = \lambda_M(E - A)$. A subset $A \subseteq E$ is *k-separating* if $\lambda_M(A) \le k$. A partition $(A, E - A)$ is called a *k-separation* if A is k-separating and both $|A|, |E - A| \ge k$. Geometrically, the spans of the two sides of a k-separation intersect in a subspace of rank less than k. See in Fig. 2 and Lemma 2.1. For $n > 1$, a matroid M is called *n-connected* if it has no k-separation for $k = 1, 2, \ldots, n-1$, and $|E(M)| \ge 2n - 2$. We say that a matroid is *connected* if it is 2-connected.

In a corresponding graph view, the connectivity function $\lambda_G(F)$ of an edge subset $F \subseteq E(G)$ equals the number of vertices of G incident both with F and with $E(G) - F$. (Then $\lambda_G(F) = \lambda_{M(G)}(F)$ provided both sides of the separation are connected in G.) That is close to the traditional view of vertex cuts in a graph, but there are small technical differences: For example, a triangle forms a 3-separation in a graph, though there is no proper vertex cut involved in.

The advantage of the above presented view of connectivity (so called Tutte connectivity) lies in the fact that it is preserved under duality. Notice, as an exercise, that a matroid is 2-connected by the above definition if and only if every two elements belong to a common circuit.

2.1 Branch-Decomposition

A *sub-cubic tree* is a tree in which all vertices have degree at most three. (We do not use the word ternary because such trees are actually sub-binary in the sense of the next section.) Let $\ell(T)$ denote the set of leaves of a tree T.

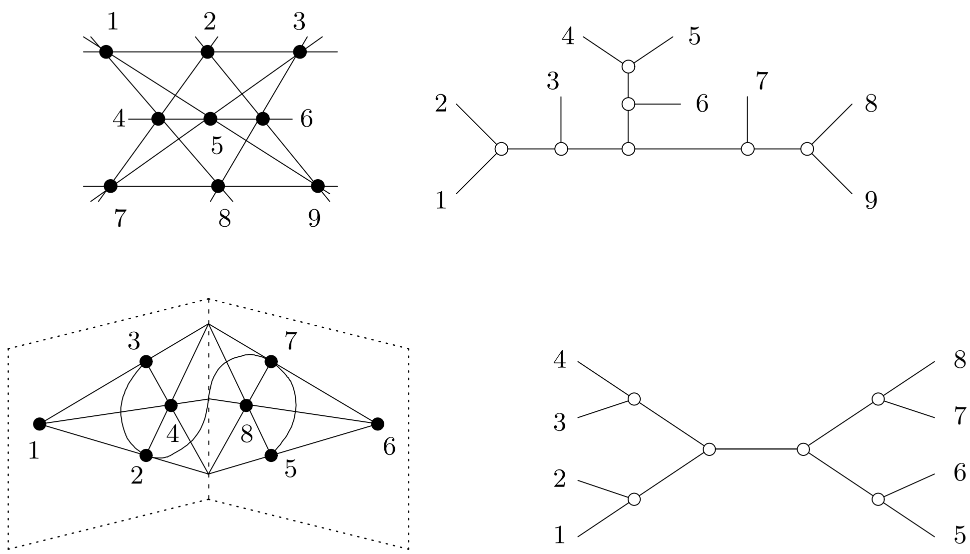

Fig. 3. Two examples of width-3 branch-decompositions of the Pappus matroid (top left, in rank 3) and of the binary affine cube (bottom left, in rank 4). The lines in matroid pictures show dependencies among elements.

Let M be a matroid on the ground set $E = E(M)$. A *branch-decomposition* of M is a pair (T, τ) where T is a sub-cubic tree, and τ is an injection of E into $\ell(T)$, called *labeling*. Let e be an edge of T, and T_1, T_2 be the connected components of $T - e$. We say the e *displays* the partition (A, B) of E where $A = \tau^{-1}(\ell(T_1))$, $B = \tau^{-1}(\ell(T_2))$. The *width* of an edge e in T is $\omega_T(e) = \lambda_M(A) = \lambda_M(B)$. The width of the branch-decomposition (T, τ) is maximum of the widths of all edges of T, and the *branch-width* of M is the minimal width over all branch-decompositions of M. If T has no edge, then we take its width as 0.

Some examples of branch-decompositions are presented in Fig. 3. Branch-width of matroids is preserved under duality by definition. We remark that the branch-width of a graph is defined analogously, using the above connectivity function λ_G. Clearly, $\lambda_{M(G)}(F) \leq \lambda_G(F)$ in a connected graph, but these numbers may not be equal if the subgraph induced by F or by $E(G) - F$ is not connected. It is still an open conjecture that the branch-width of a graph G equals the branch-width of its cycle matroid $M(G)$.

2.2 Represented Matroids

We now turn our attention to matroids represented over a fixed field $\mathbb{F}$. This is a crucial part of our introductory definitions. A *representation* of a matroid M is a matrix $\boldsymbol{A}$ whose columns correspond to the elements of M, and maximal linearly independent subsets of columns form the bases of M. We denote by $M(\boldsymbol{A})$ the matroid represented by a matrix $\boldsymbol{A}$.

We denote by $PG(n, \mathbb{F})$ the *projective geometry (space)* obtained from the vector space $\mathbb{F}^{n+1}$. See [22, Section 6.1,6.3] for an overview of projective spaces and of (in)equivalence of matroid representations. For a set $X \subseteq PG(n, \mathbb{F})$, we denote by $\langle X \rangle$ the span (affine closure) of X in the space. The rank $\mathrm{r}(X)$ of X is the maximal cardinality of a linearly independent subset of X. (This definition coincides with matroid rank.) A projective transformation is a mapping between two projective spaces over $\mathbb{F}$ that is induced by a linear transformation between the underlying vector spaces. Clearly, the matroid $M(\boldsymbol{A})$ represented by a matrix $\boldsymbol{A}$ is unchanged when columns are scaled by non-zero elements of $\mathbb{F}$. Hence we may view a loopless matroid representation $M(\boldsymbol{A})$ as a multiset of points in the projective space $PG(n, \mathbb{F})$ where n is the rank of $M(\boldsymbol{A})$.

We call a finite multiset of points in a projective space over $\mathbb{F}$ a *point configuration*; and we represent a loop in a point configuration by the empty subspace $\emptyset$. Two point configurations P_1, P_2 in projective spaces over $\mathbb{F}$ are *projectively equivalent* if there is a non-singular projective transformation between the projective spaces that maps P_1 onto P_2 bijectively. (Loops are mapped only to loops.) One may think that we do not have to include the word "bijectively" in the description since non-singular projective transformations are always injective on the points, but, in fact, we have to do this to handle multiple-elements in multisets. Two (labeled) point configurations over $\mathbb{F}$ are projectively equivalent in our sense if and only if, in the language of [22, Chapter 6], the matrix representations are equivalent without use of $\mathbb{F}$-automorphisms, otherwise called strongly equivalent in matroid theory.

We define an *$\mathbb{F}$-represented matroid* to be a projective equivalence class of point configurations over $\mathbb{F}$. Obviously, all point configurations in one equivalence class belong to the same isomorphism class of matroids. When we want to deal with an $\mathbb{F}$-represented matroid, we actually pick an arbitrary point configuration from the equivalence class. Standard matroidal terms are inherited from matroids to represented matroids. To clearly distinguish between a matroid and a represented matroid, we sometimes add the adjective *abstract* to the former one. We do not label points in our configurations, and so we are dealing with unlabeled matroid elements, which is in correspondence with the MS_M theory defined in Section 6.

Since several inequivalent representations over a fixed field may exist for the same matroid, one abstract matroid may have more than one distinct $\mathbb{F}$-represented matroids. As a clear example, we present in Fig. 4 two point config-

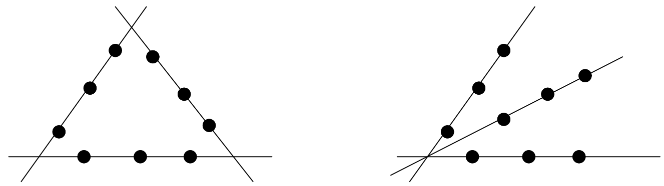

Fig. 4. Two non-equivalent representations of a 9-element rank-3 matroid.

urations representing the same 9-element rank-3 matroid which are not "equivalent" in any reasonable geometric meaning of equivalence. Actually, one should imagine the relation between an abstract matroid and a represented matroid as the relation between an abstract graph and a graph embedded on a surface — both an embedded graph and a represented matroid carry some additional geometric information over the abstract ones.

For the sake of completeness, we now show a basic geometric property of the connectivity function in matroid representations.

Lemma 2.1. *Let M be an $\mathbb{F}$-represented matroid, and $F \subseteq E = E(M)$. Then $\lambda_M(F) = \mathrm{r}\left(\langle F\rangle \cap \langle E-F\rangle\right) + 1$.*

Proof. We use modularity of the rank function in a projective geometry:

$$\lambda_M(F) = \mathrm{r}_M(F) + \mathrm{r}_M(E-F) - \mathrm{r}(M) + 1 =$$
$$= \mathrm{r}(F) + \mathrm{r}(E-F) - \mathrm{r}\left(\langle F\rangle \cup \langle E-F\rangle\right) + 1 = \mathrm{r}\left(\langle F\rangle \cap \langle E-F\rangle\right) + 1$$

∎

3 Parse Trees for Matroids

In this section we introduce our basic formal tool — the parse trees for represented matroids of bounded branch-width. We use this tool to link the matroids with formal languages and automata. Loosely speaking, a parse tree shows how to "build up" the matroid along the tree using only fixed amount of information at each tree node, and so it forms a suitable background for dynamic programming.

We are inspired by analogous boundaried graphs and parse trees known for handling graphs of bounded tree-width (see for example [1] or [11, Section 6.4]): A boundaried graph is a graph with a distinguished subset of labeled vertices. (The purpose is that only the boundary vertices can be "accessed from outside".) Then, simply speaking, a graph has tree-width at most $t-1$ if and only if it can be composed from small pieces by gluing them on boundaries of size at most t. We similarly define boundaried represented matroids, in which the boundary is a distinguished subspace of the representation, and composition operators that are used to glue representations together. However, matroids are more difficult to handle than graphs, and they bring some new problems not appearing in graphs, mainly with (in)equivalence of representations.

3.1 Tree Automata

For our arguments we need a slightly extended type of a usual automaton, that reads "words" given as labeled trees instead of sequences. Such a "tree automaton" starts its processing in the leaves of the given tree, and it moves towards the tree root. Formal definitions are next.

A *rooted ordered sub-binary tree* is such that each of its vertices has at most two sons that are ordered as "left" and "right". (If there is only one son, then it may be either left or right.) Let Σ be a finite alphabet. We denote by Σ^{**} the class of rooted ordered sub-binary trees with vertices labeled by symbols from Σ. When defining tree automata, we follow [11, Section 6.1], but we restrict our attention only to sub-binary trees.

A deterministic finite leaf-to-root *tree automaton* is $\mathcal{A} = (K, \Sigma, \delta_t, q_0, F)$, where a set of states K, an alphabet Σ, an initial state q_0, and accepting states F are like in a classical automaton. The transition function δ_t is defined as a mapping from $K \times K \times \Sigma$ to K. Let the function $t\text{-}eval_{\mathcal{A}}$ for $\mathcal{A}$ be defined recursively by $t\text{-}eval_{\mathcal{A}}(\emptyset) = q_0$, and $t\text{-}eval_{\mathcal{A}}(T) = \delta_t(t\text{-}eval_{\mathcal{A}}(T_l), t\text{-}eval_{\mathcal{A}}(T_r), a)$ for $T \in \Sigma^{**}$, where T_l, T_r is the left and right, respectively, subtree of the root of T, and where a is the root symbol. A tree T is accepted by $\mathcal{A}$ if $t\text{-}eval_{\mathcal{A}}(T) \in F$. A tree language $\mathcal{L} \subseteq \Sigma^{**}$ is *finite state* if it is accepted by a finite tree automaton.

3.2 Boundaried Matroids

All matroids throughout this section are $\mathbb{F}$-represented for some fixed field $\mathbb{F}$. It is necessary to consider represented matroids here since the notion of a "k-sum" is not well defined for abstract matroids if $k > 2$. Hence, for simplicity, if we say a "(represented) matroid", then we mean an $\mathbb{F}$-represented matroid. If we speak about a projective space, we mean a projective geometry over the field $\mathbb{F}$, including the empty subspace $\emptyset$ for representing loops. Let $[s, t]$ denote the set $\{s, s+1, \ldots, t\}$.

The following definition presents a possible way of formalizing the notion of a "matroid with a boundary". (Since matroids have no vertices unlike graphs, we have to introduce some special elements that define the matroid boundary.)

Definition. A pair $\bar{N} = (N, \delta)$ is called a *t-boundaried (represented) matroid* if the following holds: $t \geq 0$ is an integer, N is a represented matroid, and $\delta : [1, t] \to E(N)$ is an injective mapping such that $\delta([1, t])$ is an independent set in N.

We call $J(\bar{N}) = E(N) - \delta([1, t])$ the *internal elements* of $\bar{N}$, elements of $\delta([1, t])$ the *boundary points* of $\bar{N}$, and t the *boundary rank* of $\bar{N}$. We denote by $\partial(\bar{N})$ the boundary subspace spanned by $\delta([1, t])$. In particular, the boundary points are not loops. We say that that the boundaried matroid $\bar{N}$ is *based on* the represented matroid $N \setminus \delta([1, t])$, the *internal* matroid of $\bar{N}$, which is the restriction of N to $J(\bar{N})$. The notion of a t-boundaried represented matroid is similar to "rooted configurations" defined in [12], but it is more flexible.

