

An approximation algorithm for the partial vertex cover problem in hypergraphs

Mourad El Ouali · Helena Fohlin ·
Anand Srivastav

Published online: 26 September 2014
© Springer Science+Business Media New York 2014

Abstract Let $\mathcal{H} = (V, \mathcal{E})$ be a hypergraph with set of vertices V , $n := |V|$ and set of (hyper-)edges \mathcal{E} , $m := |\mathcal{E}|$. Let l be the maximum size of an edge, Δ be the maximum vertex degree and D be the maximum edge degree. The k -partial vertex cover problem in hypergraphs is the problem of finding a minimum cardinality subset of vertices in which at least k hyperedges are incident. For the case of $k = m$ and constant l it known that an approximation ratio better than l cannot be achieved in polynomial time under the unique games conjecture (UGC) (Khot and Ravev J Comput Syst Sci, 74(3):335–349, 2008), but an l -approximation ratio can be proved for arbitrary k (Gandhi et al. J Algorithms, 53(1):55–84, 2004). The open problem in this context has been to give an αl -ratio approximation with $\alpha < 1$, as small as possible, for interesting classes of hypergraphs. In this paper we present a randomized polynomial-time approximation algorithm which not only achieves this goal, but whose analysis exhibits approximation phenomena for hypergraphs with $l \geq 3$ not visible in graphs: if Δ and l are constant, and $2 \leq l \leq 4\Delta$, we prove for l -uniform hypergraphs a ratio of $l \left(1 - \frac{l-1}{4\Delta}\right)$, which tends to the *optimal* ratio 1 as $l \geq 3$ tends to 4Δ . For the larger class of hypergraphs where $l, l \geq 3$, is not constant, but D is a constant, we show a

Parts of this paper appeared in preliminary form as an extended abstract in the proceedings of the 1st Mediterranean Conference on Algorithms, Kibbutz Ein Gedi, Isreal, December 2012 (MedAlg 2012), E. Guy and R. Dror, eds., Lecture Notes in Computer Science 7659, pp. 174–187, Springer Verlag 2012.

M. E. Ouali (✉) · A. Srivastav
Department of Computer Science, University of Kiel, Kiel, Germany
e-mail: meo@informatik.uni-kiel.de

A. Srivastav
e-mail: asr@informatik.uni-kiel.de

H. Fohlin
Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden
e-mail: Helena.Fohlin@lio.se

ratio of $l(1 - 1/6D)$. Finally for hypergraphs with non-constant l , but constant Δ , we get a ratio of $l(1 - \frac{2-\sqrt{3}}{6\Delta})$ for $k \geq m/4$, leaving open the problem of finding such an approximation for $k < m/4$.

Keywords Combinatorial optimization · Approximation algorithms · Hypergraphs · Vertex cover · Probabilistic methods

1 Introduction

A hypergraph $\mathcal{H} = (V, \mathcal{E})$ consists of a set V , say $|V| = n$, and a set \mathcal{E} of subsets of V , $|\mathcal{E}| = m$. We call the elements of V vertices and the elements of \mathcal{E} (hyper-)edges. The k -partial vertex cover in hypergraphs can be stated as follows. A set $X \subseteq V$ is called a k -partial vertex cover for \mathcal{H} if at least k edges of \mathcal{H} are incident in X . The (unweighted) k -partial vertex cover problem for hypergraphs is to find a k -vertex cover of minimum cardinality. If k is equal to the number of hyperedges, we have the well-known hitting set problem in hypergraphs. For graphs it is the vertex cover problem, whose approximation complexity has been studied for nearly 4 decades.

Among the motivations to study hypergraphs are not only the natural question of generalizing graph theory to hypergraphs (Berge 1989), but also relevant applications in areas where hypergraphs are most natural, e.g. data structures in computational geometry like ϵ -nets (Matousek and Wagner 2004), which are a kind of hitting sets, or discrepancy theory (Matousek 2010).

