Maximum number of colors in hypertrees of bounded degree

Csilla Bujtás [1] Zsolt Tuza [1,2]

[1] Department of Computer Science and Systems Technology
University of Pannonia, Veszprém, Hungary

[2] Alfréd Rényi Institute of Mathematics
Hungarian Academy of Sciences, Budapest, Hungary

Latest update on 2014–5–23

Abstract

The upper chromatic number $\overline{\chi}(\mathcal{H})$ of a hypergraph $\mathcal{H} = (X, \mathcal{E})$ is the maximum number of colors that can occur in a vertex coloring $\varphi : X \to \mathbb{N}$ such that no edge $E \in \mathcal{E}$ is completely multicolored. A hypertree (also called arboreal hypergraph) is a hypergraph whose edges induce subtrees on a fixed tree graph. It has been shown that on hypertrees it is algorithmically hard not only to determine exactly but also to approximate the value of $\overline{\chi}$, unless $\mathsf{P} = \mathsf{NP}$. In sharp contrast to this, here we prove that if the input is restricted to hypertrees $\mathcal{H}$ of bounded maximum vertex degree, then $\overline{\chi}(\mathcal{H})$ can be determined in linear time if an underlying tree is also given in the input. Consequently, $\overline{\chi}$ on hypertrees is fixed parameter tractable in terms of maximum degree.

Keywords: hypergraph, hypertree, arboreal hypergraph, vertex coloring, C-coloring, upper chromatic number.

AMS 2000 Subject Classification: 05C15, 05C65, 05C85

1 Introduction

In this note our subject is a coloring parameter of hypergraphs which was introduced in several equivalent versions during the last four decades.

Formally, a *hypergraph* is a pair $\mathcal{H} = (X, \mathcal{E})$ where X is a set called *vertex set* and $\mathcal{E}$ is a set system over X, called *edge set.* Throughout, we will assume that $|E| \geq 2$ holds for each $E \in \mathcal{E}$, moreover we use the notations $n = |X|$ and $m = |\mathcal{E}|$ for the number of vertices and edges, respectively.

Definition 1 A *C-coloring* of a hypergraph $\mathcal{H} = (X, \mathcal{E})$ is an assignment of colors to the vertices, $\varphi : X \to \mathbb{N}$, such that each edge $E \in \mathcal{E}$ contains at least two vertices of the same color. The *upper chromatic number* $\overline{\chi}(\mathcal{H})$ of $\mathcal{H}$ is the maximum number of colors that can occur in a C-coloring of $\mathcal{H}$. We say that a C-coloring $\varphi : X \to \mathbb{N}$ is an *optimal C-coloring* if it uses exactly $\overline{\chi}(\mathcal{H})$ distinct colors.

1.1 Some related notions and facts

We will use the following standard terminology:

- A graph $G = (V, E)$ is a *host graph* of the hypergraph $\mathcal{H} = (X, \mathcal{E})$ if $V = X$ and each edge of $\mathcal{H}$ induces a connected subgraph in G. A *host tree* is a host graph which is a tree. A hypergraph admitting a host tree is called *hypertree*[1].

- A vertex set $S \subseteq X$ is a *transversal*[2] of $\mathcal{H} = (X, \mathcal{E})$ if it intersects every edge of $\mathcal{H}$. The minimum cardinality of a transversal of $\mathcal{H}$ is the *transversal number* and denoted by $\tau(\mathcal{H})$. The complement of a transversal is a set $S' \subseteq X$, which contains no edge of $\mathcal{H}$ entirely. Such a vertex set S' is called *independent set*[3] and the maximum cardinality $\alpha(\mathcal{H})$ of an independent set is the *independence number* of $\mathcal{H}$.

- The *decrement* of $\mathcal{H} = (X, \mathcal{E})$ is defined as $dec(\mathcal{H}) = |X| - \overline{\chi}(\mathcal{H})$, and similarly the decrement of a C-coloring $\varphi : X \to \mathbb{N}$ is $dec(\varphi) = |X| - |\varphi(X)|$ (introduced in [2]).