The basic operation we will use is the *boundary sum* $\bar{\oplus}$ of the next definition, illustrated in Fig. 5.

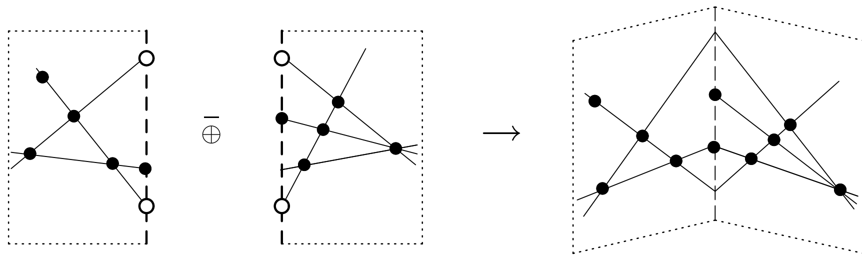

Fig. 5. An example of a boundary sum of two 2-boundaried matroids. The internal elements are drawn as solid dots, the boundary points as hollow dots, and the boundary subspaces of rank 2 are drawn with thick dashed lines. Thin lines show matroid dependencies. The resulting sum is a matroid represented on two intersecting planes in rank 4 (aka "3-dimensional picture" on the right).

Definition. Let $\bar{N}_1 = (N_1, \delta_1)$, $\bar{N}_2 = (N_2, \delta_2)$ be two t-boundaried represented matroids. We denote by $\bar{N}_1 \bar{\oplus} \bar{N}_2 = N$ the represented matroid defined as follows: Let Ψ_1, Ψ_2 be projective spaces such that the intersection $\Psi_1 \cap \Psi_2$ has rank exactly t. Suppose that, for $i = 1, 2$, $P_i \subset \Psi_i$ is a point configuration representing N_i, such that $P_1 \cap P_2 = \delta_1([1, t]) = \delta_2([1, t])$, and $\delta_2(j) = \delta_1(j)$ for $j \in [1, t]$. Then N is the matroid represented by $(P_1 \cup P_2) - \delta_1([1, t])$.

Informally, the boundary sum $\bar{N}_1 \bar{\oplus} \bar{N}_2 = N$ on the ground set $E(N) = J(\bar{N}_1) \dot{\cup} J(\bar{N}_2)$ is obtained by gluing the representations of N_1 and N_2 on a common subspace (the boundary) of rank t, so that the boundary points of both are identified in order and then deleted (see Fig. 5). Keep in mind that a point configuration is considered as a multiset. It is a matter of elementary linear algebra to verify that the boundary sum is well defined with respect to equivalence of point configurations. Clearly, $\bar{\oplus}$ is a commutative operation. It is not difficult to prove the following "kind of associativity" of $\bar{\oplus}$ (see Fig. 6):

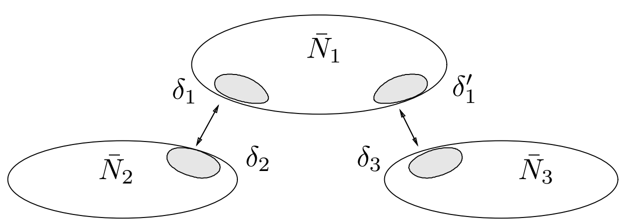

Fig. 6. Illustration to Lemma 3.1.

Lemma 3.1. *Let $\bar{N}_i = (N_i, \delta_i)$, $i = 1, 2, 3$ be t-boundaried represented matroids. Suppose that $\delta_1' : [1, t] \to E(N_1)$ is a mapping such that $\bar{N}_2' = (\bar{N}_1 \bar{\oplus} \bar{N}_2, \delta_1')$ is a t-boundaried matroid. Then $\bar{N}_3' = (N_3', \delta_1)$ where $N_3' = (N_1, \delta_1') \bar{\oplus} \bar{N}_3$ is a t-boundaried matroid as well. Moreover, $\bar{N}_2' \bar{\oplus} \bar{N}_3 = \bar{N}_2 \bar{\oplus} \bar{N}_3'$.* ∎

We write "$\leq t$-boundaried" to mean t'-boundaried for some $0 \leq t' \leq t$. We now define a composition operator (over the field $\mathbb{F}$) which will be used to generate large boundaried matroids from small pieces (Fig. 7).

Definition. A $\leq t$*-boundaried composition* operator is defined as a quadruple $\odot = (R, \gamma_1, \gamma_2, \gamma_3)$, where R is a represented matroid, $\gamma_i : [1, t_i] \to E(R)$ is an injective mapping for $i = 1, 2, 3$ and some fixed $0 \leq t_i \leq t$, each $\gamma_i([1, t_i])$ is an independent set in R, and $\big(\gamma_i([1, t_i]) : i = 1, 2, 3\big)$ is a partition of $E(R)$.
The $\leq t$-boundaried composition operator $\odot$ is a binary operator applied to a t_1-boundaried represented matroid $\bar{N}_1 = (N_1, \delta_1)$ and to a t_2-boundaried represented matroid $\bar{N}_2 = (N_2, \delta_2)$. The result of the composition is a t_3-boundaried represented matroid $\bar{N} = (N, \gamma_3)$, written as $\bar{N} = \bar{N}_1 \odot \bar{N}_2$, where a matroid N is defined using boundaried sums: $N' = \bar{N}_1 \bar{\oplus} (R, \gamma_1)$, $N = (N', \gamma_2) \bar{\oplus} \bar{N}_2$.

Speaking informally, a boundaried composition operator is a bounded-rank configuration with three boundaries distinguished by $\gamma_1, \gamma_2, \gamma_3$, and with no other internal points. The meaning of a composition $\bar{N} = \bar{N}_1 \odot \bar{N}_2$ is that, for $i = 1, 2$, we glue the represented matroid N_i to R, matching $\delta_i([1, t_i])$ with $\gamma_i([1, t_i])$ in order. The result is a t_3-boundaried matroid $\bar{N}$ with boundary $\gamma_3([1, t_3])$. One may shortly write the result as $\bar{N} = \big(\big(\bar{N}_1 \bar{\oplus} (R, \gamma_1), \gamma_2\big) \bar{\oplus} \bar{N}_2, \gamma_3\big)$. Notice that, in general, there are more than one boundaried composition operators with the same ranks. For reference we denote by $t_i(\odot) = t_i$, by $R(\odot) = R$, and by $\gamma_i(\odot) = \gamma_i$. We now show that the meaning of the operands and the resulting boundary in a boundaried composition operator could be easily exchanged.

Lemma 3.2. *Suppose that $\odot = (R, \gamma_1, \gamma_2, \gamma_3)$ and $\odot' = (R, \gamma_1, \gamma_3, \gamma_2)$ are $\leq t$-boundaried composition operators, and that $\bar{N}_i$, $i = 1, 2, 3$ are $t_i(\odot)$-boundaried represented matroids. Then $(\bar{N}_1 \odot \bar{N}_2) \bar{\oplus} \bar{N}_3 = (\bar{N}_1 \odot' \bar{N}_3) \bar{\oplus} \bar{N}_2$.*

Proof. This is a simple formal manipulation with the previous definition using Lemma 3.1:

$$(\bar{N}_1 \odot \bar{N}_2) \bar{\oplus} \bar{N}_3 = \big(\big(\bar{N}_1 \bar{\oplus} (R, \gamma_1), \gamma_2\big) \bar{\oplus} \bar{N}_2, \gamma_3\big) \bar{\oplus} \bar{N}_3 =$$
$$= \big(\big(\bar{N}_1 \bar{\oplus} (R, \gamma_1), \gamma_3\big) \bar{\oplus} \bar{N}_3, \gamma_2\big) \bar{\oplus} \bar{N}_2 = (\bar{N}_1 \odot' \bar{N}_3) \bar{\oplus} \bar{N}_2 \qquad \blacksquare$$

3.3 Parse Trees

Now we present the main outcome of this section — that it is possible to generate all represented matroids of branch-width at most $t + 1$ using $\leq t$-boundaried composition operators on parse trees.

Let $\bar{\Omega}_t$ denote the *empty* t-boundaried matroid (Ω, δ_0) where $t \geq 0$ and $\delta_0([1, t]) = E(\Omega)$ (t will often be implicit in the context). If $\bar{N} = (N, \delta)$ is an arbitrary t-boundaried matroid, then $\bar{N} \bar{\oplus} \bar{\Omega}_t$ is actually the restriction of $\bar{N}$ to $E(N) - \delta([1, t])$. Let $\bar{\Upsilon}$ denote the *single-element* 1-boundaried matroid (Υ, δ_1) where $E(\Upsilon) = \{x, x'\}$ are two parallel elements, and $\delta_1(1) = x'$. Let $\bar{\Upsilon}_0$ denote the *loop* 0-boundaried matroid (Υ_0, δ_0) where $E(\Upsilon_0) = \{z\}$ is a loop, and δ_0

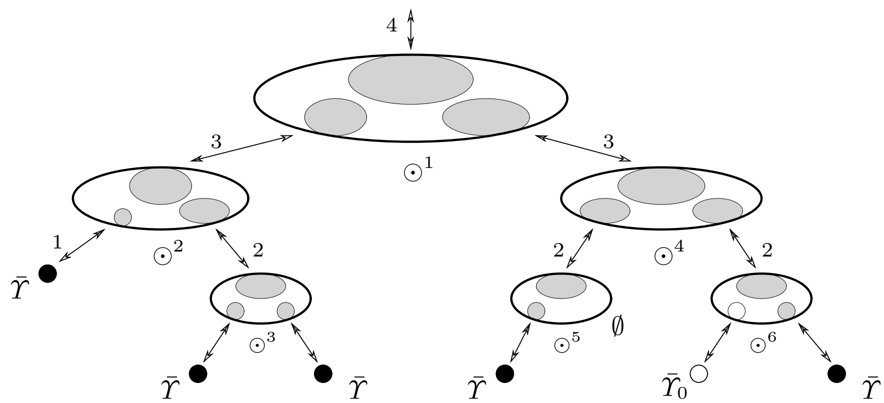

Fig. 7. An example of a boundaried parse tree. The ovals represent composition operators, with shaded parts for the boundaries and edge-numbers for the boundary ranks. (E.g. $\odot^4 = (R^4, \gamma_1^4, \gamma_2^4, \gamma_3^4)$ where $\gamma_1^4, \gamma_2^4 : [1,2] \to E(R^4)$, $\gamma_3^4 : [1,3] \to E(R^4)$.)

is an empty mapping. Let $\mathcal{R}_t^{\mathbb{F}}$ denote the set of all $\leq t$-boundaried composition operators over the field $\mathbb{F}$.

We set $\Pi_t = \mathcal{R}_t^{\mathbb{F}} \cup \{\bar{\Upsilon}, \bar{\Upsilon}_0\}$ where $\mathbb{F}$ is considered fixed. Let $T \in \Pi_t^{**}$ be a rooted ordered sub-binary tree with vertices labeled by the alphabet Π_t. Considering a vertex v of T; we set $\varrho(v) = 1$ if v is labeled by $\bar{\Upsilon}$, $\varrho(v) = 0$ if v is labeled by $\bar{\Upsilon}_0$, and $\varrho(v) = t_3(\odot)$ if v is labeled by $\odot$. We call T a *$\leq t$-boundaried parse tree* if the following are true:

- only leaves of T are labeled by $\bar{\Upsilon}$ or $\bar{\Upsilon}_0$;
- if a vertex v of T labeled by a composition operator $\odot$ has no left (no right) son, then $t_1(\odot) = 0$ $(t_2(\odot) = 0)$;
- if a vertex v of T labeled by $\odot$ has left son u_1 (right son u_2), then $t_1(\odot) = \varrho(u_1)$ $(\, t_2(\odot) = \varrho(u_2)\,)$.

Informally, the boundary ranks of composition operators and/or single-element terminals must "agree" across each edge. Notice that $\bar{\Upsilon}$ or $\bar{\Upsilon}_0$ are the only labels from Π_t that "create" elements of the resulting represented matroid $P(T)$ in the next definition. See an illustration example in Fig. 7.

Definition. Let T be a $\leq t$-boundaried parse tree. The $\leq t$-boundaried represented *matroid* $\bar{P}(T)$ *parsed by* T is recursively defined as follows:

- if T is an empty tree, then $\bar{P}(T) = \bar{\Omega}_0$;
- if T has one vertex labeled by $\bar{\Upsilon}$ (by $\bar{\Upsilon}_0$), then $\bar{P}(T) = \bar{\Upsilon}$ $(= \bar{\Upsilon}_0)$;
- if the root r of T is labeled $\odot_r$, and r has a left subtree T_1 and a right subtree T_2 (possibly empty trees), then $\bar{P}(T) = \bar{P}(T_1) \odot^r \bar{P}(T_2)$.

The composition is well defined according to the parse-tree description in the previous paragraph. The represented matroid parsed by T is $P(T) = \bar{P}(T) \,\bar{\oplus}\, \bar{\Omega}$.

