Discrete Mathematics 338 (2015) 1714–1721

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Approximability of the upper chromatic number of hypergraphs[☆]

Csilla Bujtás [a,*], Zsolt Tuza [a,b]

[a] Department of Computer Science and Systems Technology, University of Pannonia, Veszprém, Hungary
[b] Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary

ARTICLE INFO

Article history:
Received 30 October 2013
Accepted 10 August 2014
Available online 29 August 2014

Keywords:
Approximation ratio
Hypergraph
Hypertree
C-coloring
Upper chromatic number
Multiple hitting set

ABSTRACT

A C-coloring of a hypergraph $\mathcal{H} = (X, \mathcal{E})$ is a vertex coloring $\varphi : X \to \mathbb{N}$ such that each edge $E \in \mathcal{E}$ has at least two vertices with a common color. The related parameter $\overline{\chi}(\mathcal{H})$, called the upper chromatic number of $\mathcal{H}$, is the maximum number of colors in a C-coloring of $\mathcal{H}$. A hypertree is a hypergraph which has a host tree T such that each edge $E \in \mathcal{E}$ induces a connected subgraph in T. Notations n and m stand for the number of vertices and edges, respectively, in a generic input hypergraph.

We establish guaranteed polynomial-time approximation ratios for the difference $n - \overline{\chi}(\mathcal{H})$, which is $2 + 2\ln(2m)$ on hypergraphs in general, and $1 + \ln m$ on hypertrees. The latter ratio is essentially tight as we show that $n - \overline{\chi}(\mathcal{H})$ cannot be approximated within $(1 - \epsilon)\ln m$ on hypertrees (unless $\mathsf{NP} \subseteq \mathsf{DTIME}\,(n^{\mathcal{O}(\log\log n)})$). Furthermore, $\overline{\chi}(\mathcal{H})$ does not have $\mathcal{O}(n^{1-\epsilon})$-approximation and cannot be approximated within additive error $o(n)$ on the class of hypertrees (unless $\mathsf{P} = \mathsf{NP}$).

1. Introduction

In this paper we study a hypergraph coloring invariant, termed upper chromatic number and denoted by $\overline{\chi}(\mathcal{H})$, which was first introduced by Berge (cf. [4]) in the early 1970's and later independently by several other authors [14,1,17] from different motivations. The present work is the very first one concerning approximation algorithms on it.

We also consider the complementary problem of approximating the difference $n - \overline{\chi}$, the number of vertices minus the upper chromatic number. For an upper bound on its approximability, one of our main tools is an approximation ratio that we establish for the 2-transversal number of hypergraphs. As problems of this type are of interest in their own right, we also prove an approximation ratio in general for the minimum size of multiple transversals, i.e., sets of vertices intersecting each edge in a prescribed number of vertices at least. Earlier results allowed to select a vertex into the set several times; we prove bounds for the more restricted scenario where the set does not include any vertex more than once.

1.1. Notation and terminology

A *hypergraph* $\mathcal{H} = (X, \mathcal{E})$ is a set system, where X denotes the set of vertices and each edge $E_i \in \mathcal{E}$ is a nonempty subset of X. Here we also assume that for each edge E_i the inequality $|E_i| \geq 2$ holds, moreover we use the standard notations $|X| = n$ and $|\mathcal{E}| = m$. A hypergraph $\mathcal{H}$ is said to be *r-uniform* if $|E_i| = r$ for each $E_i \in \mathcal{E}$.

☆ Research supported in part by the Hungarian Scientific Research Fund, OTKA grant T-81493, and 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.

* Corresponding author.
E-mail address: bujtas@dcs.vein.hu (Cs. Bujtás).

http://dx.doi.org/10.1016/j.disc.2014.08.007
0012-365X/© 2014 Elsevier B.V. All rights reserved.

We shall also consider hypergraphs with restricted structure, where some kind of host graphs are assumed. A hypergraph $\mathcal{H} = (X, \mathcal{E})$ admits a *host graph* $G = (X, E)$ if each edge $E_i \in \mathcal{E}$ induces a connected subgraph in G. The edges of the host graph G will be referred to as *lines*. Particularly, $\mathcal{H}$ is called *hypertree* or *hyperstar* if it admits a host graph which is a tree or a star, respectively. Note that under our condition, which forbids edges of size 1, $\mathcal{H}$ is a hyperstar if and only if there exists a fixed vertex $c^* \in X$ (termed the center of the hyperstar) contained in each edge of $\mathcal{H}$.

A *C-coloring* of $\mathcal{H}$ is an assignment $\varphi : X \to \mathbb{N}$ such that each edge $E \in \mathcal{E}$ has at least two vertices of a common color (that is, with the same image). The *upper chromatic number* $\overline{\chi}(\mathcal{H})$ of $\mathcal{H}$ is the maximum number of colors that can be used in a C-coloring of $\mathcal{H}$. We note that in the literature the value $\overline{\chi}(\mathcal{H}) + 1$ is also called the 'cochromatic number' or 'heterochromatic number' of $\mathcal{H}$ with the terminology of Berge [4, p. 151] and Arocha et al. [1], respectively. A C-coloring φ with $|\varphi(X)| = \overline{\chi}(\mathcal{H})$ colors will be referred to as an *optimal coloring* of $\mathcal{H}$. The *decrement* of $\mathcal{H} = (X, \mathcal{E})$, introduced in [2], is defined as $\mathrm{dec}(\mathcal{H}) = n - \overline{\chi}(\mathcal{H})$. Similarly, the decrement of a C-coloring $\varphi : X \to \mathbb{N}$ is meant as $\mathrm{dec}(\varphi) = |X| - |\varphi(X)|$. For results on C-coloring see the recent survey [8].