Some simple relations satisfied by every hypergraph $\mathcal{H} = (X, \mathcal{E})$ are:

[1] same as 'arboreal hypergraph' in part of the literature

[2] same as 'hitting set' or 'vertex cover'

[3] also called 'stable set', but some papers use the two terms differently for hypergraphs

- The Gallai-type equality $\alpha(\mathcal{H}) + \tau(\mathcal{H}) = |X|$ follows directly from the definitions.
- Selecting one vertex from each color class of a C-coloring is independent, therefore $\overline{\chi}(\mathcal{H}) \leq \alpha(\mathcal{H})$. Equivalently, $dec(\mathcal{H}) \geq \tau(\mathcal{H})$ always holds.
- The upper chromatic number is additive with respect to vertex-disjoint union, hence the optimal C-colorings of a disconnected hypergraph are composed of those of its connected componens.

Throughout this paper, the host tree of a hypertree $\mathcal{H}$ is denoted by T, moreover T is considered with an arbitrarily fixed root r. To avoid misunderstanding, we refer to the edges of T as *lines* while we keep the term *edge* for the edges of $\mathcal{H}$. Additionally, we use the following notation for each vertex x_i of T:

- T_i – the subtree of T rooted at x_i, induced by those vertices x_j for which the x_j–r path contains x_i (also including $j = i$),
- X_i – the vertex set of T_i,
- $\mathcal{H}_i$ – the subhypergraph of $\mathcal{H}$ induced by X_i, consisting of those edges of $\mathcal{H}$ which are entirely contained in X_i.

1.2 Earlier results

C-coloring of hypergraphs was proposed by Berge and first investigated in detail by Sterboul [9] in the early 1970s. The parameter equal to $\overline{\chi}(\mathcal{H}) + 1$ is called the 'cochromatic number' and the 'heterochromatic number' of $\mathcal{H}$ due to Berge [3, p. 151] and Arocha *et al.* [1], respectively. For more about history, relations to other important notions, and for earlier results on C-coloring we refer to the survey [4].

Another track of related research started in 1993 when Voloshin [10, 11] introduced the concept of *mixed hypergraphs*. This combines two opposite types of coloring conditions making difference between two families $\mathcal{D}$ and $\mathcal{C}$ of edges. Then, in a proper vertex coloring, each D-edge must have at least two vertices with different colors whilst each C-edge must have at least two vertices with a common color. This leads to two fundamental parameters instead of one: the minimum and the maximum number of colors, which are called the lower and the upper chromatic number of a mixed hypergraph, respectively. Clearly, a proper coloring of a mixed hypergraph which has only C-edges is just a C-coloring in our terminology.

For the determination of the upper chromatic number of mixed hypergraphs, there is no polynomial-time $o(n)$-approximation algorithm in general [6] and there is no polynomial-time approximation scheme (PTAS) even if the maximum degree is fixed to be 2, unless $\mathsf{P} = \mathsf{NP}$ [7]. On the other hand, for mixed hypergraphs of maximum degree 2 the upper chromatic number has a linear-time $\frac{5}{3}$-approximation and a $O(m^3 + n)$-time $\frac{3}{2}$-approximation [7]. All the results cited above assume the presence of D-edges in the input mixed hypergraph.

More related to our current subject, approximability results for C-coloring of hypergraphs were obtained in [5]. As proved there, $dec(\mathcal{H})$ is $(2+2\ln 2m)$-approximable on the class of all hypergraphs, and it is $(1+\ln m)$-approximable on the class of hypertrees. On the other hand, assuming that $\mathsf{NP} \not\subseteq \mathsf{DTIME}$ $(n^{O(log\ log\ n)})$, for every $\epsilon > 0$ one cannot approximate $dec(\mathcal{H})$ within $(1 - \varepsilon) \ln m$, neither $\overline{\chi}(\mathcal{H})$ within $O(n^{1-\varepsilon})$ in polynomial time. Both negative results are valid already on the class of hypertrees and the latter one also holds if all the edges are of cardinality 3. Further, on the class of hypertrees in which every edge is of size at most 7, neither $\overline{\chi}(\mathcal{H})$ nor $\alpha(\mathcal{H}) - \overline{\chi}(\mathcal{H})$ can be approximated within additive error $o(n)$, unless $\mathsf{P} = \mathsf{NP}$.