Proposition 3.3. *The set $\mathcal{R}_t^{\mathbb{F}}$ (and hence also the tree alphabet Π_t) is finite if $\mathbb{F}$ is a finite field.*

Proof. Let $\odot = (R, \gamma_1, \gamma_2, \gamma_3)$ be a $\leq t$-boundaried composition operator. Clearly, the represented matroid R has at most $3t$ elements and rank at most $3t$ as well. (In fact, for further applications one may restrict the set of composition operators to those with R of rank at most $2t$.) So if t and R are fixed, then there are finitely many choices for the mappings $\gamma_1, \gamma_2, \gamma_3$. For finite $\mathbb{F}$ and each value of t, there are finitely many matrices in $\mathbb{F}^{r\times c}$, $r, c \leq 3t$, and hence finitely many choices for the represented matroid R. ∎

This proposition clearly shows why we deal with parse trees of matroids represented over finite fields in this paper. We remark that, although there are finitely many abstract matroids on at most $3t$ elements up to isomorphism, some of them have infinitely many inequivalent representations over an infinite field $\mathbb{F}$. Moreover, as we show in Section 7, even simple properties of branch-width-3 matroids over the rationals may not be recognizable by finite tree automata.

We say that a branch-decomposition (T, τ) of a matroid is *reduced* if all leaves of T are labeled, and all non-leaves have degree exactly 3. It is easy to see that any branch-decomposition can be turned into a reduced one. We say that a t-boundaried represented matroid $\bar{M}$ is spanning if the boundary subspace $\partial(\bar{M})$ is contained in the span $\langle J(\bar{M})\rangle$ of the internal points of $\bar{M}$. We say that a $\leq t$-boundaried parse tree T is *spanning* if, for each nonempty subtree T_1 of T, the boundaried matroid $\bar{P}(T_1)$ is spanning and nonempty.

Theorem 3.4. *An $\mathbb{F}$-represented matroid M has branch-width at most $t+1$ if and only if M is parsed by some $\leq t$-boundaried spanning parse tree.*

Proof. Let T be a $\leq t$-boundaried parse tree, and let $M = P(T)$. We define a mapping $\tau : E(M) \to \ell(T)$ by $\tau(x) = v$ if x is the element of M created by the boundaried matroid $\bar{\Upsilon}$ or $\bar{\Upsilon}_0$ labeling v in T. We claim that (T, τ) is a width-$(t+1)$ branch-decomposition of the matroid M. Indeed, consider an edge e of T, and denote by T_e the subtree of T below e. Let $F = E(P(T_e))$ and $F' = E - F$. By the definition of a boundary sum (gluing the root of T_e in the parse tree), it is $\partial(\bar{P}(T_e)) \supseteq \langle F\rangle \cap \langle F'\rangle$, and so $\lambda_M(F) = \mathrm{r}\left(\langle F\rangle \cap \langle F'\rangle\right) + 1 \leq \mathrm{r}\left(\partial(\bar{P}(T_e))\right) + 1 \leq t+1$ by Lemma 2.1.

Conversely, suppose that (T, τ) is a reduced width-$(t+1)$ branch-decomposition of a matroid M. We choose an arbitrarily edge e_0 of T, subdivide e_0 with a new vertex r, and make the resulting tree T' rooted at r. We then label each leaf of T' by the boundaried matroid $\bar{\Upsilon}$ or $\bar{\Upsilon}_0$ accordingly. Let P be a point configuration representing M in a projective space Ψ (over $\mathbb{F}$). For an edge e of T', we denote by (A_e, A'_e) the separation of P displayed by e in the branch-decomposition (T, τ). If e was obtained by subdividing e_0, we consider the separation induced by e_0 instead. We denote by Ψ_e the subspace $\Psi_e = \langle A_e\rangle \cap \langle A'_e\rangle$, by B_e an arbitrary basis of Ψ_e, and by $\beta_e : [1, |B_e|\,] \to B_e$ some bijection.

Consider now a vertex v of T' that is not a leaf. Let e_1, e_2 be edges joining v with its left and right sons, respectively, and let e_3 be the edge joining v with its parent. We denote by P_v the point configuration $B_{e_1} \cup B_{e_2} \cup B_{e_3}$, and by R_v the matroid represented by P_v. We take an empty set instead of B_{e_3}

if $v = r$. Finally, we define the composition operator $\odot^v = (R_v, \beta_{e_1}, \beta_{e_2}, \beta_{e_3})$. Since (T, τ) is a width-$(t+1)$ branch-decomposition, we know that $\odot^v$ is a $\le t$-boundaried composition operator. Clearly, for the above labeling, the parse tree T' is spanning, and the represented matroid parsed by T' is M. ∎

4 An Analogue of the Myhill-Nerode Theorem

We now return back to the theory of tree automata, and apply it to our boundaried parse trees for represented matroids. This section uses standard automata-theoretical arguments, following [11, Section 6.4], and it is presented mainly for formal completeness of our paper. From now on, we consider only finite fields $\mathbb{F}$.

We start with the classical Myhill-Nerode theorem for tree automata. We denote by $\Sigma^{**}_{(x)}$ the class of all rooted ordered sub-binary trees labeled by $\Sigma \cup \{x\}$ such that exactly one vertex which is a leaf is labeled by $x \notin \Sigma$. For $T_0 \in \Sigma^{**}$, $T \in \Sigma^{**}_{(x)}$, we denote by $T_0 \bullet_x T$ the tree obtained from T by replacing the leaf of label x with the subtree T_0. Suppose that $\mathcal{L} \subset \Sigma^{**}$ is a tree language. For $T_1, T_2 \in \Sigma^{**}$ we define $T_1 \sim_{\mathcal{L}} T_2$ if and only if the following holds $T_1 \bullet_x T \in \mathcal{L} \iff T_2 \bullet_x T \in \mathcal{L}$ for all $T \in \Sigma^{**}_{(x)}$. Obviously, $\sim_{\mathcal{L}}$ is an equivalence on Σ^{**}.

Theorem 4.1. (Myhill-Nerode theorem for tree automata, e.g. [11, Section 6.1]) *A tree language $\mathcal{L} \subset \Sigma^{**}$ is finite state if and only if the equivalence $\sim_{\mathcal{L}}$ has finite index over Σ^{**}.*

We also need to prove the next technical property of parse trees for matroids.

Lemma 4.2. *Let T be a $\le t$-boundaried parse tree, let v be a vertex of T, and let T_v denote the subtree of T rooted at v. Then there exists a $\le t$-boundaried parse tree T' such that $P(T) = \bar{P}(T_v) \bar{\oplus} \bar{P}(T')$. Moreover, the tree T' depends only on $T - V(T_v)$, but not on T_v.*

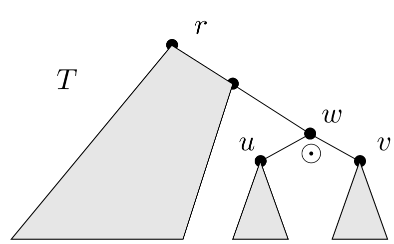

Fig. 8. An illustration to the proof of Lemma 4.2.

Proof. We prove this statement by induction on the distance between v and the root r in T. If $v = r$, then the claim is trivial. Otherwise, let w be the parent of v in T. Up to symmetry, we may assume that u is a left son of w and v is the right son of w. By the inductive assumption, there is a $\le t$-boundaried parse tree T'_w such that $P(T) = \bar{P}(T_w) \bar{\oplus} \bar{P}(T'_w)$. (See in Fig. 8.)

We define the tree T'_v as follows: the root of T'_v is a new vertex w', the left subtree of w' is T_u, and the right subtree of w' is T'_w. Let $\odot = (R, \gamma_1, \gamma_2, \gamma_3)$ be the composition operator labeling w in T. We label w' in T'_v by the composition operator $\odot' = (R, \gamma_1, \gamma_3, \gamma_2)$. Then $\bar{P}(T'_v) = \bar{P}(T_u) \odot' \bar{P}(T'_w)$, and $\bar{P}(T_v) \bar{\oplus} \bar{P}(T'_v) = \big(\bar{P}(T_u) \odot' \bar{P}(T'_w)\big) \bar{\oplus} \bar{P}(T_v) = \big(\bar{P}(T_u) \odot \bar{P}(T_v)\big) \bar{\oplus} \bar{P}(T'_w)$ $= \bar{P}(T_w) \bar{\oplus} \bar{P}(T'_w) = P(T)$ using Lemma 3.2.

The second conclusion immediately follows from the fact that we have not used information about T_v when defining T'_v. ∎

Let $\mathcal{B}_t$ be the set of all $\mathbb{F}$-represented matroids that have branch-width at most t. Let $\mathcal{T}_t \subset \Pi^{**}_{t-1}$ be the language of all $\leq(t-1)$-boundaried parse trees over $\mathbb{F}$ with the alphabet Π_{t-1}, and let $\bar{\mathcal{B}}_t$ be the set of all $\leq(t-1)$-boundaried matroids parsed by the trees from $\mathcal{T}_t$. We know from Theorem 3.4 that $N \in \mathcal{B}_t$ if and only if $N = \bar{N}' \bar{\oplus} \bar{\Omega}$ for some $\bar{N}' \in \bar{\mathcal{B}}_t$. However, notice that not all $\leq(t-1)$-boundaried matroids based on members of $\mathcal{B}_t$ necessarily belong to $\bar{\mathcal{B}}_t$.

It is easy to see that the set of all $\leq(t-1)$-boundaried parse trees $\mathcal{T}_t$ is finite state. Suppose that $\mathcal{M}$ is a set of represented matroids. We shortly say that the set $\mathcal{M}$ is finite state if the collection of all parse trees parsing the members of $\mathcal{M}$ is finite state. We say that $\mathcal{M}$ is *t-width finite state*, $t \geq 1$, if the restriction $\mathcal{M} \cap \mathcal{B}_t$ is finite state. Moreover, we define an equivalence $\approx_{\mathcal{M},t}$ for $\bar{N}_1, \bar{N}_2 \in \bar{\mathcal{B}}_t$ as follows: $\bar{N}_1 \approx_{\mathcal{M},t} \bar{N}_2$ if and only if the boundary ranks of $\bar{N}_1, \bar{N}_2$ are equal, and if $\bar{N}_1 \bar{\oplus} \bar{M} \in \mathcal{M} \iff \bar{N}_2 \bar{\oplus} \bar{M} \in \mathcal{M}$ for all $\bar{M} \in \bar{\mathcal{B}}_t$ of the same boundary rank as $\bar{N}_1, \bar{N}_2$.

Theorem 4.3. *Let $t \geq 1$, and $\mathbb{F}$ be a finite field. A set of $\mathbb{F}$-represented matroids $\mathcal{M}$ is t-width finite state if and only if the equivalence $\approx_{\mathcal{M},t}$ has finite index over $\bar{\mathcal{B}}_t$.*

Proof. Let t be now fixed, and let $\Sigma = \Pi_{t-1}$. We denote by $\mathcal{V} \subseteq \mathcal{T}_t$ the collection of all parse trees corresponding to members of $\mathcal{M}_t = \mathcal{M} \cap \mathcal{B}_t$. The tree language $\mathcal{V}$ is finite state if and only if $\mathcal{M}$ is t-finite state by the definition. By Theorem 4.1, $\mathcal{V}$ is finite state if and only if the equivalence $\sim_{\mathcal{V}}$ has finite index over Σ^{**}. Recall that, for $T_1, T_2 \in \Sigma^{**}$, we define $T_1 \sim_{\mathcal{V}} T_2$ if and only if the following holds $T_1 \bullet_x T \in \mathcal{V} \iff T_2 \bullet_x T \in \mathcal{V}$ for all $T \in \Sigma^{**}_{(x)}$. Hence it is enough to prove that $\sim_{\mathcal{V}}$ has infinite index if and only if $\approx_{\mathcal{M},t}$ has infinite index.

Suppose the latter — that there exist infinitely many $\bar{N}_k \in \bar{\mathcal{B}}_t$, $k \in \mathbb{N}$ such that, for all indices $i \neq j$, there is $\bar{M}_{i,j} \in \bar{\mathcal{B}}_t$ for which $\bar{N}_i \bar{\oplus} \bar{M}_{i,j} \in \mathcal{M}$ but $\bar{N}_j \bar{\oplus} \bar{M}_{i,j} \notin \mathcal{M}$, or vice versa. We may assume without loss of generality that all $\bar{N}_k$, $k \in \mathbb{N}$ are t'-boundaried matroids where $0 \leq t' \leq t-1$ is fixed. Let us denote by $T_k \in \mathcal{T}_t$ the parse tree for $\bar{N}_k$, and by $U_{i,j} \in \mathcal{T}_t$ the parse tree for $\bar{M}_{i,j}$. Let $\odot^o = (R^o, \gamma^o_1, \gamma^o_2, \gamma^o_3)$ be a composition operator such that $\gamma^o_3 : \emptyset \to E(R^o)$, $\gamma^o_1, \gamma^o_2 : [1, t'] \to E(R^o)$, and $\gamma^o_1(n), \gamma^o_2(n)$ are parallel elements for each $n \in [1, t']$.

We define a new tree $W_{i,j} \in \Sigma^{**}_{(x)}$ as follows: $W_{i,j}$ has a root r labeled $\odot^o$, the right subtree of r is $U_{i,j}$, and the left son of r is a leaf labeled x. Notice that $P(T_i \bullet_x W_{i,j}) = \bar{N}_i \bar{\oplus} \bar{M}_{i,j}$, etc. Thus $T_i \bullet_x W_{i,j} \in \mathcal{V}$ but $T_j \bullet_x W_{i,j} \notin \mathcal{V}$, or vice versa; and so the trees $W_{i,j}$ certify that the parse trees T_k, $k \in \mathbb{N}$ are pairwise nonequivalent in $\sim_{\mathcal{V}}$.