1.1 Previous work

Consider a minimization problem. For $\rho \geq 1$ we say that a polynomial-time algorithm achieves an ρ -approximation or an approximation ratio of ρ , if it computes for all instances a solution of value at most $\rho \cdot \text{Opt}$, where Opt is the value of an optimal solution to the problem.

For the k -partial vertex cover problem in *graphs* it has been an open question for which graphs the well-known 2-approximation can be improved. For graphs with maximum vertex degree at most a constant Δ , Gandhi et al. (2004) gave the first algorithm with approximation ratio smaller than 2. Further improvements have been obtained by Halperin and Srinivasan (2002).

For hypergraphs the hitting set and the set cover problem have been investigated intensively in the context of polynomial-time approximations (Chvátal 1979; Hochbaum 1982; Lovász 1975; Duh and Fürer 1997; Halperin 2000; Krivelevich 1997) as well as non-approximability (Alon et al. 2006; Feige 1998; Lund and Yannakakis 1994; Raz and Safra 1997).

For the hitting set problem in l -uniform hypergraphs with constant l an approximation with a ratio strictly smaller than l cannot be achieved in polynomial time under the unique games conjecture (UGC) (Khot and Regev 2008). Since the hitting set problem in hypergraphs is a special case of the k -vertex cover problem in hypergraphs, namely with $k = m$, this hardness of approximation holds also for the k -vertex cover problem in

hypergraphs. On the other hand, for the k -partial vertex cover problem in hypergraphs with edge size at most l (l not necessarily constant) [Gandhi et al. \(2004\)](#) gave a polynomial time l -approximation based on the primal/dual approach. Thus polynomial-time approximations below the l -ratio for significant classes of hypergraphs are complexity-theoretic and algorithmically interesting and would extend the approximation theory for the k -partial vertex cover problem from graphs to hypergraphs. Our work is a contribution to this specific task, and to an approximation theory for problems in hypergraphs.

1.2 Our contribution

We give a randomized algorithm of hybrid type, combining linear programming (LP) based randomized rounding and afterwards greedy repairing to ensure the feasibility of the randomized solution. While the algorithm is a straightforward generalization of the algorithm of [Gandhi et al. \(2004\)](#) from graphs to hypergraphs, the main challenge is its analysis in different hypergraph settings. In general, the analysis of hybrid randomized algorithms is a difficult task, as the probabilistic analysis has to combine different dependent subroutines (e.g. maximum graph bisection [Frieze and Jerrum \(1997\)](#), maximum graph partitioning [Feige and Langberg \(2001\)](#); [Jäger and Srivastav \(2005\)](#) and vertex cover and partial vertex cover problem for *graphs* [Gandhi et al. \(2004\)](#); [Halperin \(2000\)](#)). Similar problems appear also in our proofs.

We show the following approximation results: If both Δ and l are constant, $2 \leq l \leq 4\Delta$, we prove for a l -uniform hypergraph with maximum vertex degree Δ an approximation ratio of $l(1 - \frac{l-1}{4\Delta})$. This ratio tends to the *optimal* ratio 1 as l is at least 3 and tends to 4Δ . Thus for hypergraphs with large l the approximability is much better than for graphs, matching our intuition that large hyperedges may cause many intersections and thus a few vertices might be sufficient to get a k -partial vertex cover. The analysis is based on variance computation and the Chebyshev–Cantelli bound. For the larger class of hypergraphs, where $l, l \geq 3$, is not constant, the above mentioned approach does not work. Among this class we consider hypergraphs with a constant maximum edge degree D and show an approximation ratio of $l(1 - 1/6D)$. Here again variance computation and the Chebyshev–Cantelli bound are invoked, but more work is needed to minimize the probability of an infeasible k -partial vertex cover. Finally for hypergraphs with non-constant $l, l \geq 3$, but constant Δ , a similar analysis as above can be carried out, but with the severe restriction $l \leq 2$, which does not reveal any new insight. Interestingly, we can overcome this restriction with the bounded difference inequality (based on the Azuma–Hoeffding bound) and prove for any $l, l \geq 3$ and $k \geq m/4$ a ratio of $l(1 - \frac{2-\sqrt{3}}{\Delta})$.