The main result of our paper is that the determination of $\overline{\chi}(\mathcal{H})$ becomes tractable on hypertrees if some (any) upper bound is put on the vertex degrees. It may be of interest to note that it is not necessary to assume any bounds on the size of edges; an efficient algorithm can be designed without this kind of restriction.

Theorem 1 *For every $d \in \mathbb{N}$, there exists a polynomial-time algorithm that solves the following optimization and search problems:*

> *Given an input hypertree $\mathcal{H}$ of maximum vertex degree at most d, determine the upper chromatic number $\overline{\chi}(\mathcal{H})$ and find a C-coloring of $\mathcal{H}$ with exactly $\overline{\chi}(\mathcal{H})$ colors.*

Moreover, if a host tree T of maximum degree at most $\binom{d}{\lfloor d/2 \rfloor} + 1$ is also given in the input, then the running time of the algorithm is linear.

This theorem will be proved in Section 3. In the next Section 2 we describe some general properties of C-colorings of hypertrees, which will be important in proving the main result. Among them, we should invite attention to our All-Optimality Lemma (Proposition 4) which states that every rooted hypertree admits a C-coloring in which every rooted subhypertree accommodates the maximum possible number of colors and at the same time it satisfies a useful connectivity condition, too.

2 The structure of optimal C-coloring of hypertrees

The observations presented in this section are simple but will be very useful, and also are interesting on their own right. They can be applied to reduce the number of candidates for optimal C-colorings in hypertrees substantially, and will simplify several arguments later on. We introduce two notions; the first of them makes sense not only for hypertrees but also for any hypergraph.

Definition 2 *(Connected C-coloring)* Let $\mathcal{H}$ be a hypergraph over a host graph G. A C-coloring φ of $\mathcal{H}$ is *connected* if each color class induces a connected subgraph of G.

Definition 3 *(All-optimal C-coloring)* Let $\mathcal{H} = (X, \mathcal{E})$ be a hypertree over a rooted host tree T. Then, a C-coloring φ of $\mathcal{H}$ is *all-optimal* if for each $x_i \in X$, the set X_i (of vertices in the rooted subtree T_i) contains $\overline{\chi}(\mathcal{H}_i)$ distinct colors.

Note that every all-optimal C-coloring uses precisely $\overline{\chi}(\mathcal{H})$ colors.

The next assertion states that any C-coloring of a hypertree can be transformed to a connected C-coloring, without decreasing the number of colors. (This is false in general if cycles are allowed in the host graph.)

Proposition 2 *If a hypertree $\mathcal{H}$ admits a C-coloring with k colors, then it also has a connected C-coloring with k colors, over any fixed host tree T.*

Proof We view the host tree T of $\mathcal{H}$ as a rooted tree, and consider any $\overline{\chi}$-coloring of $\mathcal{H}$. Traverse T in preorder and let a be the first vertex whose color, say α occurred earlier, but not on its parent b; i.e., vertex b separates some connected components of this color class. We denote by A the set of vertices in the subtree rooted at a, by β the color of b, and set $B = X \setminus A$.

Inside A, we recolor all occurrences of color α with β. This operation does not decrease the number of colors as there exists a vertex outside A which is colored with α. We claim that all edges of $\mathcal{H}$ remain properly colored. Indeed, if an edge is inside B, its colors have not been changed. Inside A, only α has been switched to β, therefore any monochromatic vertex pair remains monochromatic after recoloring. Finally, if an edge meets both A and B, then it contains the monochromatic pair $\{a, b\}$.