Now suppose the former — that there are infinitely many trees $T_k \in \Sigma^{**}$, $k \in \mathbb{N}$ such that, for all indices $i \neq j$, there is $W_{i,j} \in \Sigma^{**}_{(x)}$ for which $T_i \bullet_x W_{i,j} \in \mathcal{V}$ but $T_j \bullet_x W_{i,j} \notin \mathcal{V}$, or vice versa. We may assume without loss of generality that $T_k \in \mathcal{T}_t$ are parse trees for all $k \in \mathbb{N}$ since all trees from $\Sigma^{**} - \mathcal{T}_t$ are equivalent in $\sim_{\mathcal{V}}$, and that all $\bar{P}(T_k)$, $k \in \mathbb{N}$ are t'-boundaried matroids where $0 \leq t' \leq t-1$ is fixed. Consider, up to symmetry, indices $i \neq j$ such that $T_i \bullet_x W_{i,j} \in \mathcal{V} \subseteq \mathcal{T}_t$. Our assumptions now imply that also $T_j \bullet_x W_{i,j} \in \mathcal{T}_t$, as can be easily checked from the definition of a $\leq(t-1)$-boundaried parse tree.

We denote by $\bar{N}_k = \bar{P}(T_k)$. By Lemma 4.2, there exists a tree $U_{i,j} \in \mathcal{T}_t$ such that $P(T_i \bullet_x W_{i,j}) = \bar{N}_i \bar{\oplus} \bar{P}(U_{i,j})$ and $P(T_j \bullet_x W_{i,j}) = \bar{N}_j \bar{\oplus} \bar{P}(U_{i,j})$. Let $\bar{M}_{i,j} = \bar{P}(U_{i,j})$. Recall that $T_i \bullet_x W_{i,j} \in \mathcal{V}$ but $T_j \bullet_x W_{i,j} \notin \mathcal{V}$, or vice versa, for all $i \neq j$. Then, by the definition of $\mathcal{V}$, the t'-boundaried matroids $\bar{M}_{i,j} \in \bar{\mathcal{B}}_t$ certify that the t'-boundaried matroids $\bar{N}_k \in \bar{\mathcal{B}}_t$, $k \in \mathbb{N}$ belong to pairwise different equivalence classes of $\approx_{\mathcal{M},t}$. ∎

Remark. The property "$\approx_{\mathcal{M},t}$ has finite index" is called sometimes "t-cutset regularity" of $\mathcal{M}$ [1] (also noted in [11, Definition 6.76]). We add an informal remark to this interesting concept since its formal definition may sound confusing. The true meaning of $\approx_{\mathcal{M},t}$ having finite index is that, regardless of a choice of $\bar{N} \in \bar{\mathcal{B}}_t$, only a bounded amount of information relevant to membership in $\mathcal{M}$ may "cross" the boundary of $\bar{N}$.

5 Monadic Second-Order Logic

Our paper has been inspired by the so-called "MS_2-theorem" (Theorem 5.1) for graphs. This theorem is a high-level theoretical tool for establishing that various hard graph properties, expressible in monadic second-order theory of graphs, are t-width finite state, i.e. recognizable by a finite tree automaton for fixed tree-width. This section introduces monadic second-order logic and the associated theories of graphs and matroids, and reviews some important results.

5.1 MS_2 Theory of Graphs

We shortly write MSO to stand for *monadic second-order logic*. The language of MSO logic applied over incidence graphs forms the *monadic second-order theory* MS_2 of graphs. Precisely, the syntax of MS_2 includes variables for vertices, edges, and their sets, the quantifiers $\forall, \exists$ applicable to these variables, the logical connectives $\wedge, \vee, \neg$, and the next binary relations:

1. $=$, the equality for vertices, edges, and their sets,
2. $v \in W$, where v is a vertex and W is a vertex set variables,
3. $e \in F$, where e is an edge and F is an edge set variables,
4. $\mathrm{inc}(v, e)$, where v is a vertex variable and e is an edge variable, and the relation tells whether v is incident with e.

We remark that the above language is sometimes called an "extended MSO logic" of graphs, since it is allowed to quantify over both vertices and edges. Opposed to that, the so called "basic" monadic second-order theory MS_1 of graphs results by using the same MSO logic over adjacency graphs. (An adjacency graph has a binary relation for edges, i.e. edges are not objects unlike vertices.) The expressive power of MS_1 is known to be strictly weaker than that of MS_2.

Parse trees for graphs of bounded tree-width are informally sketched in Section 3. Let $\mathcal{G}$ be a graph family. Analogously to the previous section, we say that $\mathcal{G}$ is *t-width finite state* if the subset of all tree-width-t members of $\mathcal{G}$ is finite state. The MS_2-theorem (in a tree-automata formulation) follows.

Theorem 5.1. (Courcelle [6, 7]) *If $\mathcal{G}$ is a family of graphs described by a sentence in MS_2, then $\mathcal{G}$ is t-width finite state for every $t \geq 1$.*

Similar results were obtained also by Arnborg, Lagergren, and Seese [2], and later by Borie, Parker, and Tovey [5] (in explicit algorithmic formulations).

5.2 MS_M Theory of Matroids

As noted in the introduction, the concept of a branch-width is very close to that of a tree-width in graphs. If we want to extend Theorem 5.1 to matroids, we have to introduce an MSO logic of matroids. Recall that there is no analogue of graph vertices in matroids, instead, matroid elements are analogous to graph edges. So it has better sense to extend to matroids the MS_2 theory, but "without" vertices. (See also Section 8 for our definition of MSO logic of "edge-only" graphs.)

Let us consider a matroid as a structure formed by the elements and a relation for independent subsets. Applying the language of MSO logic to such matroid structures gives the *monadic second-order theory MS_M of matroids*:

Definition. The syntax of MS_M includes variables for matroid elements and element sets, the quantifiers $\forall, \exists$ applicable to these variables, the logical connectives $\wedge, \vee, \neg$, and the next predicates:

1. $=$, the equality for elements and their sets,
2. $e \in F$, where e is an element and F is an element set variables,
3. $\mathit{indep}(F)$, where F is an element set variable, and the predicate is true iff F is an independent set in the matroid.

Moreover, we write $\phi \rightarrow \psi$ to stand for $\neg\phi \vee \psi$, and $X \subseteq Y$ for $\forall x (x \in Y \vee x \notin X)$.

Notice that the "universe" of a formula (the model in logic terms) in the above theory is a finite (abstract) matroid. To give a better feeling for the expressive power of the MS_M logic, we show a few additional basic matroid predicates. Recall that an independent set in a matroid is a subset of a basis (for example, an acyclic subset in a cycle matroid of a graph), and that a circuit in a matroid is a minimal set not contained in any basis.

– We write $\mathit{basis}(B) \equiv \mathit{indep}(B) \wedge \forall D \big(B \not\subseteq D \vee B = D \vee \neg\, \mathit{indep}(D) \big)$ to express the fact that a basis is a maximal independent set.

– Similarly, we write $\mathit{circuit}(C) \equiv \neg\, \mathit{indep}(C) \wedge \forall D \big(D \not\subseteq C \vee D = C \vee \mathit{indep}(D) \big)$, saying that C is dependent, but all proper subsets of C are independent.

It is, of course, possible to define an MSO theory of matroids using any one of $\mathit{indep}, \mathit{basis}, \mathit{circuit}$ as the atomic predicate, and to express the other two predicates similarly as above.

One may also define *counting* MS_2 or MS_M logics by adding the predicates $\mathrm{mod}_{p,q}$ for integer constants $0 \leq p < q$, where the interpretation of $\mathrm{mod}_{p,q}(X)$ for an arbitrary set variable X is that $|X| \bmod q = p$. Let us shortly denote the counting versions by CMS_2 or CMS_M, respectively. The CMS_2 logic is known to be stronger [7] than plain MS_2 over all graphs. The full statement of Theorem 5.1 is actually formulated for CMS_2.

We show next that the language of CMS_M is at least as powerful as that of CMS_2. Notice, however, that such a translation from graphs to their cycle matroids is not straightforward since non-isomorphic graphs of low connectivity may have isomorphic cycle matroids. (Likewise trees having cycle matroids with all independent sets.) Let $G \uplus H$ denotes the graph obtained from disjoint copies of G and H by adding all edges between them.

Theorem 5.2. *Let G be a loopless multigraph, and let M be the cycle matroid of $G \uplus K_3$. Then any sentence about G in CMS_2 can be expressed as a sentence about M in CMS_M.*

Since this theorem does not belong to the core results of our paper, we postpone its proof till Section 8. On the other hand, we remark that since the predicate $\mathit{indep}(F)$ can be formulated in MS_2, the expressive power of MS_M over cycle matroids of graphs is essentially the same as that of MS_2 over graphs.

In addition to Theorem 5.2, we show some simple interesting matroidal properties in MS_M. Connectivity of a matroid can be expressed as $\mathit{connected} \equiv \forall e, f\, \exists C \left(e, f \in C \wedge \mathit{circuit}(C) \right)$, which means that every two elements lie in a common circuit. The graph property of being Hamiltonian has a quite complex MS_2 expression. (Hamiltonicity even cannot be formulated in MS_1.) On the other hand, in the matroid language a Hamiltonian cycle is a spanning circuit, i.e. a circuit containing a basis. So we may write $\mathit{hamiltonian} \equiv \exists C, B \big(\mathit{circuit}(C) \wedge \mathit{basis}(B) \wedge B \subseteq C \big)$.

Other matroidal properties in MS_M are described in [17], such as k-connectivity, branch-width k, paving matroids (a property important in design theory), binary transversal matroids, or (importantly) matroid minors. Many interesting properties can be then described using the minor relation, like all minor-closed matroid properties subject to bounded branch-width and representability [17]. For an illustration, we show a sentence describing that a matroid has a triangle (a circuit of 3 elements) as a minor: $\exists C\, \exists x, y, z \big(\mathit{circuit}(C) \wedge x, y, z \in C \wedge x \neq y \neq z \neq x \big)$.

6 MS_M-Theorem for Represented Matroids

We are now ready to formulate and prove a natural extension of Theorem 5.1 to MS_M logic over represented matroids. We follow the ideas of the Abrahamson-Fellows' alternative proof [1] of Theorem 5.1, in the form published in [11, Section 6.5]. (Necessity of assuming matroids represented over a fixed finite field here is justified by a further negative result in Corollary 8.6.)

Theorem 6.1. *Let $\mathbb{F}$ be a finite field. If $\mathcal{M}$ is a set of $\mathbb{F}$-represented matroids described by a sentence in the logic CMS_M over matroids, then $\mathcal{M}$ is t-width finite state for every $t \geq 1$.*

Recall the sets $\mathcal{B}_t$ and $\bar{\mathcal{B}}_t$ defined for $\mathbb{F}$, and the equivalence relation $\approx_{\mathcal{M},t}$ defined over $\bar{\mathcal{B}}_t$ by $\mathcal{M}$, from Section 4. Let ϕ be the sentence in CMS_M that describes the set $\mathcal{M}$. Let us fix an integer $0 \leq u \leq t-1$ for the rest of the proof, and write $\approx_\phi$ for $\approx_{\mathcal{M},t}$ restricted to exactly u-boundaried matroids over $\mathbb{F}$. We are going to prove the theorem by induction on the length of ϕ.

In order to use induction, we must slightly generalize the setting. We allow a formula ϕ with (possible) free variables, and we associate with such ϕ an equivalence relation $\approx_\phi^{\cdot}$ on the set of all u-boundaried represented matroids over $\mathbb{F}$ that are partially "equipped" with distinguished elements and sets corresponding to the free variables $Free(\phi)$ in ϕ. The relation $\approx_\phi^{\cdot}$ naturally generalizes the relation $\approx_{\mathcal{M},t}$ to partially equipped boundaried matroids defined next.

Let $Free(\phi) = Fr(\phi) \cup FR(\phi)$ be the partition of the free variables into those $Fr = Fr(\phi)$ for elements and those $FR = FR(\phi)$ for sets of elements. We define a *partial equipment signature* as a triple $\sigma = (Fr, FR, f)$ where $f : Fr \to \{0,1\}$. A boundaried represented matroid $\bar{M}$ is said to be *σ-partially equipped* if it has distinguished elements and element sets assigned to the free variables in σ. Formally, for each variable $X \in FR$ there is a distinguished subset $S_X \subseteq J(\bar{M})$ of internal elements of $\bar{M}$, and for each variable $x \in Fr$ such that $f(x) = 0$ there is a distinguished internal element $e_x \in J(\bar{M})$. Nothing is assigned to variables $x \in Fr$ such that $f(x) = 1$. We say that a partial equipment signature $\sigma' = (Fr, FR, f')$ is a *complement* of σ if $f'(x) = 1 - f(x)$ for all $x \in Fr$.

The importance of the complemented partial equipment signature σ' lies in the following fact: If $\bar{N}, \bar{N}'$ are boundaried matroids such that $\bar{N}$ is σ-partially equipped and $\bar{N}'$ is σ'-partially equipped, then the free variables from $Free = Fr \cup FR$ have consistent full interpretation over the whole matroid $\bar{N} \bar{\oplus} \bar{N}'$. A represented matroid M is *fully equipped* for ϕ if all free variables from $Free(\phi)$ have interpretation in $E(M)$. We write $M \models \phi$ to mean that the formula ϕ is true on M with the associated full interpretation for $Free(\phi)$.