An open problem for hypergraphs arising from our work is the question whether for a hypergraph with maximum vertex degree Δ , Δ a constant, but arbitrary edge size l , an approximation ratio of $l(1 - \frac{c}{\Delta})$, $c > 0$ a constant, can be shown for any k , in particular for $k < m/4$.

Note that the approach to lift the analysis in [Gandhi et al. \(2004\)](#) from the k -vertex cover problem in *graphs* to hypergraphs fails. Among the reasons are the different hypergraphs settings with different hypergraph parameters (e.g. l, D and Δ) requiring

a situation adapted application of different concentration results in combination with combinatorial arguments.

1.3 Outline of the paper

The paper is organised as follows: In Sect. 2 we give the necessary definitions and state some probabilistic tools. In Sect. 3 we present the randomized algorithm for the k -partial vertex cover for hypergraphs and state estimations for the expectation and the variance of the random variables under consideration. In Sect. 4 we analyse the algorithm for l -uniform hypergraphs where l and Δ are constants. In Sect. 5 resp. 6 we analyse the approximation ratio for hypergraphs with constant D resp. Δ .

2 Preliminaries and definitions

2.1 Graph-theoretical notions

For a natural number n let $[n]$ denote the set $\{1, \dots, n\}$. Let $\mathcal{H} = (V, \mathcal{E})$ be a (finite) hypergraph, where V is a finite set, called the set of vertices, and \mathcal{E} is the set of hyperedges (or simply edges) consisting of subsets of V . For $v \in V$ we define the degree of a vertex v by $d(v) := |\{E \in \mathcal{E}; v \in E\}|$ and the maximum vertex degree by $\Delta := \max\{d(v); v \in V\}$. For a set $X \subseteq V$ we denote by $\mathcal{E}/X := \{F \in \mathcal{E}; X \cap F \neq \emptyset\}$ the set of edges incident to X . For an edge $E \in \mathcal{E}$ let $\text{deg}(E) := |\mathcal{E}/E|$ be the (edge-) degree of E and $D := \max_{E \in \mathcal{E}} \text{deg}(E)$ is the maximum edge degree. Note that $\Delta \leq D \leq l\Delta$. For $l \in \mathbb{N}$ we call the hypergraph \mathcal{H} l -uniform resp. l -bounded, if $|E| = l$ resp. $|E| \leq l$ for all $E \in \mathcal{E}$. In this notion l is not assumed to be a constant. If l is assumed to be constant, we will make it explicitly clear. Otherwise, we will say that l is non-constant. It is convenient to order the vertices and hyperedges, $V = \{v_1, \dots, v_n\}$ and $\mathcal{E} = \{E_1, \dots, E_m\}$, and to identify vertices and edges with their indices. This is an useful abbreviation of the notation especially when we model the optimization problem under consideration as an integer linear program and look at its linear programming relaxation.

2.2 Concentration inequalities

We will frequently use the Chebychev–Cantelli inequality:

Theorem 1 (see [Motwani and Raghavan \(1995\)](#), p 64) *Let X be a non-negative random variable with finite mean $\mathbb{E}(X)$ and variance $\text{Var}(X)$. Then for any $a > 0$ we have*

$$\Pr(X \geq \mathbb{E}(X) + a) \leq \frac{\text{Var}(X)}{\text{Var}(X) + a^2} \tag{1}$$

$$\Pr(X \leq \mathbb{E}(X) - a) \leq \frac{\text{Var}(X)}{\text{Var}(X) + a^2}. \tag{2}$$

A further useful concentration result used in this paper is the bounded differences inequality:

Theorem 2 (see [McDiarmid \(1989\)](#), p 216) *Let $X = (X_1, X_2, \dots, X_n)$ be a family of independent random variables with X_k taking values in a set A_k for each k . Suppose that the real-valued function f defined on $A_1 \times A_2 \cdots \times A_n$ satisfies $|f(x) - f(x')| \leq c_k$ for all pairs of vectors x and x' from $A_1 \times A_2 \cdots \times A_n$ that differ only in the k -th coordinate, $k = 1, \dots, n$. Then for any $t > 0$ it holds*

$$\Pr [f(X) \leq \mathbb{E}(f(X)) - t] \leq \exp\left(\frac{-2t^2}{\sum_{k=1}^n c_k^2}\right).$$

Let X be the sum of finitely many 0/1 random variables, i.e. $X = X_1 + \dots + X_n$. For a pair $i, j \in \{1, \dots, n\}$ we say that i and j are dependent, if X_i and X_j are not independent. Let Γ be the set of all unordered dependent pairs i, j , i.e. 2-element sets $\{i, j\}$, and let $\gamma = \sum_{\{i,j\} \in \Gamma} \mathbb{E}(X_i X_j)$. The following estimate on the variance of such a sum of dependent random variables can be proved as in the book of [Alon and Spencer \(2000\)](#), p 41:

Lemma 1 *We have $\text{Var}(X) \leq \mathbb{E}(X) + 2\gamma$.*

For a sum of independent random variables we will use the large deviation inequalities due to [Angluin and Valiant \(1979\)](#):

Theorem 3 (see [McDiarmid \(1989\)](#), p 200) *Let X_1, \dots, X_n be independent 0/1-random variables and $\mathbb{E}(X_i) = p_i$ for all $i = 1, \dots, n$. Let $X = \sum_{i=1}^n X_i$ with $\mu = \mathbb{E}(X)$. For any $\beta > 0$ it holds*

$$\Pr(X \geq (1 + \beta) \cdot \mu) \leq \exp\left(-\frac{\beta^2 \mu}{3}\right).$$

3 The algorithm and estimation of mean and variance

The input is a l -bounded hypergraph \mathcal{H} , $l \geq 2$. At the moment we do not assume D or Δ to be constants. Before we proceed to the randomized algorithm, let us rule out trivial cases of the problem solvable to optimality in polynomial time (not optimizing the time complexity here).

Proposition 1 *Let \mathcal{H} be a hypergraph with n vertices and m hyperedges. If k is a constant, then the k -partial vertex cover can be solved to optimality in $O(kmn^k)$ -time.*

Proof Since an optimal k -partial vertex cover has cardinality at most k , we can find such a cover by examining all sets of vertices of cardinality r , $r \leq k$. Their number is $\sum_{r=1}^k \binom{n}{r} \leq O(kn^k)$. The test whether or not such a r -set is a k -partial vertex cover requires the test of at most all the m edges, so the total complexity is at most $O(kmn^k)$. \square

As already said, we will identify the vertices and edges of \mathcal{H} by their indices. An integer linear programming formulation of the k -partial vertex cover in \mathcal{H} is the following:

$$\begin{aligned}
 & \min \sum_{j=1}^n x_j \\
 \text{(ILP-k-VC)} \quad & \sum_{j=1}^n a_{ij}x_j \geq z_i \quad \text{for all } i \in [m] := \{1, \dots, m\} \\
 & \sum_{i=1}^m z_i \geq k \\
 & x_j, z_i \in \{0, 1\} \quad \text{for all } i \in [m], j \in [n].
 \end{aligned}$$

Its linear programming (LP) relaxation LP- k -VC is given by relaxing the integrality constraints to $x_j, z_i \in [0, 1] \forall i \in [m], j \in [n]$. Let Opt resp. Opt^* be the value of an optimal solution to ILP- k -VC resp. to LP- k -VC. Let $x^* \in [0, 1]^n$ and $z^* \in [0, 1]^m$ be an optimal solution of LP- k -VC for the variables $x_j, z_i, j \in [n]$ and $i \in [m]$. Note that an optimal solution of a linear program can be computed in polynomial time, for example with the ellipsoid method of [Khachian \(1979\)](#) or more efficiently with the interior-point method of [Karmarkar \(1984\)](#).