Continuing the process, after every recoloring, vertices with the properties of a occur later and later, until all color classes become connected in the host tree, and a required $\overline{\chi}$-coloring is found. □

In calculating the coloring invariants of hypergraphs, connected C-colorings have some advantage, as one can easily observe:

Proposition 3 *If φ is a connected C-coloring of hypergraph $\mathcal{H}$ over a fixed host graph G, then*

- (i) *the number of colors in φ is equal to the number of monochromatic components in G;*
- (ii) *if $\mathcal{H}$ is a hypertree, then the decrement of φ is the number of monochromatic lines in the host tree.*

The observation above makes it possible to switch our view from vertex coloring to a selection of monochromatic lines. This leads to an argument strengthening Proposition 2 as follows.

Proposition 4 **(All-Optimality Lemma)** *Every hypertree $\mathcal{H}$, over any host tree T, admits a C-coloring which is connected and all-optimal.*

Proof Due to Proposition 3(ii), to obtain a connected and optimal C-coloring, we have to select a set F of lines in the host tree such that each edge $E \in \mathcal{E}$ contains at least one line $f \in F$, moreover, $|F|$ is as small as possible. In analogy to optimal C-colorings, we call F *optimal* if it consists of the monochromatic lines of an optimal connected C-coloring of $\mathcal{H}$.

Having fixed a preorder traversal $x_n, x_{n-1}, \ldots, x_1$ of the host tree, let F be optimal with the additional property that the index i is as small as possible such that the restriction of F to each of $T_n, T_{n-1}, \ldots, T_i$ is optimal. Clearly, i is well-defined because $i \leq n$ holds for any optimal F. If $i = 1$ holds, then the proof is done as the corresponding C-coloring of $\mathcal{H}$ is all-optimal. Suppose for a contradiction that $i > 1$. Then, by assumption, F is not optimal in T_{i-1}, i.e. an optimal F_{i-1} for T_{i-1} has fewer lines than F has inside T_{i-1}. Let F' be the set of lines obtained from F by removing $F \cap E(T_{i-1})$, inserting all lines of F_{i-1}, and also inserting the line that joins x_{i-1} to its parent. This modification does not increase the number of lines in any T_j with $i - 1 \leq j \leq n$; in particular, $|F'| \leq |F|$ holds and F' is optimal in each of those T_j. This will contradict the choice of F (by the minimality of i) once we show that F' determines a C-coloring of $\mathcal{H}$.

Let $E_k \in \mathcal{E}$ be any edge of $\mathcal{H}$. If E_k is disjoint from T_{i-1}, then its intersection with F' is the same as that with F, hence it is C-colored. If E_k is entirely contained in T_{i-1}, then it is C-colored by F_{i-1} which is part of F'. Finally, if neither of those two situations occur, then E_k contains both x_{i-1} and its parent, hence is C-colored by the newly inserted line that joins those two vertices of the host tree. □

Although the argument above is non-constructive, we shall see in the proof of the main result that a connected all-optimal C-coloring can be found by a polynomial-time algorithm, too, whenever the maximum degree of input hypergraphs is bounded above. Observe, on the other hand, that all-optimality is guaranteed for every hypertree, with no assumption on vertex degrees.

3 Linear-time algorithm for bounded degree and edge size

The goal of this section is to prove the main result of the paper. Before designing an efficient algorithm, we first show how the too high degrees of a host tree can be decreased to respect a well-defined upper bound. The essential part of the coloring algorithm will then be described in the second subsection.

3.1 Host tree construction and degree reduction

As a preparation for the proof of Theorem 1, here we prove that hypertrees of bounded degree always admit host trees of bounded degree, and the latter can also be found efficiently. We begin with noting that *some* host tree can be determined in polynomial time.

Lemma 5 *If $\mathcal{H}$ is a hypertree, then a host tree of $\mathcal{H}$ can be constructed in polynomial time.*

Proof We apply the following lemma from pages 187–188 of [3]:

($\star$) If $\mathcal{H} = (X, \mathcal{E})$ is a hypertree (without 1-element edges), then there exists a vertex x such that all the edges of $\mathcal{H}$ containing x have a common vertex $y \neq x$.