A *transversal* (also called hitting set or vertex cover) is a subset $T \subseteq X$ which meets each edge of $\mathcal{H} = (X, \mathcal{E})$, and the minimum cardinality of a transversal is the *transversal number* $\tau(\mathcal{H})$ of the hypergraph. An *independent set* (or stable set) is a vertex set $I \subseteq X$, which contains no edge of $\mathcal{H}$ entirely. The maximum size of an independent set in $\mathcal{H}$ is the *independence number* (or stability number) $\alpha(\mathcal{H})$. It is immediate from the definitions that the complement of a transversal is an independent set and vice versa, so the Gallai-type equality $\tau(\mathcal{H}) + \alpha(\mathcal{H}) = n$ holds for each hypergraph. Remark that selecting one vertex from each color class of a C-coloring yields an independent set, therefore $\overline{\chi}(\mathcal{H}) \leq \alpha(\mathcal{H})$ and, equivalently, $\mathrm{dec}(\mathcal{H}) \geq \tau(\mathcal{H})$.

More generally, a *k-transversal* is a set $T \subseteq X$ such that $|E_i \cap T| \geq k$ for every $E_i \in \mathcal{E}$. A 2-transversal is sometimes called double transversal or strong transversal, and its minimum size is the *2-transversal number* $\tau_2(\mathcal{H})$ of the hypergraph.

For an optimization problem and a constant $c > 1$, an algorithm $\mathcal{A}$ is called a *c-approximation algorithm* if, for every feasible instance $\mathcal{I}$ of the problem,

- if the value has to be minimized, then $\mathcal{A}$ delivers a solution of value at most $c \cdot \mathrm{Opt}(\mathcal{I})$;
- if the value has to be maximized, then $\mathcal{A}$ delivers a solution of value at least $\mathrm{Opt}(\mathcal{I})/c$.

Throughout this paper, an approximation algorithm is always meant to be one with polynomial running time on every instance of the problem. We say that a value has guaranteed approximation ratio c if it has a c-approximation algorithm. In the other case, when no c-approximation algorithm exists, we say that the value cannot be approximated within ratio c. For a function $f(n, m)$, an $f(n, m)$-approximation algorithm and the related notions can be defined similarly. A polynomial-time approximation scheme, abbreviated as PTAS, means an algorithm for every fixed $\varepsilon > 0$ which is a $(1 + \varepsilon)$-approximation and whose running time is a polynomial function of the input size (but any function of $1/\varepsilon$ may occur in the exponent).

For further terminology and facts we refer to [4,6,15] in the theory of graphs, hypergraphs, and algorithms, respectively. The notations $\ln x$ and $\log x$ stand for the natural logarithm and for the logarithm in base 2, respectively.

1.2. Approximability results on multiple transversals

The transversal number $\tau(\mathcal{H})$ of a hypergraph can be approximated within ratio $(1 + \ln m)$ by the classical greedy algorithm (see e.g. [15]). On the other hand, Feige [10] proved that $\tau(\mathcal{H})$ cannot be approximated within $(1 - \epsilon)\ln m$ for any constant $0 < \epsilon < 1$, unless $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\mathcal{O}(\log\log n)})$. As relates to the k-transversal number, in [15] a $(1 + \ln m)$-approximation is stated under the less restricted setting which allows multiple selection of vertices in the k-transversal. In the context of coloring, however, we cannot allow repetitions of vertices. For this more restricted case, when the k-transversal consists of pairwise different vertices, we prove a guaranteed approximation ratio $(1 + \ln(km))$.

In fact we consider a more general problem, where the required minimum size of the intersection $E_i \cap T$ can be prescribed independently for each $E_i \in \mathcal{E}$.

Theorem 1. *Given a hypergraph $\mathcal{H} = (X, \mathcal{E})$ with m edges $E_1, \ldots, E_m$ and positive integers $w_1, \ldots, w_m$ associated with the edges, the minimum cardinality of a set $S \subset X$ satisfying $|S \cap E_i| \geq w_i$ for all $1 \leq i \leq m$ can be approximated within $\sum_{i=1}^{W} 1/i < 1 + \ln W$, where $W = \sum_{i=1}^{m} w_i$.*

This result, proved in the next section, implies a guaranteed approximation ratio $(1 + \ln 2m)$ for $\tau_2(\mathcal{H})$.

1.3. Approximability results on the upper chromatic number

The problem of determining the upper chromatic number is NP-hard, already on the class of 3-uniform hyperstars. On the other hand, the problems of determining $\overline{\chi}(\mathcal{H})$ and finding a $\overline{\chi}(\mathcal{H})$-coloring are fixed-parameter tractable in terms of maximum vertex degree on the class of hypertrees [9].

A notion closely related to our present subject was introduced by Voloshin [16,17] in 1993. A *mixed hypergraph* is a triple $\mathcal{H} = (X, \mathcal{C}, \mathcal{D})$ with two families of subsets called $\mathcal{C}$-edges and $\mathcal{D}$-edges. By definition, a coloring of a mixed hypergraph is an assignment $\varphi : X \to \mathbb{N}$ such that each $\mathcal{C}$-edge has two vertices of a common color and each $\mathcal{D}$-edge has two vertices of distinct colors. Then, the minimum and the maximum possible number of colors, that can occur in a coloring of $\mathcal{H}$, is