Definition. Let σ be a partial equipment signature for a formula ϕ, and let σ' be the complement of σ. Suppose that $\bar{N}_1, \bar{N}_2$ are σ-partially equipped u-boundaried matroids. We define $\bar{N}_1 \approx_\phi^\sigma \bar{N}_2$ if and only if $\bar{N}_1 \bar{\oplus} \bar{M} \models \phi \iff \bar{N}_2 \bar{\oplus} \bar{M} \models \phi$ for every σ'-partially equipped u-boundaried matroid $\bar{M}$.

For reference when speaking about the equivalence classes of $\approx_\phi^\sigma$, we call the boundaried matroids $\bar{N}_i$ from the previous definition as "left-hand side" matroids. We may now precisely formulate our induction statement.

Lemma 6.2. *Let ϕ be a formula in the MS_M logic of matroids, and let σ be a partial equipment signature for ϕ. Then $\approx_\phi^\sigma$ has finite index on the universe of σ-partially equipped u-boundaried matroids.*

Proof. Let $\sigma = (Fr, FR, f)$ where $Free(\phi) = Fr \cup FR$ as above. Unless stated otherwise, we implicitly consider σ-partially equipped u-boundaried matroids. We first show that the statement holds for atomic formulas. The empty formula is trivial. The proofs for equality formulas are easy. Say, if ϕ is $x = y$ where $x, y \in Fr$ and $f(x) = 0$, $f(y) = 1$, then ϕ is true in no interpretation (equipment) of x, y by definition of the equipment signature, and so $\approx_\phi^\sigma$ has index 1. If $f(x) = f(y) = 1$ for the same ϕ, then again $\approx_\phi^\sigma$ has index 1 since the outcome of $x = y$ depends only on an interpretation of x, y in $\bar{M}$. If $f(x) = f(y) = 0$, then $\approx_\phi^\sigma$ has two equivalence classes; one of them contains all left-hand side matroids with an interpretation of x, y as $e_x = e_y$, and the other one contains all remaining partially equipped matroids. Similarly to the last case, if ϕ is $X = Y$ where $X, Y \in FR$, then $\approx_\phi^\sigma$ has two equivalence classes.

If ϕ is $\mathrm{mod}_{p,q}(X)$ where $X \in FR$, then $\approx_\phi^\sigma$ clearly has index q since the equivalence classes of $\approx_\phi^\sigma$ are given by the values of $|S_X| \bmod q$, where S_X is an interpretation of the variable X in the left-hand side matroids. If ϕ is $x \in X$, then $\approx_\phi^\sigma$ has index 1 for $f(x) = 1$, while $\approx_\phi^\sigma$ has index 2 for $f(x) = 0$. In the former case, the outcome of $x \in X$ depends only on an interpretation of x and X in $\bar{M}$, and so all left-hand side matroids are equivalent in the above definition. In the latter case, one equivalence class of $\approx_\phi^\sigma$ contains all those matroids with an interpretation $e_x \in S_X$, and the other equivalence class of $\approx_\phi^\sigma$ contains all remaining partially equipped matroids.

The only really interesting atomic formula is $\phi \equiv \mathit{indep}(X)$ where $X \in FR$. This is not surprising as the predicate *indep* determines the whole matroid structure, and it is the only place in the proof where we use the fact that the boundaries of our matroids have fixed rank.

Claim 6.3. For $\phi \equiv \mathit{indep}(X)$, the equivalence $\approx_\phi^\sigma$ is of (bounded) index $1 + p(u, \mathbb{F})$ where $p(n, \mathbb{F})$ stands for the number of all distinct subspaces of the finite projective space $PG(n, \mathbb{F})$.

Proof. Suppose that $\bar{N}_i = (N_i, \delta_i)$, $i = 1, 2$ are σ-partially equipped boundaried matroids, and that $F_X^i \subseteq J(\bar{N}_i)$ is an interpretation of the set variable X in $\bar{N}_i$. If F_X' is an interpretation of X in a matroid $\bar{M}$, then, by the definition of a boundary sum, a linear dependency among elements of $F_X^i \cup F_X'$ in $\bar{N}_i \bar{\oplus} \bar{M}$ is fully determined by the sets F_X^i, F_X' themselves, and the intersections of $\langle F_X^i \rangle$ and $\langle F_X' \rangle$ with the common boundary of the sum. So, $\bar{N}_1 \approx_\phi^\sigma \bar{N}_2$ if and only if either both F_X^1 and F_X^2 are dependent, or the subspaces $\langle F_X^1 \rangle \cap \partial(\bar{N}_1)$ and $\langle F_X^2 \rangle \cap \partial(\bar{N}_2)$ are equivalent in the linear transformation $\mathcal{L}$ matching the boundary points of $\bar{N}_1$ to those of $\bar{N}_2$ in order (i.e. $\delta_2(j) = \mathcal{L}(\delta_1(j))$ for $j \in [1, u]$).

Hence the claim follows easily from the assumption that the boundaries of $\bar{N}_1$ and $\bar{N}_2$ have bounded rank u. □

For the inductive step, we consider that the formula ϕ is created from shorter formula(s) in one of the following ways: $\phi \equiv \neg\phi_1$, $\phi_1 \wedge \phi_2$, $\exists x \phi_1(x)$, or $\exists X \phi_1(X)$, where $x \in Fr(\phi_1)$ or $X \in FR(\phi_1)$ in the latter cases. One may easily express the $\vee$ or $\forall$ symbols using these. We assume by induction that $\approx_{\phi_1}^{\sigma_1}$ ($\approx_{\phi_2}^{\sigma_2}$) has finite index, where a signature σ_1 is inherited from σ for ϕ_1 (see below for case-by-case details). The first case of $\phi \equiv \neg\phi_1$ is quite easy to resolve — the equivalence $\approx_\phi^\sigma$ is, in fact, the same as $\approx_{\phi_1}^\sigma$. Let us look at the second case $\phi \equiv \phi_1 \wedge \phi_2$.

Claim 6.4. Let $\phi \equiv \phi_1 \wedge \phi_2$, and let σ_i denote the restriction of σ to $Free(\phi_i)$, for $i = 1, 2$. (Notice that $Free(\phi) = Free(\phi_1) \cup Free(\phi_2)$.) Assume the equivalence $\approx_{\phi_i}^{\sigma_i}$, $i = 1, 2$, is of index q_i. Then the index of $\approx_\phi^\sigma$ is at most $q_1 \cdot q_2$.

Proof. Suppose that $\bar{N}_1 \not\approx_\phi^\sigma \bar{N}_2$ are two σ-partially equipped boundaried matroids, and that $\bar{M}$ is such that $\bar{N}_1 \bar{\oplus} \bar{M} \models \phi$ but $\bar{N}_2 \bar{\oplus} \bar{M} \models \neg\phi$. Then, for some $i \in \{1, 2\}$, the boundaried matroids $\bar{N}_1, \bar{N}_2, \bar{M}$ (considered now with the restricted σ_i-partial equipments) must satisfy $\bar{N}_1 \bar{\oplus} \bar{M} \models \phi_i$, but $\bar{N}_2 \bar{\oplus} \bar{M} \models \neg\phi_i$. So $\bar{N}_1 \not\approx_{\phi_i}^{\sigma_i} \bar{N}_2$ for some $i \in \{1, 2\}$, and the claim follows. (One may say that the equivalence classes of $\approx_\phi^\sigma$ are unions of the equivalence classes of the intersection $\approx_{\phi_1}^{\sigma_1} \cap \approx_{\phi_2}^{\sigma_2}$, after formally adding an arbitrary – meaningless, equipment of $Free(\phi) - Free(\phi_i)$ to the σ_i-partial equipments, $i = 1, 2$.) □

Claim 6.5. Let $\phi \equiv \exists X \phi_1(X)$ for $X \in FR(\phi_1)$, and let $\sigma_1 = (Fr, FR \cup \{X\}, f)$. Assume the equivalence $\approx_{\phi_1}^{\sigma_1}$ is of index c. Then the index of $\approx_\phi^\sigma$ is less than 2^c.

Proof. Again, suppose that (arbitrary) $\bar{N}_1 \not\approx_\phi^\sigma \bar{N}_2$ and $\bar{M}$ are such that $\bar{N}_1 \bar{\oplus} \bar{M} \models \phi$, but $\bar{N}_2 \bar{\oplus} \bar{M} \models \neg\phi$. Let σ_1' be the complemented signature to σ_1. We shortly write $\bar{N}[X{=}S]$ for the σ_1-partially equipped matroid obtained from σ-partially equipped $\bar{N}$ by interpreting X as $S \subseteq J(\bar{N})$. Then our assumption $\bar{N}_1 \bar{\oplus} \bar{M} \models \phi \equiv \exists X \phi_1$ says that there exist $S_X \subseteq J(\bar{N}_1)$ and $S_X' \subseteq J(\bar{M})$ such that $\bar{N}_1[X = S_X] \bar{\oplus} \bar{M}[X = S_X'] \models \phi_1$. On the other hand, $\bar{N}_2 \bar{\oplus} \bar{M} \models \neg\phi$ implies that $\bar{N}_2[X = T_X] \bar{\oplus} \bar{M}[X = S_X'] \models \neg\phi_1$ for all $T_X \subseteq J(\bar{N}_2)$, and so it is $\bar{N}_1[X{=}S_X] \not\approx_{\phi_1}^{\sigma_1} \bar{N}_2[X{=}T_X]$.

We now, in a search for a contradiction, look at the problem from the other side. Let the equivalence classes of $\approx_{\phi_1}^{\sigma_1}$ be $\mathcal{C}^1, \mathcal{C}^2, \ldots, \mathcal{C}^c$. For a σ-partially equipped matroid $\bar{N}$, we define a nonempty set $Ix(\bar{N}) \subseteq \{1, 2, \ldots, c\}$ as follows: $i \in Ix(\bar{N})$ if and only if $\bar{N}[X = S] \in \mathcal{C}^i$ for some $S \subseteq J(\bar{N})$. If there were 2^c pairwise incomparable σ-partially equipped matroids in the relation $\approx_\phi^\sigma$, then some two of them, say $\bar{N}_1 \not\approx_\phi^\sigma \bar{N}_2$, would get $Ix(\bar{N}_1) = Ix(\bar{N}_2)$ by the pigeon-hole principle. However, by the argument of the previous paragraph – $\bar{N}_1[X{=}S_X] \not\approx_{\phi_1}^{\sigma_1} \bar{N}_2[X{=}T_X]$ for some $S_X \subseteq J(\bar{N}_1)$ and all $T_X \subseteq J(\bar{N}_2)$, we have that $j \in Ix(\bar{N}_1) - Ix(\bar{N}_2)$ where j is such that $\bar{N}_1[X{=}S_X] \in \mathcal{C}^j$. This final contradiction proves our claim. □

Claim 6.6. Let $\phi \equiv \exists x \phi_1(x)$ for $x \in Fr(\phi_1)$, and let $\sigma_1 = (Fr \cup \{x\}, FR, f_1)$, $\sigma_2 = (Fr \cup \{x\}, FR, f_2)$ where $f_1(x) = 0$, $f_2(x) = 1$. Assume the equivalences

$\approx^{\sigma_1}_{\phi_1}$ and $\approx^{\sigma_2}_{\phi_1}$ are of indices c_1 and c_2, respectively. Then the index of $\approx^{\sigma}_{\phi}$ is at most $2^{c_1}c_2 + 1 - c_2$.

Proof. Let $\bar{N}$ be a σ-partially equipped u-boundaried matroid. Firstly, recall from Claim 6.5 the notation $\bar{N}[x{=}e]$ for a matroid $\bar{N}$ with an additional partial equipment of the variable x. Notice that a σ_2-partial equipment of $\bar{N}$ does not interpret the variable x inside $J(\bar{N})$, and so $\bar{N}$ may be view as well as a σ_2-partially equipped matroid.

Suppose that $\bar{N}_1 \not\approx^{\sigma}_{\phi} \bar{N}_2$ and $\bar{M}$ are such that $\bar{N}_1 \bar{\oplus} \bar{M} \models \phi$ but $\bar{N}_2 \bar{\oplus} \bar{M} \models \neg\phi$. In other words, $\bar{N}_1 \bar{\oplus} \bar{M} \models \exists x \phi_1(x)$ but $\bar{N}_2 \bar{\oplus} \bar{M} \models \forall x \neg\phi_1(x)$. Suppose also that both $J(\bar{N}_1)$ and $J(\bar{N}_2)$ are nonempty. Let $e_x \in E(\bar{N}_1 \bar{\oplus} \bar{M})$ be the interpretation of the variable x that satisfies ϕ_1 over $\bar{N}_1 \bar{\oplus} \bar{M}$. In particular, ϕ_1 is false over $(\bar{N}_2 \bar{\oplus} \bar{M})[x = e_x]$. If $e_x \in J(\bar{M})$, then immediately $\bar{N}_1 \not\approx^{\sigma_2}_{\phi_1} \bar{N}_2$. Otherwise, it is $e_x \in J(\bar{N}_1)$, and we are in a situation analogous to Claim 6.5: $\bar{N}_1[x{=}e_x] \bar{\oplus} \bar{M} \models \phi_1$, but $\bar{N}_2[x{=}f_x] \bar{\oplus} \bar{M} \models \neg\phi_1$ for all $f_x \in J(\bar{N}_2)$.