3.1 The Algorithm

Algorithm 1: k -VC(\mathcal{H})

Input : $k \in \mathbb{N}$ and a hypergraph $\mathcal{H} = (V, \mathcal{E})$

Output: A k -partial vertex cover C

1. **Initialisation.** $C := \emptyset$. Choose $\epsilon \in (0, 1)$ and $\lambda = 1/(1 - \epsilon)$ such that $\lambda \geq 1$.
 2. **LP-Relaxation.** Solve the LP relaxation of ILP- k -VC with a polynomial-time algorithm with an optimal solution $(x_j^*)_{j \in [n]}$. Then $\text{Opt}^* = \sum_{j \in [n]} x_j^*$ should denote its value.
 3. **Definition of S -sets.** Set $S_1 := \{j \in [n] \mid x_j^* = 1\}$, $S_{\geq} := \{j \in [n] \mid 1 - x_j^* \geq \frac{1}{\lambda}\}$ and $S_{<} := \{j \in [n] \mid 1 - x_j^* < \frac{1}{\lambda}\}$.
 4. **Pre-Cover.** Take all vertices of S_1 and S_{\geq} into C .
 5. **Randomized Rounding.** For all $j \in S_{<}$ add vertex j to C with probability λx_j^* , independently.
 6. **Repairing.**
 - a) If $|\{E \in \mathcal{E} \mid E \cap C \neq \emptyset\}| \geq k$, then return C .
 - b) If $|\{E \in \mathcal{E} \mid E \cap C \neq \emptyset\}| < k$, then pick at most $k - |\{E \in \mathcal{E} \mid E \cap C \neq \emptyset\}|$ additional vertices from arbitrary not covered edges in C .
 7. Return C .
-

The algorithm extends the randomized algorithm of Gandhi, Khuller and Srinivasan [Gandhi et al. \(2004\)](#) from graphs to hypergraphs. While the extension is quite natural, the analysis needs efforts beyond [Gandhi et al. \(2004\)](#) as already mentioned in the introduction.

3.2 Computation of expectation and variance

We give some basic estimates. Note that at this moment we have not to comment on the randomized rounding parameter ϵ as the proofs in this section work for any $\epsilon \in (0, 1)$.

But the choice of ϵ will become a crucial point when we apply the algorithm to the different hypergraph families in later sections.

Let X_1, \dots, X_n be $\{0, 1\}$ -random variables defined as follows. For $j \in [n]$ let

$$X_j = \begin{cases} 1 & \text{if vertex } v_j \text{ was picked by randomized rounding or if } v_j \in S_1 \cup S_{\geq} \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Note that the random variables X_1, \dots, X_n are independent and

$$Y := \sum_{j=1}^n X_j \quad (4)$$

is the cardinality of the set C after randomized rounding. For all $i \in [m]$ we define the $\{0, 1\}$ -random variables Z_i as follows:

$$Z_i = \begin{cases} 1 & \text{if the the hyperedge } E_i \text{ is covered after the rounding step} \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

The sum of the Z_i 's

$$W := \sum_{j=1}^m Z_j \quad (6)$$

is the number of covered hyperedges before entering the repairing step. Note that in this step at most $k - W$ further vertices are added to C . Hence, for the expected size of the final cover C returned by the algorithm we have

Proposition 2

$$\mathbb{E}(|C|) \leq \mathbb{E}(Y) + \mathbb{E}(\max\{k - W, 0\}). \quad (7)$$

For the computation of the expectation of W we need the following lemma that gives the exact solution of a constrained optimization problem (Lemma 2.2, Peleg et al. (1997)).

Lemma 2 For all $n \in \mathbb{N}$, $\alpha > 0$ and $x_1, \dots, x_n, z \in [0, 1]$ with $\sum_{i=1}^n x_i \geq z$ and $\alpha x_i < 1$ for all $i \in \mathbb{N}$, we have $\prod_{i=1}^n (1 - \alpha x_i) \leq (1 - \alpha \frac{z}{n})^n$, and this bound is the tight maximum.