For a hypertree $\mathcal{T}_i = (Y_i, \mathcal{E}_i)$ and for a vertex $x \in Y_i$, let $\mathcal{E}_i(x)$ denote the set of all edges containing x. Now, we set $\mathcal{T}_1 = \mathcal{H}$, $Y_1 = X$, $\mathcal{E}_1 = \mathcal{E}$ and consider the algorithm whose ith step (for $1 \leq i \leq n-1$) is determined as follows.

- Select a vertex pair $x_i, y_i \in Y_i$ for which $\mathcal{E}_i(x_i) \subseteq \mathcal{E}_i(y_i)$ holds and there is no vertex $z \in Y_i$ with $\mathcal{E}_i(z) \subsetneq \mathcal{E}_i(x_i)$. Further, we define $p(x_i) = y_i$, which means that y_i will be the parent of x_i in the rooted host tree under construction.

- If $i < n-1$, delete x_i from each edge of $\mathcal{T}_i$ and delete the possibly arising 1-element edges; that is, $\mathcal{T}_{i+1} = (Y_{i+1}, \mathcal{E}_{i+1})$ where

$$Y_{i+1} = Y_i \setminus \{x_i\} \quad and \quad \mathcal{E}_{i+1} = \{E \setminus \{x_i\} : E \in \mathcal{E}_i \ \wedge \ |E \setminus \{x_i\}| > 1\}.$$

Due to ($\star$), if $\mathcal{T}_i$ is a hypertree, such a vertex pair (x_i, y_i) exists. Moreover, if $\mathcal{T}_i$ is a hypertree with a host tree T, then deleting x_i from T and joining its children to y_i, a host tree of $\mathcal{T}_{i+1}$ is obtained. Thus, $\mathcal{T}_{i+1}$ is a hypertree as well. Therefore, the procedure can be executed and the graph on the vertex set X and with the edge set $\{x_i p(x_i) : 1 \leq i \leq n-1\})$ is a host tree of $\mathcal{H}$. □

We now turn to degree reduction.

Lemma 6 *If a hypertree $\mathcal{H}$ has maximum vertex degree at most d, then there exists a rooted host tree of $\mathcal{H}$ in which every vertex has at most $\binom{d}{\lfloor d/2 \rfloor}$ children. Moreover, from any host tree T of $\mathcal{H}$, another host tree with such small degrees can be constructed in $O(n^2)$ time.*

Proof Consider a hypertree $\mathcal{H} = (X, \mathcal{E})$ with $\Delta(\mathcal{H}) \leq d$, and let T be any rooted host tree of $\mathcal{H}$. Suppose that there is a vertex x with more than $D = \binom{d}{\lfloor d/2 \rfloor}$ children in T. We assume that, among all such vertices, x is nearest to the root. We denote by $x_1, \ldots, x_k$ the children of x, and consider $x_1, \ldots, x_{D+1}$.

By assumption, the set $\mathcal{E}_x \subseteq \mathcal{E}$ of all edges involving x contains not more than d elements. For each i in the range $1 \leq i \leq D+1$ we set $\mathcal{E}_i = \{E \in \mathcal{E}_x \mid x_i \in E\}$. This can be viewed as a set system over at most d elements, namely over the members of $\mathcal{E}_x$. Hence, by Sperner's theorem [8], there exist two children x_i and x_j of x such that $\mathcal{E}_j \subseteq \mathcal{E}_i$. In this situation, we replace the line xx_j with $x_i x_j$ in T.

We are going to prove that the obtained T' is also a host tree of $\mathcal{H}$. Let $E \in \mathcal{E}$ be an arbitrary edge of $\mathcal{H}$. Since T is a host tree, the subgraph induced by E in T is connected. We show that E induces a subtree in T', too.

- If $x_j \notin E$ then the subtree induced by E remains unchanged.
- If $x_j \in E$ but $x \notin E$ then E is contained entirely in the subtree of T rooted in x_j, hence it induces a connected subgraph also in T'.
- If $x_j \in E$ and $x \in E$ then $x_i \in E$ also holds and the path xx_ix_j can replace the omitted line xx_j in T'. Hence the subgraph induced by E remains connected in T'.