Now, looking for a contradiction, we assume that there are $2^{c_1}c_2 + 2 - c_2$ pairwise incomparable σ-partially equipped matroids in the relation $\approx^{\sigma}_{\phi}$. Then at least $2^{c_1}c_2 + 1 - c_2 = (2^{c_1} - 1)c_2 + 1$ of those are not equal to the empty u-boundaried matroid $\bar{\Omega}_u$, and out of them at least 2^{c_1} pairwise incomparable matroids with respect to $\approx^{\sigma}_{\phi}$ belong to the same equivalence class of $\approx^{\sigma_2}_{\phi_1}$. Let us denote their set by $\mathcal{N}$. (Hence for each pair in $\mathcal{N}$, the latter conclusion of the previous paragraph applies.) Denoting the equivalence classes of $\approx^{\sigma_1}_{\phi_1}$ by $\mathcal{C}^1, \mathcal{C}^2, \ldots, \mathcal{C}^{c_1}$, we again define a nonempty set $Ix(\bar{N}) \subseteq \{1, 2, \ldots, c_1\}$ as follows: $i \in Ix(\bar{N})$ if and only if $\bar{N}[x{=}e] \in \mathcal{C}^i$ for some $e \in J(\bar{N})$. Then some pair, say $\bar{N}_1, \bar{N}_2 \in \mathcal{N}$, satisfies $Ix(\bar{N}_1) = Ix(\bar{N}_2)$ by the pigeon-hole principle. However, that contradicts the above latter conclusion; $\bar{N}_1[x{=}e_x] \bar{\oplus} \bar{M} \models \phi_1$ for some $e_x \in J(\bar{N}_1)$, but $\bar{N}_2[x{=}f_x] \bar{\oplus} \bar{M} \models \neg\phi_1$ for all $f_x \in J(\bar{N}_2)$. □

We have finished all necessary steps in the inductive proof of the lemma. ∎

We apply Lemma 6.2 to the (closed) formula ϕ describing the set $\mathcal{M}$ in Theorem 6.1 and to $u = 0, 1, \ldots, t-1$, showing that $\approx_{\mathcal{M},t}$ has finite index for each u when restricted to exactly u-boundaried matroids. (Notice that we have established above more than what was required — we have proved that $\approx_{\mathcal{M},t}$ has finite index over all u-boundaried matroids, not only over u-boundaried matroids from $\bar{\mathcal{B}}_t$.) Hence $\approx_{\mathcal{M},t}$ has finite index for each t, and the proof of Theorem 6.1 is finished by Theorem 4.3.

Remark. We note in passing, that the proofs of Lemma 6.2 and of Theorem 4.3 are constructive in the following sense: If ϕ is a given MS_M formula, then we can construct the equivalence classes of $\approx_{\phi}$ as defined above. These equivalence classes then, essentially, provide the states of the constructed finite tree automaton in Theorem 4.3. Hence there is an algorithm that computes this finite tree automaton for the given width t, finite field $\mathbb{F}$, and formula ϕ.

Remark. It is, on the other hand, possible to give a similar proof of Theorem 6.1 in the setting of the logic method of "interpretation" in labeled trees. That is the approach originally taken for graphs by Arnborg, Lagergren and Seese [2]. The general interpretability method for arbitrary structures is also surveyed in [20]. We prefer the direct combinatorial approach in our paper.

7 Computing Aspects

In this section we review some complexity implications of our results. (More algorithmic applications can be found in [17].) We first include a remark on matroid complexity in general: An n-element (abstract) matroid carries an amount of information which is exponential in n, and so to speak about complexity of matroid algorithms, one has to decide about the input representation of a matroid. We use vector representations of matroids over a field $\mathbb{F}$, and hence an n-element matroid is given by an $n \times r$ matrix where r is the matrix rank, which typically means an input size of order $\Omega(n^2)$. On the other hand, a matroid parse tree of bounded width over a finite field has only linear $O(n)$ size.

Since the computation of a finite tree automaton may be easily emulated by a linear-time algorithm, we get the following corollary of Theorem 6.1.

Corollary 7.1. *Let $t \geq 1$, let $\mathbb{F}$ be a finite field, and let ϕ be a CMS_M sentence. There is an algorithm that, for given $\leq t$-boundaried matroid parse tree T over $\mathbb{F}$, decides in linear time whether ϕ is true for the matroid $P(T)$ parsed by T.*

To use Corollary 7.1 in a practical computation, we first have to construct a parse tree for the given $\mathbb{F}$-represented matroid of bounded branch-width.

Theorem 7.2. (PH [19]) *Let $\mathbb{F}$ be a finite field, and $t \geq 1$. Given an $\mathbb{F}$-represented n-element matroid M of branch-width t, one can construct in time $O(n^3)$ a $\leq(3t)$-boundaried spanning parse tree T over $\mathbb{F}$ such that $P(T) \simeq M$.*

Unfortunately, the parse tree constructed in this theorem does not necessarily have the optimal width, but there is a computable finite set of forbidden minors for the class of matroids of branch-width at most t for each t by [13]. Hence we can construct an MS_M sentence ϕ_k such that $M \models \phi_k$ iff M has branch-width at most k, and conclude:

Corollary 7.3. (PH [19]) *Let $\mathbb{F}$ be a finite field, and $t \geq 1$. The problem to decide whether given $\mathbb{F}$-represented n-element matroid M has branch-width at most t is fixed-parameter tractable, i.e. it can be solved in time $O(f(t) \cdot n^3)$.*

Moreover, the main result of [12] implies that, for any minor-closed matroid family $\mathcal{M}$ and a fixed t, there are only finitely many $\mathbb{F}$-representable forbidden minors for $\mathcal{M}$ of branch-width at most t. (This result is non-constructive.) Hence:

Theorem 7.4. (PH [17]) *Let $\mathbb{F}$ be a finite field. For every minor-closed matroid family $\mathcal{M}$, and $t \geq 1$, there is an algorithm deciding whether $M \in \mathcal{M}$ on an $\mathbb{F}$-represented matroid M of branch-width at most t in time $O(n^3)$.*

In particular, one may compute matroid tree-width [21] using Theorem 6.1. Besides applications based directly on the theorem, one may use the machinery of matroid parse trees from Section 3 for solving other problems, like for computing the Tutte polynomial of a represented matroid [18].

Let us finish this section with a note on decidability of MSO theories of graphs and matroids. A theory (with an associated class of models) is said to be decidable if there is an algorithm that for any given sentence ψ decides whether ψ is true on all models of the theory. The following result is due to Seese.

Theorem 7.5. (Seese [27]) *Suppose that $\mathcal{G}$ is a family of graphs with a decidable MS_2 theory. Then there is a number t such that the tree-width of each $G \in \mathcal{G}$ is at most t.*

Notice the difference of this theorem from Theorem 5.1; here we do not care about an algorithmic power or time needed to decide a particular sentence, but we have to verify a sentence over the whole (usually infinite) family $\mathcal{G}$ instead over a given one graph. The theorem does not seem to directly apply to our matroid case, but Seese's proof idea does so — using a so called "excluded grid" theorem for matroids [14], we manage to prove:

Theorem 7.6. (PH and Seese [20]) *Suppose that $\mathcal{M}$ is a family of matroids representable over a finite field $\mathbb{F}$. If the MSO theory of $\mathcal{M}$ is decidable, then there is a number t such that the branch-width of each $M \in \mathcal{M}$ is at most t.*

At this point it is interesting to mention also a recent theorem of Courcelle and Oum [10] that graph families with a decidable C_2MS_1 theory have bounded clique-width (a result similar to Theorem 7.5), a topic which is further discussed in Section 8.4.

8 Final Notes and Related Concepts

In the last section we discuss a close relation of matroid MS_M theories to MS_2 theories of graphs, and subsequently of our Theorem 6.1 to Theorem 5.1 of Courcelle. Then we exhibit some limits of possible extensions of our research to general matroids, and conclude with showing some interestingly related "width" parameters of graphs and matrices. The main purpose is to set our results in the context of recent (and current) active research in discrete mathematics.

8.1 Edge-only graphs

As noted in the introduction, the concept of a matroid branch-width is very close to that of a graph tree-width. If we want to relate our Theorem 6.1 to Theorem 5.1, we have to find a natural correspondence between CMS_2 and CMS_M. The underlying correspondence of the objects is shown in the definition of the cycle matroid of a graph. One trouble is that there are no analogues of graph vertices in matroids. In this context we consider a structure that we call an *edge-only graph*, that is a graph which has only edges and a "star" relation, where a *star* in a graph is the set of all edges incident with one vertex (a center of the star).

The syntax of the monadic second-order logic MS_e of edge-only graphs includes variables for edges and edge sets, the quantifiers $\forall, \exists$ applicable to these variables, the logical connectives $\wedge, \vee, \neg$, and the next predicates:

1. $=$, the equality for edges and their sets,
2. $e \in F$, where e is an edge and F is an edge set variables,
3. $\mathrm{star}(F)$, where F is an edge set variable, and $\mathrm{star}(F)$ is true iff F is the set of all edges (incl. loops) incident with some vertex in the graph.

It is clear that any sentence in MS_e is expressible in MS_2. For example, $\mathrm{star}(F) \equiv \exists v \forall e\, (\mathrm{inc}(v, e) \leftrightarrow e \in F)$. Observe that if a connected graph has more than 2 vertices, then two stars F_1, F_2 of edges are equal if and only if they are centered at the same vertex, and so the stars fully describe vertices of the graph. It is known that the previously mentioned MS_1 logic is strictly weaker than MS_2. On the other hand, we show that the MS_e logic is equally strong as MS_2 over connected graphs on more than 2 vertices.

Lemma 8.1. *Let G be a connected multigraph on more than 2 vertices, and let $U \subseteq V(G)$ be a set of vertices. Choose $u \in U$, and denote by X_U the set of all edges of G that have exactly one end in U. Let $v \in V(G)$ be arbitrary, and denote by F_u, F_v the stars of edges centered at u, v, respectively. If $U = \emptyset$, choose $F_u = \emptyset$. Then the following are true:*

- *$v \in U$ if and only if $F_v = F_u$, or if there exists a path $P \subset G$, such that $|E(P) \cap F_u| = |E(P) \cap F_v| = 1$, and $|E(P) \cap X_U|$ is even.*
- *The previous characterization of $v \in U$ can be expressed with an MS_e predicate $\mathsf{vertexinset}(F_v; X_U, F_u)$.*

Proof. Since G is connected, there is a path joining u to v in G, and it is straightforward to verify the required properties. If $U = \emptyset$, then $F_u = \emptyset$ as well, and the condition $|E(P) \cap F_u| = 1$ is always false. Conversely, assume an existence of the path P. Since the sets F_u, F_v contain all edges incident with vertices u, v, respectively, it must be that u and v are the ends of P. Then the parity condition on $|E(P) \cap X_U|$ guarantees that both ends of P belong to the same side of the bipartition $(U, E(G) - U)$, and so $v \in U$.

In the following proof we use several shortcuts: We write $|X| = 1$ to stand for $\exists x (x \in X) \wedge \forall x, y (x \notin X \vee y \notin X \vee x = y)$, similarly $|X| = 2$ for $\exists x, y (x, y \in X \wedge x \neq y) \wedge \forall x, y, z (x \notin X \vee y \notin X \vee z \notin X \vee x = y \vee x = z \vee z = y)$, and $x \in X \cap Y$ for $x \in X \wedge x \in Y$. We write $\mathit{disjunion}(X; Y, Z) \equiv \forall x \big(x \notin X \vee x \in Y \vee x \in Z\big) \wedge \forall x \big(x \in X \vee x \notin Y \wedge x \notin Z\big) \wedge \forall x \big(x \notin Y \vee x \notin Z\big)$ to mean that X is a disjoint union of Y, Z. The fact that a subset E of edges induces a connected subgraph is written as $\mathit{connedges}(E) \equiv \forall X, Y \big(\neg\, \mathit{disjunion}(E; X, Y) \vee \exists F (\mathrm{star}(F) \wedge F \cap X \neq \emptyset \wedge F \cap Y \neq \emptyset)\big)$.

Let us now formulate the first part of the predicate $\mathit{vertexinset}$; saying that $v = u$ or there exists a path P with ends u and v: $\mathit{vertexinset}(F_v; X_U, F_u) \equiv$

$\big[F_u = F_v \vee \exists E_P \big[\, connedges(E_P) \wedge |F_u \cap E_P| = 1 \wedge |F_v \cap E_P| = 1 \wedge \forall F \big(\neg\, \mathrm{star}(F) \vee |F \cap E_P| = 0 \vee |F \cap E_P| = 2 \vee F = F_u \vee F = F_v\big) \wedge even\_intersection \,\big]\big]$. Now it remains to express by $even\_intersection$ that the set E_P has an even intersection with X_U. (That would be easy using a counting predicate, but we want to stay in plain MS_e.) We write $even\_intersection \equiv \exists L_1, L_2\, \exists e_u \big(\, disjunion(E_P \cap X_U; L_1, L_2) \wedge u \in E_P \cap F_u \wedge \forall e(e \notin L_1 \vee succonpath(e, L_2)) \wedge \forall e((e \notin L_2 \wedge e \neq e_u) \vee succonpath(e, L_1) \vee succonpath(e, F_v))\big)$, where $succonpath(e, L)$ means that the edge e is succeeded on the path P in the direction from u to v by an edge $f \in L$ among $L_1 \cup L_2 \cup F_v$ (it may possibly be $f = e$). It is $succonpath(e, L) \equiv \exists L'\, \exists f \big(L' \subseteq E_P \wedge connedges(L') \wedge e \in L' \wedge f \in L' \cap L \wedge \forall N \big(N \not\subseteq E_P \vee \neg\, connedges(N) \vee N \cap F_u = \emptyset \vee f \notin N \vee e \in N\big) \wedge \forall x \big(x = e \vee x = f \vee x \notin L_1 \cup L_2 \cup F_v \vee x \notin L'\big)\big)$. ∎

Theorem 8.2. *Let G be a connected multigraph on more than 2 vertices. Then any sentence about G in MS_2 is expressible in MS_e.*

Proof. Let ϕ be a closed MS_2 sentence. The task is to modify the formula ϕ in a way such that parts using vertex and vertex-set variables are equivalently expressed using certain new edge-set variables.