For the expectation and the variance the following lemma will serve as a technical backbone.

Lemma 3 Let D and Δ be as above, not assumed to be constants, and let $\epsilon \in (0, 1)$.

- (i) $(1 - (1 - \epsilon)x)^2 \leq 1 - x(1 - \epsilon^2)$ for all $x \in [0, 1]$.
- (ii) $\mathbb{E}(W) \geq (1 - \epsilon^2)k$.

- (iii) (1) $\text{Var}(W) \leq (D + 1)\mathbb{E}(W)$, (2) $\text{Var}(|C|) \leq l\Delta\mathbb{E}(|C|)$.
- (iv) $\text{Opt}^* \geq \frac{k}{\Delta} \geq \frac{k}{D}$.
- (v) For $\lambda \geq 1$, $\text{Opt}^* \leq \mathbb{E}(Y) \leq \lambda\text{Opt}^*$.
- (vi) $\sum_{v \in V} d(v)^2 \leq \Delta lm$.

Proof (i) Straightforward calculations show that assertion (i) is equivalent to

$$x(1 + \epsilon^2 - 2\epsilon) \leq 1 + \epsilon^2 - 2\epsilon.$$

But this is true for any $x \in [0, 1]$.

(ii) Let $i \in [m]$ and $|E_i| = r_i$. If there is a $j \in E_i$ with $\lambda x_j^* \geq 1$, then $j \in S_1 \cup S_{\geq}$. Thus j belongs to the cover C , $Z_i = 1$, and trivially $\Pr(Z_i = 0) = 0$. Otherwise, if $\lambda x_j^* < 1$ for all $j \in E_i$, we have

$$\begin{aligned} P(Z_i = 0) &= \prod_{j \in E_i} (1 - \lambda x_j^*) \stackrel{\text{Lem2}}{\leq} \left(1 - \frac{\lambda z_i^*}{r_i}\right)^{r_i} \\ &\leq \left(1 - \frac{\lambda z_i^*}{l}\right)^{r_i} = (1 - (1 - \epsilon)z_i^*)^{r_i} \\ &\leq (1 - (1 - \epsilon)z_i^*)^2 \stackrel{\text{Lem3(i)}}{\leq} 1 - z_i^*(1 - \epsilon^2), \end{aligned}$$

so $P(Z_i = 0) \leq 1 - z_i^*(1 - \epsilon^2)$, and we get

$$\begin{aligned} \mathbb{E}(W) &= \sum_{i=1}^m \Pr(Z_i = 1) = \sum_{i=1}^m (1 - \Pr(Z_i = 0)) \\ &\geq \sum_{i=1}^m (1 - (1 - z_i^*(1 - \epsilon^2))) = \sum_{i=1}^m z_i^*(1 - \epsilon^2) = (1 - \epsilon^2) \underbrace{\sum_{i=1}^m z_i^*}_{\geq k} \\ &\geq (1 - \epsilon^2)k. \end{aligned}$$

(iii) Part (1) According to the notation of Lemma 1, let $\Gamma = \{\{i, j\}; i, j \in [m] \text{ are dependent}\}$ and put $\gamma = \sum_{\{i,j\} \in \Gamma} \mathbb{E}(Z_i Z_j)$. Since the Z_i 's are functions of the random variables X_k , $k \in E_i$, for every $E_i, E_j \in \mathcal{E}$ the random variables Z_i, Z_j are not independent iff the hyperedges E_i and E_j have a non-empty intersection. Thus, for a fixed E_i , there are at the most D random variables Z_j possibly depending on Z_i . For every $E_i, E_j \in \mathcal{E}$ we have

$$\begin{aligned} \mathbb{E}(Z_i Z_j) &= \Pr(Z_i = 1 \wedge Z_j = 1) \\ &\leq \min\{\Pr(Z_i = 1), \Pr(Z_j = 1)\} \\ &\leq \frac{\Pr(Z_i = 1) + \Pr(Z_j = 1)}{2}. \end{aligned}$$