Thus, the degree of x is decreased by one, and the modification cannot cause any increase in degrees of vertices not farther from the root than x. Repeatedly applying this procedure, the vertices having more than $\binom{d}{\lfloor d/2 \rfloor}$ children will be at greater and greater distance apart from the root. After finite number of steps we obtain a host tree T^* satisfying the requirements.

Since we restrict our attention to the first $D+1$ children of x, for fixed d a reducible pair (i, j) in the neighborhood of x can be detected in constant time. Define now the function

$$f(T) := \sum_{e \in E^*(T)} (n - d_T(r, e))$$

where n is the number of vertices, $E^*(T)$ is the set of lines rooted at vertices of degree larger than D in the current host tree, and $d_T(r, e)$ denotes the distance of line e from the root r. Originally we have $f(T) < n^2$, the reduction decreases the value of $f(T)$ in each step, and we have nothing to do once we reach $f(T) = 0$. Thus, the entire procedure runs in time $O(n^2)$. In fact it may take $\Theta(n^2)$, e.g. if the initial host tree is a star. □

3.2 Proof of Theorem 1

Here we complete the proof of the main result of the paper, by designing a linear-time algorithm that determines $\overline{\chi}$ on all hypertrees of bounded degree over host trees of low maximum degree.

Preliminaries. Consider a hypertree $\mathcal{H}$ with $\Delta(\mathcal{H}) \leq d$ and its rooted host tree T, moreover, introduce the notation $D = \binom{d}{\lfloor d/2 \rfloor} + 1$. Due to Lemma 6 we may assume that $\Delta(T) \leq D$. From this point on, we use D as a universal upper bound for the vertex degrees in both T and $\mathcal{H}$. We also assume that the indexing $x_1, x_2, \ldots, x_n$ of vertices is consistent with a postorder traversal of T.

Based on Proposition 2 it suffices to consider C-colorings in which each color class induces a connected subgraph of T. Such color partitions are unambiguously characterized by the monochromatic lines of the host tree. Recall that T_i denotes the subtree rooted at x_i, moreover $X_i = V(T_i)$ and $\mathcal{H}_i = \{E \in \mathcal{E} \mid E \subseteq X_i\}$ have been introduced for every $1 \leq i \leq n$. Now, we partition the set of edges containing x_i into two classes:

- $\mathcal{H}_i^*$ — the set of edges whose highest (closest-to-root) vertex is x_i;
- $\mathcal{M}_i$ — the set of edges which contain x_i and the parent of x_i, too.

As a preliminary observation, we note that the set systems $\mathcal{H}_i$, $\mathcal{H}_i^*$, and $\mathcal{M}_i$ can be generated in linear time.

The algorithm. During the process of selecting monochromatic lines, in each step we investigate the optimal C-colorings of T_i, in increasing order of i. Bookkeeping will include:

- the value $w_i := dec(\mathcal{H}_i)$;
- the family $\mathfrak{F}_i$ of subcollections $\mathcal{F} \subseteq \mathcal{M}_i$ with the following property: T_i admits a C-coloring with just w_i monochromatic lines, in which each edge of $\mathcal{F}$ contains a monochromatic line, and $\mathcal{F}$ is maximal under inclusion with this property.

If these properties are ensured for all i, then the value w_n at the root is equal to $n - \overline{\chi}(\mathcal{H})$.

The steps of the computation can be executed as follows.

1. If x_i is a leaf of T, then let $w_i = 0$ and $\mathfrak{F}_i = \{\emptyset\}$. (Since $\mathcal{H}_i = \emptyset$, we have $w_i = dec(\mathcal{H}_i)$ and no edge of $\mathcal{M}_i$ contains any monochromatic lines.)

2. If the set of children of x_i is $S_i = \{x_{i_1}, \ldots, x_{i_s}\}$ $(s \geq 1)$, and when w_{i_ℓ} and $\mathfrak{F}_{i_\ell}$ have already been determined for all x_{i_ℓ} $(1 \leq \ell \leq s)$, then we apply the following procedure.

   (a) Initially let $k = 0$ and $\mathfrak{F}_i = \{\emptyset\}$.