We formally rewrite ϕ to an MS_e sentence ϕ' in the following way:

- All edge and edge-set variables of ϕ are left untouched. All universal quantifiers in ϕ are expressed using existential quantifiers.
- Every equality predicate of vertex-set variables is equivalently replaced as

$$V = W \;\rightarrow\; \forall x (x \in V \leftrightarrow x \in W)\,.$$

- Each occurrence of a vertex variable v is replaced with a new edge-set variable F_v, meaning the star of edges centered at v. (Recall that $F_v = F_w$ implies $v = w$ in a connected multigraph on more than 2 vertices.) Specifically, a quantifier over v is rewritten as

$$\exists v (\ldots) \;\rightarrow\; \exists F_v \big[\, \mathrm{star}(F_v) \wedge (\ldots) \big]\,,$$

  and a vertex-edge incidence relation as

$$\mathrm{inc}(v, e) \;\rightarrow\; e \in F_v\,.$$

- Finally, it remains to replace vertex-set variables; a variable U is replaced simply with a pair X_U, Y_U of edge-set variables, with no additional restrictions. Namely, we rewrite

$$\exists U (\ldots) \;\rightarrow\; \exists X_U, Y_U (\ldots)\,,$$

  and then we change each $\in U$ relation to

$$x \in U \;\rightarrow\; vertexinset(F_x; X_U, Y_U)\,.$$

Let the formula resulting from ϕ by iterative application of the above rules be ϕ'. Now we argue why $G \models \phi \iff G \models \phi'$ over any connected graph G on more than 2 vertices. The first two rewriting steps clearly preserve the equivalence of models, and the third step does as well since edge stars uniquely express their central vertices. Possible question may arise about the fourth step – replacing a vertex-set variable U with an arbitrary pair X_U, Y_U. If $G \models \phi$ with a (partial) choice $U = U_0 \subseteq V(G)$ at $\exists U \ldots$, then, according to Lemma 8.1, $G \models \phi'$ would be satisfied with $X_U = X_{U_0}$ and $Y_U = F_u$ for any $u \in U_0$, or $Y_U = \emptyset$ if $U_0 = \emptyset$. If, on the other hand, $G \models \phi'$ with a choice $X_U = X_1, Y_U = Y_1 \subseteq E(G)$, then $G \models \phi$ with a choice $U = U_1$ at $\exists U \ldots$, where the set $U_1 = \{x \in V(G) : G \models \mathit{vertexinset}(F_x; X_1, Y_1)\}$. ∎

8.2 Interpreting CMS_2 in CMS_M

Recall that if M is the cycle matroid of a graph G, then a set $F \subseteq E(M)$ is independent if and only if the edges F induce no circuits in G. Notice that if a loopless multigraph G is not 3-connected, then its cycle matroid may not fully describe G. Similarly the positions of loops in a multigraph are not described by the matroid.

Lemma 8.3. *Let M be the cycle matroid of a loopless 3-connected multigraph G. Then any sentence about G in MS_2 can be expressed as a sentence about M in MS_M.*

Proof. Using Theorem 8.2, we only need to express the predicate star(F) in MS_M. A star of edges F in a 3-connected graph is described in the matroid language as a nonseparating cocircuit. (A cocircuit is a circuit in the dual matroid.) Since the complement of a cocircuit is a hyperplane (a maximal non-spanning set) from the definition, the complement $\bar{F} = E(G) - F$ is a hyperplane inducing a connected submatroid. We reformulate this back in the graph language for clarity: A set of edges $F \subseteq E(G)$ in a 3-connected graph G is a star if and only if, $G - F$ is a maximal subgraph containing no spanning tree of G, and each two edges of F are connected by a cycle in F, i.e. F induces a 2-connected subgraph of G.

Thus we may write star(F) in the language MS_M as $\exists X \big[\forall x (x \in F \leftrightarrow x \notin X) \wedge \mathit{hyperplane}(X) \wedge \mathit{connected}(X) \big]$, where $\mathit{hyperplane}(X) \equiv \forall B \big(B \not\subseteq X \vee \neg\, \mathit{basis}(B) \big) \wedge \forall Y \big(X = Y \vee X \not\subseteq Y \vee \exists B (B \subseteq Y \wedge \mathit{basis}(B)) \big)$, and $\mathit{connected}(X) \equiv \forall e, e' \big(e \notin X \vee e' \notin X \vee \exists C (C \subseteq X \wedge e, e' \in C \wedge \mathit{circuit}(C)) \big)$. ∎

In order to apply Lemma 8.3 to an arbitrary graph G, we have to make the graph 3-connected by adding extra vertices. Let $G \uplus H$ denotes the graph obtained from disjoint copies of G and H by adding all edges between them.

Lemma 8.4. *Let G be a loopless multigraph. Then, in the CMS_2 logic, any sentence ϕ about G can be equivalently written as a sentence ϕ' about $G \uplus K_3$. Moreover, ϕ' can be formulated such that there are no counting predicates $\mathrm{mod}_{p,q}(U)$ in ϕ' for a vertex-set variable U.*

Proof. We write $\phi' \equiv \exists a_1, a_2, a_3 \big(\mathit{apex}(a_1) \wedge \mathit{apex}(a_2) \wedge \mathit{apex}(a_3) \wedge \phi^o(a_1, a_2, a_3) \big)$ to "exclude" the three additional vertices a_1, a_2, a_3 from $G \uplus K_3$.

Here the predicate $apex(a) \equiv \forall v \big[v = a \vee \big[\exists e \left(\operatorname{inc}(v,e) \wedge \operatorname{inc}(a,e) \right) \wedge \forall e, f \big(e = f \vee \neg \operatorname{inc}(a,e) \vee \neg \operatorname{inc}(a,f) \vee \neg \operatorname{inc}(v,e) \vee \neg \operatorname{inc}(v,f) \big) \big] \big]$ says that the vertex a is adjacent to every other vertex by exactly one edge. (If there are more than three vertices x satisfying $apex(x)$ in $G \uplus K_3$, then all of them belong to the same orbit of the automorphism group, and hence we do not care which three of them we select.) The formula $\phi^o(a_1, a_2, a_3)$ is constructed from ϕ as follows.

- All universal quantifiers in ϕ are expressed using existential quantifiers.
- All occurrences of existential quantifiers are modified in the following ways: for vertices $\exists v\, \psi$ is replaced with $\exists v\, (v \neq a_1, a_2, a_3 \wedge \psi)$, similarly $\exists U\, \psi$ is replaced with $\exists U\, (a_1, a_2, a_3 \notin U \wedge \psi)$, and for edges $\exists e\, \psi$ is replaced with $\exists e \left[(\bigwedge_{i=1,2,3} \neg \operatorname{inc}(a_i, e)) \wedge \psi \right]$, similarly $\exists F\, \psi$ is replaced with $\exists F \left[\forall e \in F (\bigwedge_{i=1,2,3} \neg \operatorname{inc}(a_i, e)) \wedge \psi \right]$.
- Each occurrence of a counting predicate $\operatorname{mod}_{p,q}(U)$ on a vertex-set variable U is replaced with $\exists F \big[\operatorname{mod}_{p,q}(F) \wedge \forall e \big(e \notin F \vee \operatorname{inc}(a_1, e) \big) \wedge \forall e \exists v \big(e \notin F \vee v \in U \wedge \operatorname{inc}(v,e) \big) \wedge \forall v \exists e \big(v \notin U \vee e \in F \wedge \operatorname{inc}(v,e) \big) \big]$, which relates the vertex set U with the set of edges F joining the first apex vertex a_1 to the vertices of U, that means $|U| = |F|$. ∎

In fact, one may use a more involved construction with subdividing all edges of G in the above proof, thus handling also possible loops in G. However, the arguments would be too complicated to be presented here. We are now ready to finish this part.

Proof of Theorem 5.2. Let ϕ be a CMS_2 sentence over a graph G. We first apply Lemma 8.4 to rewrite ϕ as ϕ' over $G' = G \uplus K_3$ and without use of counting predicates of vertex sets. The we apply Lemma 8.3 to ϕ' and G', ignoring the edge-set counting predicates in ϕ'. Finally, we formally replace the original edge-set counting predicates of ϕ' with the corresponding CMS_M counting predicates. (These counting predicates literally stay the same as they were in ϕ'.) The resulting sentence ϕ'' over the matroid $M = M(G')$ is thus equivalent to ϕ over G; $G \models \phi \iff M \models \phi''$. ∎

Remark. If G is a loopless graph of tree-width t, then the graph $G' = G \uplus K_3$ has tree-width at most $t + 3$, and so branch-width at most $t + 4$ by [24]. Hence branch-width of the cycle matroid $M = M(G')$ of G' is at most $t + 4$ as well. The cycle matroids of graphs are representable over any field, as already noted above. Therefore Courcelle's Theorem 5.1 for (loopless) graphs follows from Theorem 6.1 and Theorem 5.2.

8.3 Matroids over Infinite Fields

After reading this paper, one may ask why we formulate Theorem 6.1 only for matroids represented over finite fields, especially when the MS_M theory speaks about abstract matroids. One reason for this has already been told above — there seems to be no natural definition of a "boundary sum" of abstract matroids, and so no natural way to derive parse trees for abstract matroids.

Though, one may still speculate that (s)he considers matroid parse trees represented over infinite fields, and that (s)he possibly expands each parse tree with more nodes in order to reduce the necessary composition operators down to a finite set. Would then it be possible to extend Theorem 6.1? We claim that it is not — another, and more important, reason to consider *only finite fields* is explained in the next example: We are going to show that a quite simple matroidal property generates an equivalence relation of an infinite index on the ≤2-boundaried $\mathbb{Q}$-represented matroids, and so this property cannot be recognized by a finite tree automaton due to Theorem 4.3.

We say that a matroid M is *identically self-dual* if, for each basis $B \subseteq E(M)$ of M, the set $E(M) - B$ is also a basis of M. It is easy to describe this property in MS_M logic. We write $\mathit{id\_selfdual} \equiv \forall X\, \exists Y\, \big(\forall e(e \in X \leftrightarrow x \notin Y) \wedge (\mathit{basis}(X) \rightarrow \mathit{basis}(Y))\big)$. Recall also, for a matroid family $\mathcal{M}$, the equivalence relation $\approx_{\mathcal{M},t}$ on $\bar{\mathcal{B}}_t$ from Section 4; considering now the rational field $\mathbb{Q}$.

The class of matroids called "spikes" more or less explicitly appears in several papers in matroid structure theory, for example [12]. For simplicity, we restrict our attention to $\mathbb{Q}$-represented spikes. Let $\boldsymbol{D}(x_1, \ldots, x_n) = [d_{i,j}]_{i=1}^n$ denote an $n \times n$ matrix such that $n \geq 3$, $d_{i,j} = 1$ if $i \neq j \in [1, n]$, and $d_{i,i} = x_i$ for $i \in [1, n]$. Each *rank-n spike S representable over* $\mathbb{Q}$ is represented by the matrix $\boldsymbol{A} = \big[\boldsymbol{I}_n \,|\, \boldsymbol{D}(x_1, \ldots, x_n)\big]$ for $x_1, \ldots, x_n \in \mathbb{Q} - \{1\}$. Denote the elements of S by $e_1, \ldots, e_n, f_1, \ldots, f_n$ in order corresponding to the columns of $\boldsymbol{A}$. Then the rows of $\boldsymbol{D}(x_1, \ldots, x_n)$ naturally correspond to the elements $e_1, \ldots, e_n$.

Lemma 8.5. *Let $\mathcal{S}$ be the set of all $\mathbb{Q}$-represented identically self-dual spikes. Then the equivalence relation $\approx_{\mathcal{S},3}$ has an infinite index over $\bar{\mathcal{B}}_3$.*

Proof. We may easily construct a branch-decomposition (T, τ) of any spike S: Let T' be an arbitrary cubic tree with n leaves, and let T be obtained from T' by adding two new leaves l'_i, l''_i, $i \in [1, n]$ to each leaf l_i of T'. Let $\tau(e_i) = l'_i$ and $\tau(f_i) = l''_i$. For any subset $K \subseteq [1, n]$ and $F = \bigcup_{i \in K} \{e_i, f_i\}$, the connectivity of F is $\lambda_S(F) \leq 3$, and so the width of (T, τ) is 3. Hence $\mathcal{S} \subset \mathcal{B}_3$.