So

$$\begin{aligned} \gamma &= \sum_{\{i,j\} \in \Gamma} \mathbb{E}(Z_i Z_j) \\ &\leq \sum_{\{i,j\} \in \Gamma} \frac{\Pr(Z_i = 1) + \Pr(Z_j = 1)}{2} \\ &\leq \sum_{i=1}^m \frac{D}{2} \Pr(Z_i = 1) = \frac{D}{2} \mathbb{E}(W). \end{aligned}$$

With Lemma 1 we conclude

$$\text{Var}(W) \leq \mathbb{E}(W) + 2\gamma \leq \mathbb{E}(W) + D\mathbb{E}(W) = (D + 1)\mathbb{E}(W).$$

(iii) Part (2) We modify the proof of (iii) part (1). Let U_j be the 0/1 random variable, which is 1 iff the vertex j is contained in the final cover C after repairing. According to the notation of Lemma 1, let $\Gamma = \{\{i, j\}; i, j \in [n] \text{ are dependent}\}$ and let $\gamma = \sum_{\{i,j\} \in \Gamma} \mathbb{E}(U_i U_j)$. Furthermore for every pair $U_i, U_j \in \mathcal{E}$, U_i, U_j are not independent, if the selection of vertex i into the final cover depends on the selection of vertex j into the final cover, and vice versa. We have

$$\begin{aligned} \mathbb{E}(U_i U_j) &= \Pr(U_i = 1 \wedge U_j = 1) \\ &\leq \min\{\Pr(U_i = 1), \Pr(U_j = 1)\} \\ &\leq \frac{\Pr(U_i = 1) + \Pr(U_j = 1)}{2}. \end{aligned}$$

So

$$\begin{aligned} \gamma &= \sum_{\{i,j\} \in \Gamma} \mathbb{E}(U_i U_j) \\ &\leq \sum_{\{i,j\} \in \Gamma} \frac{\Pr(U_i = 1) + \Pr(U_j = 1)}{2} \\ &\leq \sum_{i=1}^n \frac{\Delta(l - 1)}{2} \Pr(U_i = 1) = \frac{\Delta(l - 1)}{2} \mathbb{E}(|C|). \end{aligned}$$

The last inequality is based on the following argument: for every pair $U_i, U_j \in \mathcal{E}$, U_i, U_j are not independent, if the selection of vertex i into C depends on the selection of vertex j , and vice versa. The selection of vertex i can be affected by vertices contained in edges incident in i and these are at most $\Delta(l - 1)$.

With Lemma 1 we conclude

$$\text{Var}(|C|) \leq \mathbb{E}(|C|) + 2\gamma \leq \mathbb{E}(|C|) + \Delta(l - 1)\mathbb{E}(|C|) = (\Delta(l - 1) + 1)\mathbb{E}(|C|) \leq l\Delta\mathbb{E}(|C|).$$

(iv) Since x^* and z^* are optimal solutions of LP- k -VC, we have

$$k \leq \sum_{i=1}^m z_i^* \leq \sum_{i=1}^m \sum_{j \in E_i} x_j^* = \sum_{j=1}^n d(v_j)x_j^* \leq \Delta \sum_{j=1}^n x_j^* = \Delta \cdot \text{Opt}^* \leq D \cdot \text{Opt}^*.$$

(v) For a set $X \subset V$, we define $\text{Opt}^*(X) := \sum_{j \in X} x_j^*$. By using the LP relaxation and the definition of the sets S_1, S_{\geq} and $S_{<}$, and since $\lambda \geq 1$, we get

$$\text{Opt}^* \leq \underbrace{|S_1| + |S_{\geq}| + \lambda \text{Opt}^*(S_{<})}_{=\mathbb{E}(Y)} \leq \underbrace{|S_1|}_{\leq \lambda \text{Opt}^*(S_1)} + \underbrace{|S_{\geq}|}_{\leq \lambda \text{Opt}^*(S_{\geq})} + \lambda \text{Opt}^*(S_{<}) \leq \lambda \text{Opt}^*.$$