   (b) Informally speaking — based on the optimal C-colorings of the children subtrees — for $k = 0, 1, \ldots, s$ we try to color properly all edges of $\mathcal{H}_i$ with precisely k monochromatic lines incident with x_i, and while doing so, we keep track of those edges in $\mathcal{M}_i$ which then become properly colored. More specifically, we consider each k-element subset $C = \{x_{j_1}, \ldots, x_{j_k}\} \subset S_i$ and each sequence $\mathfrak{S} = (\mathcal{F}_1, \ldots, \mathcal{F}_s)$ where, for all $1 \leq \ell \leq s$, we have chosen $\mathcal{F}_\ell \in \mathfrak{F}_{i_\ell}$. Every $\mathcal{F}_\ell$ corresponds to a set of monochromatic lines in T_{i_ℓ}; we extend the union of those sets with making also the k lines $x_i x_{j_q}$ monochromatic (i.e., the lines between x_i and C). Then, among the edges incident with x_i, the following set becomes properly colored:

   $$\mathcal{F} = (\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_s) \cup (\mathcal{M}_{j_1} \cup \cdots \cup \mathcal{M}_{j_k}).$$

   Hence,

   - if $\mathcal{H}_i^* \setminus \mathcal{F} \neq \emptyset$, then $\mathcal{H}_i$ is not colored properly;
   - if $\mathcal{H}_i^* \setminus \mathcal{F} = \emptyset$ holds for some C and $\mathfrak{S}$, then the k monochromatic lines joining x_i with C yield a C-coloring of $\mathcal{H}_i$. In this case if $\mathcal{F}^* = \mathcal{F} \setminus \mathcal{H}_i^*$ is not contained in any current element of $\mathfrak{F}_i$ as a subset, then we insert $\mathcal{F}^*$ into $\mathfrak{F}_i$ and remove its subsets from $\mathfrak{F}_i$ (if there occur any).

   (c) Having investigated all k-tuples of children of x_i paired with all sequences $\mathfrak{S} = (\mathcal{F}_1, \ldots, \mathcal{F}_s)$, and having obtained $\mathcal{H}_i^* \setminus \mathcal{F} \neq \emptyset$ in each case, we proceed with 2(b) putting $k := k + 1$ and $\mathfrak{F}_i := \{\emptyset\}$, as long as the updated k does not exceed s.

(d) If $\mathcal{H}_i^* \setminus \mathcal{F} = \emptyset$ occurs for some choice of C and $\mathfrak{S}$, then we obtain

$$w_i = k + \sum_{j=1}^{s} w_j.$$

We still investigate all remaining k-tuples of children of x_i in combination with all sequences $\mathfrak{S}$, but do not continue the procedure with any larger value of k afterwards. If $x_i = x_n$ (i.e., $i = n$), then the solution $w_n = dec(\mathcal{H})$ and $\overline{\chi}(\mathcal{H}) = n - w_n$ has been reached; otherwise $i := i + 1$ is taken and the steps are executed for the next vertex x_i.

Soundness of the algorithm. By Proposition 4, it is enough to consider the connected and all-optimal C-colorings of $\mathcal{H}$. The algorithm proceeds by constructing such a C-coloring φ; we prove this by induction.

In Step 1, the value $w(x_i) = 0$ corresponds to the decrement of the "subtree" rooted at a leaf vertex and $\emptyset$ is the only possible set of properly colored edges containing x_i.

Then, in Step 2 we identify the minimum number of monochromatic lines between x_i and its children $x_{i_1}, x_{i_2}, \dots, x_{i_s}$ subject to that the subtrees $\mathcal{H}_{i_1}, \mathcal{H}_{i_2}, \dots, \mathcal{H}_{i_s}$ have connected and all-optimal colorings. Clearly, it is enough to consider only those colorings of the subtrees $\mathcal{H}_{i_\ell}$ where the sets $\mathcal{F} \subseteq \mathcal{E}$ of edges containing a monochromatic line from T_{i_ℓ} are maximal under inclusion. Hence, the computed value of w_i is indeed the decrement of $\mathcal{H}_i$, for all $1 \le i \le n$. In particular, since $\mathcal{H}_n = \mathcal{E}$, the algorithm determines the exact value of $\overline{\chi}(\mathcal{H})$ correctly.