Let us look at the determinant of a matrix $\boldsymbol{D}_k = \boldsymbol{D}(y_1, \ldots, y_k)$. If more than one values $y_i = 1$ in $\boldsymbol{D}_k$, then $|\boldsymbol{D}_k| = 0$. If exactly one value $y_i = 1$ in $\boldsymbol{D}_k$, then $|\boldsymbol{D}_k| \neq 0$. Otherwise, when $y_i \neq 1$ for all $i \in [1, k]$,

$$(*) \qquad |\boldsymbol{D}(y_1, \ldots, y_k)| = \prod_{i=1}^{k} (y_i - 1) \cdot \left(1 + \sum_{i=1}^{k} \frac{1}{y_i - 1}\right) .$$

Suppose that $S = S(x_1, \ldots, x_n)$ denote a $\mathbb{Q}$-represented spike on the ground set $E = E(S) = \{e_1, \ldots, e_n, f_1, \ldots, f_n\}$, and that $X \subseteq E$. Then X is a basis of S if and only if the subdeterminant selected by the rows $\{e_1, \ldots, e_n\} - X$ and the columns $\{f_1, \ldots, f_n\} \cap X$ in the matrix $\boldsymbol{D}(x_1, \ldots, x_n)$ is nonzero.

So, using (*), we conclude the following: If there are $i, i' \in [1, n]$ such that $e_i, f_i \in X$, $\;e_{i'}, f_{i'} \notin X$, and that $|\{e_j, f_j\} \cap X| = 1$ for $j \neq i, i'$, then X is a basis of S. If $|\{e_j, f_j\} \cap X| = 1$ for all $j \in [1, n]$, then X is a basis of S iff $|\boldsymbol{D}(x_j : j \in [1, n] \text{ s.t. } f_j \in X)| \neq 0$. Otherwise, X is not a basis of S. Therefore,

S is identically self-dual if and only if, for each $K \subseteq [1,n]$, $|\boldsymbol{D}(x_j : j \in K)| = 0 \iff |\boldsymbol{D}(x_j : j \in [1,n] - K)| = 0$. (Here we declare $|\boldsymbol{D}(\emptyset)| = 1$.)

Consider now even $n \geq 8$, and an arbitrary choice $(x_1, \ldots, x_n) \in \{-1, 3\}^n$. Then $S = S(x_1, \ldots, x_n)$ is identically self-dual if and only if at most one value in the sequence $(x_1, \ldots, x_n)$ is -1, or if $\sum_{i=1}^{n} \frac{1}{x_i - 1} = -2$ which means that exactly $\frac{n}{2} - 2$ values in the sequence $(x_1, \ldots, x_n)$ equal -1.

Let $\bar{S}_1, \bar{S}_2$ be the $\leq$2-boundaried matroids with internal elements $\{e_i, f_i : i \in [1, \frac{n}{2}]\}$ and $\{e_i, f_i : i \in [\frac{n}{2} + 1, n]\}$, respectively, such that $\bar{S}_1 \bar{\oplus} \bar{S}_2 = S$. Clearly, $\bar{S}_1, \bar{S}_2 \in \bar{\mathcal{B}}_3$, and the matroid $\bar{S}_1 = \bar{S}_1(x_1, \ldots, x_{\frac{n}{2}})$ depends only on the selection of $x_1, \ldots, x_{\frac{n}{2}}$. If $(x_1, \ldots, x_{\frac{n}{2}})$ and $(x'_1, \ldots, x'_{\frac{n}{2}})$ are two sequences from $\{-1, 3\}^{\frac{n}{2}}$ which differ in their numbers of -1 values, then $\bar{S}_1(x_1, \ldots, x_{\frac{n}{2}}) \not\approx_{\mathcal{S},3} \bar{S}_1(x'_1, \ldots, x'_{\frac{n}{2}})$. Hence the index of the equivalence relation $\approx_{\mathcal{S},3}$ is at least $\frac{n}{2} - 1$ for each n. Altogether, the index of $\approx_{\mathcal{S},3}$ on the whole $\bar{\mathcal{B}}_3$ must be infinite over the rationals $\mathbb{Q}$. ∎

Notice that 3 is the smallest interesting value of branch-width in this context since a matroid of branch-width 2 is trivially representable over any field. In conjunction with Theorem 4.3 we get:

Corollary 8.6. *There is a class $\mathcal{S}$ of matroids of branch-width 3 represented over $\mathbb{Q}$ by matrices with entries from $\{-1, 1, 3\}$, such that $\mathcal{S}$ can be characterized by an MS_M sentence, but $\mathcal{S}$ cannot be recognized by a finite tree automaton.*

8.4 Related "Width" Parameters

In the context of our research, it is interesting to mention another graph "with" parameter called clique-width [8]: A graph has *clique-width* $\leq k$ if it can be constructed using k labels and the following four operations: 1) create a new vertex with label i; 2) take the disjoint union of several labeled graphs; 3) add all edges between vertices of label i and label j; and 4) relabel all vertices with label i to have label j. An expression defining a graph G built from the above four operations using k labels is a *k-expression* for G.

Clique-width generalizes tree-width or branch-width of graphs in the sense that a graph class of bounded tree-width has also bounded clique-width, but the converse is not true. (For example, complete graphs have clique-width 2.) Analogously to Theorem 5.1, any graph class definable in MS_1 is efficiently recognizable over graphs of bounded clique-width [9]. However, until [25], algorithms running on graphs of bounded clique-width needed a corresponding k-expression on the input. The first (and currently the only known) efficient approximation [25, 26] of an expression for a graph of bounded clique-width is computed using the further notion of a rank-width.

Instead of giving the full definition of rank-width [25] (which is defined in similar terms as branch-width), we mention that the rank-width of a bipartite graph equals the branch-width of a certain binary matroid — a rank-decomposition of a bipartite graph G is exactly the same as a branch-decomposition of the matroid represented by the bipartite adjacency matrix of G over $GF(2)$. Actually, the current asymptotically fastest algorithm for computing rank-width of a (general) graph [26] uses this correspondence and the matroid branch-width algorithm of [19]. This important correspondence between graph clique-width / rank-width and matroid branch-width, and between the related MS theories, have also recently found applications in the research of decidable MS_1 theories of graphs by Courcelle and Oum [10] (see also [20]).

Our theory also allows to define a notion of a matrix "width" that is invariant on line-scaling and pivoting of the matrix, or, in other words, invariant on the projective equivalence of point configurations. The hope is that matrices of small "width" are much easier to handle than general matrices, and that fast algorithms exist for problems involving these matrices.

The *branch-width* of a matrix $\boldsymbol{A}$ over $\mathbb{F}$ is the branch-width of the matroid $M(\boldsymbol{A})$. Hence the matrix branch-width of $\boldsymbol{A}$ is not changed when row operations are applied to $\boldsymbol{A}$, and so it is a robust measure of a "complexity" of $\boldsymbol{A}$. We also remark that our $\mathbb{F}$-represented matroids are in a one-to-one correspondence with linear codes over $\mathbb{F}$ since a projective equivalence of point configurations coincides with the standard equivalence of linear codes. Thus we may define a branch-width of a linear code C as the branch-width of the generator matrix of the code C.

Let us lastly mention that some authors deal with another matrix "width" parameter defined as follows. For a matrix $\boldsymbol{A} = [a_{i,j}]_{i,j=1}^{n}$, let G_A be the graph on the vertex set $\{1, \dots, n\}$ and the edge set consisting of all $\{i, j\}$ such that $a_{i,j} \neq 0$ or $a_{j,i} \neq 0$. The tree-width of the matrix $\boldsymbol{A}$ is given by the tree-width of the graph G_A. This definition was, perhaps, inspired by Choleski factorization of sparse symmetric matrices which is related to graph tree-width of the matrix. (See [3] for more details.) However, this notion of a matrix tree-width is not robust in the above sense — applying a row operation to a matrix $\boldsymbol{A}$ may dramatically change the tree-width of G_A, while the vector configuration represented by $\boldsymbol{A}$ is still the same. That is why we think that the tree-width of G_A is not a good measure of a "complexity" of the matrix / vector configuration $\boldsymbol{A}$.

The two above defined matrix "width" parameters are not related to each other. Look at the following example of a matrix $\boldsymbol{D} = \boldsymbol{J}_n - \boldsymbol{I}_n$: The graph G_D defined by this matrix is a clique, and so it has tree-width $n - 1$. On the other hand, the matroid $M([\boldsymbol{I}|\boldsymbol{D}])$, a spike, is 3-connected of branch-width 3.

Acknowledgments

The author acknowledges valuable discussions with Rod Downey and Detlef Seese on the topics of parametrized complexity, tree automata, and of MSO logic, suggestions from Geoff Whittle on matroid representations, and a suggestion on Lemma 8.4 from Jim Geelen. The author is also grateful to the NZ Marsden Fund (a grant to Geoff Whittle) and the Victoria University of Wellington for supporting his wonderful stay in New Zealand in 2000–2002, during which the major results originated. Ongoing research is now being supported by Czech research grant GAČR 201/05/050.

References

1. K.A. Abrahamson, M.R. Fellows, *Finite Automata, Bounded Treewidth, and Well-Quasiordering*, In: Graph Structure Theory, Contemporary Mathematics 147, American Mathematical Society (1993), 539–564.
2. S. Arnborg, J. Lagergren, D. Seese, *Problems easy for Tree-decomposible Graphs (extended abstract)*, Proc. 15th Colloq. Automata, Languages and Programming, Lecture Notes in Computer Science 317, Springer-Verlag (1988), 38–51.
3. H.L. Bodlaender, *A Tourist Guide through Treewidth*, Acta Cybernetica 11 (1993), 1–21.
4. H.L. Bodlaender, D.M. Thilikos, *Constructive Linear Time Algorithms for Branch-width*, Proceedings 24th ICALP, Lecture Notes in Computer Science 1256, 1997, 627–637.
5. R.B. Borie, R.G. Parker, C.A. Tovey, *Automatic generation of linear-time algorithms from predicate calculus descriptions of problems on recursively constructed graph families*, Algorithmica 7 (1992), 555–582.
6. B. Courcelle, *On context-free sets of graphs and their monadic second-order theory*, In: Graph-Grammars and Their Application to Computer Science, Proc. Int. Wkshop. on Graph Grammars 1986, Lecture Notes in Computer Science 291 (1987), 133–146.
7. B. Courcelle, *The Monadic Second-Order Logic of Graphs I. Recognizable sets of Finite Graphs* Information and Computation 85 (1990), 12–75.
8. B. Courcelle, S. Olariu, *Upper bounds to the clique width of graphs*, Discrete Appl. Math. 101 (2000), 77–114.
9. B. Courcelle, J.A. Makowsky, U. Rotics, *Linear Time Solvable Optimization Problems on Graphs of Bounded Clique-Width*, Theory Comput. Systems 33 (2000), 125–150.
10. B. Courcelle, Sang-Il Oum, *Vertex-minors, monadic second-order logic, and a conjecture by Seese*, submitted, 2005.
11. R.G. Downey, M.R. Fellows, Parametrized Complexity, Springer-Verlag New York, 1999.
12. J.F. Geelen, A.H.M. Gerards, G.P. Whittle, *Branch-Width and Well-Quasi-Ordering in Matroids and Graphs*, J. Combin. Theory Ser. B 84 (2002), 270–290.
13. J.F. Geelen, A.H.M. Gerards, N. Robertson, G.P. Whittle, *On the Excluded Minors for the Matroids of Branch-Width k*, J. Combin. Theory Ser. B 88 (2003), 261–265.
14. J.F. Geelen, A.H.M. Gerards, G.P. Whittle, *Excluding a Planar Graph from a $GF(q)$-Representable Matroid*, manuscript, 2003.
15. P. Hliněný, *It is Hard to Recognize Free Spikes*, manuscript, 2002.
16. P. Hliněný, *Branch-Width, Parse Trees, and MSO Logic for Matroids (Extended Abstract)*, In: Theoretical Aspects of Computer Science STACS 2003, Lecture Notes in Computer Science 2607, Springer Verlag Berlin (2003), 319–330.
17. P. Hliněný, *On Matroid Properties Definable in the MSO Logic*, In: Math Foundations of Computer Science MFCS 2003, Lecture Notes in Computer Science 2747, Springer Verlag Berlin (2003), 470–479.
18. P. Hliněný, *The Tutte Polynomial for Matroids of Bounded Branch-Width*, Combinatorics, Probability and Computing, to appear (2005).
19. P. Hliněný, *A Parametrized Algorithm for Matroid Branch-Width*, SIAM J. Computing, to appear (2005).
20. P. Hliněný, D. Seese, *Trees, Grids, and MSO Decidability: from Graphs to Matroids*, Theoretical Computer Sci., to appear (2005).

Extended abstract in: IWPEC 2004, Lecture Notes in Computer Science 3162, Springer Verlag Berlin (2004), 96–107.
21. P. Hliněný, G.P. Whittle, *Matroid Tree-Width*, submitted.
Ext. abstract in: Eurocomb'03, ITI Series 2003–145, Charles University 2003.
22. J.G. Oxley, Matroid Theory, Oxford University Press, 1992.
23. N. Robertson, P.D. Seymour, *Graph Minors – A Survey*, Surveys in Combinatorics, Cambridge Univ. Press 1985, 153–171.
24. N. Robertson, P.D. Seymour, *Graph Minors X. Obstructions to Tree-Decomposition*, J. Combin. Theory Ser. B 52 (1991), 153–190.
25. Sang-Il Oum, P.D. Seymour, *Approximation Algorithm to the Clique-Width of a Graph*, submitted, 2004.
26. Sang-Il Oum, *Approximating Rank-width and Clique-width Quickly*, manuscript, 2005.
27. D. Seese, *The Structure of Models of Decidable Monadic Theories of Graphs*, Ann. Pure and Appl. Logic 53 (1991), 169–195.