Time analysis. The edge sets $\mathcal{H}_i$ and $\mathcal{M}_i$ can be constructed in time proportional to $\sum_{H \in \mathcal{H}} |H| \le nD$. Since every vertex x_i in $\mathcal{H}$ and also in T has degree at most D, every $\mathfrak{F}_i$ has fewer than 2^D elements[4] For an x_i, in Step 2 we have to consider at most 2^D different sets C together with at most $(2^D)^s \le 2^{D^2}$ sequences $\mathfrak{S} = (\mathcal{F}_1, \dots, \mathcal{F}_s)$. This means not more than 2^{D^2+D} pairs $(C, \mathfrak{S})$ for each vertex x_i. Therefore, if D is fixed and a host tree of maximum vertex degree not greater than D is given in the input, the algorithm can be executed in time linear in n. □

[4] In fact, the stronger inequality $|\mathfrak{F}_i| \le \binom{D}{\lfloor D/2 \rfloor}$ also holds, since we have a Sperner family.

4 Concluding remarks

In this paper we investigated the upper chromatic number of hypertrees. We described several structural properties of optimal colorings and proved that $\overline{\chi}$ can be determined for hypertrees of bounded maximum degree in polynomial time. (With unbounded degrees the problem is known to be intractable.) An important question that remains open is as follows.

Problem 7 *What is the complexity of determining the exact value of $\overline{\chi}$ for hypertrees if the degrees are bounded only in their host trees?*

Under the assumption of bounded edge size, restricting the input to hypergraphs of bounded degree is equivalent to restriction to hypergraphs over host trees of bounded degree. For unbounded edge size, however, bounded degrees in the host tree do not guarantee the same in the hypergraph.

Acknowledgement

Research supported in part by the Hungarian Scientific Research Fund, OTKA grant T-81493, moreover by the European Union and Hungary, co-financed by the European Social Fund through the project TÁMOP-4.2.2.C-11/1/KONV-2012-0004 – National Research Center for Development and Market Introduction of Advanced Information and Communication Technologies.

References

[1] J. L. Arocha, J. Bracho and V. Neumann-Lara, On the minimum size of tight hypergraphs. *Journal of Graph Theory* 16 (1992), 319–326.

[2] G. Bacsó and Zs. Tuza, Upper chromatic number of finite projective planes. *Journal of Combinatorial Designs* 16:3 (2008), 221–230.

[3] C. Berge, *Hypergraphs*. North-Holland, 1989.

[4] Cs. Bujtás and Zs. Tuza, Maximum number of colors: C-coloring and related problems. *Journal of Geometry*, 101 (2011), 83–97.

[5] Cs. Bujtás and Zs. Tuza, Approximability of the upper chromatic number of hypergraphs, manuscript (2013).

[6] D. Král', On feasible sets of mixed hypergraphs. *Electronic Journal of Combinatorics* 11 (2004), #R19, 14 pp.

[7] D. Král', J. Kratochvíl and H.-J. Voss, Mixed hypergraphs with bounded degree: edge-coloring of mixed multigraphs. *Theoretical Computer Science* 295 (2003), 263–278.

[8] E. Sperner, Ein Satz über Untermengen einer endlichen Menge. *Mathematische Zeitschrift* 27:1 (1928), 544–548.

[9] F. Sterboul, A new combinatorial parameter. In: *Infinite and Finite Sets* (A. Hajnal et al., eds.), Colloq. Math. Soc. J. Bolyai 10, Vol. III, Keszthely 1973 (North-Holland/American Elsevier, 1975) 1387–1404.

[10] V. I. Voloshin, The mixed hypergraphs. *Computer Sci. J. Moldova* 1 (1993), 45–52.

[11] V. I. Voloshin, *Coloring Mixed Hypergraphs: Theory, Algorithms and Applications*, Fields Institute Monographs 17, Amer. Math. Soc., 2002.