(vi) Let r_j the size of the hyperedge $E_j, j \in [m]$. Then trivially $\sum_{v \in V} d(v)^2 \leq \Delta \sum_{v \in V} d(v)$. From $\sum_{v \in V} d(v) = \sum_{j=1}^m r_j$, and $l = \max_{j \in [m]} r_j$ the assertion (vi) follows. □

4 Hypergraphs with constant l and Δ

We consider in this section a l -uniform hypergraph with maximum vertex degree Δ , where l, Δ are constants. Let A be the set of vertices added by randomized rounding to C , and B be set of vertices added by the repairing step to C . Let Y, W be the random variables as defined in (4) resp. (6). Then, obviously

$$C = S_1 \cup S_{\geq} \cup A \cup B \text{ and } |C| = Y + |B| \leq Y + k - W. \tag{8}$$

To run the algorithm k -VC(\mathcal{H}) we have to choose ϵ . We define

$$\epsilon := l \text{Opt}^* / 2k \tag{9}$$

It must be ensured that $\epsilon \in (0, 1)$:

Proposition 3 *If $\epsilon > 1/2$, then a simple polynomial-time algorithm returns a k -partial vertex cover which is an approximation with ratio strictly smaller than l , without any further assumption on the degree D and Δ , and on l .*

Proof If $\epsilon > 1/2$, then $l \text{Opt}^* > k$. A k -partial vertex cover of cardinality k can be found trivially by selecting k arbitrary hyperedges and picking one vertex out of each of them. This cover approximates the optimum within a ratio strictly smaller than l . □

By Proposition 3 we may assume w.l.o.g

$$0 \leq \epsilon \leq 1/2, \text{ and thus } \lambda \geq 1 \text{ for any } l \geq 2, \tag{10}$$

and run the algorithm k -VC(\mathcal{H}) with ϵ as in (9).

Lemma 4 $\mathbb{E}(|C|) \leq l \left(1 - \frac{l}{4\Delta}\right) \text{Opt}^*$

Proof We have

$$\begin{aligned}
 \mathbb{E}(|C|) &= \mathbb{E}(Y) + \mathbb{E}(|B|) \\
 &\stackrel{\text{Lem 3(ii),(v)}}{\leq} \lambda \text{Opt}^* + k\epsilon^2 = l(1 - \epsilon)\text{Opt}^* + k\epsilon^2 \\
 &\quad l \left(1 - \frac{l\text{Opt}^*}{2k} \right) \text{Opt}^* + \frac{l^2\text{Opt}^{*2}}{4k} \\
 &\quad l \left(1 - \frac{l\text{Opt}^*}{4k} \right) \text{Opt}^* \\
 &\stackrel{\text{Lem 3(iv)}}{\leq} l \left(1 - \frac{l}{4\Delta} \right) \text{Opt}^*.
 \end{aligned}$$

□

Lemma 4 and Theorem 1 imply the following theorem.

Theorem 4 *Let l, Δ be constants and let \mathcal{H} be an l -uniform hypergraph with maximum vertex degree Δ . We further assume that $2 \leq l \leq 4\Delta$. Then the algorithm k -VC(\mathcal{H}) returns a k -partial vertex cover C such that*

$$|C| \leq l \left(1 - \frac{l-1}{4\Delta} \right) \text{Opt}^* \text{ with probability at least } 2/3.$$

Proof By Proposition 1 we may assume $k \geq 32\Delta^4$ w.l.o.g. We have

$$\begin{aligned}
 \Pr \left(|C| \geq l \left(1 - \frac{l-1}{4\Delta} \right) \text{Opt}^* \right) &= \Pr \left(|C| \geq l \left(1 - \frac{l}{4\Delta} \right) \text{Opt}^* + \frac{l\text{Opt}^*}{4\Delta} \right) \\
 &\stackrel{\text{Lem 4}}{\leq} \Pr \left(|C| \geq \mathbb{E}(|C|) + \frac{l\text{Opt}^*}{4\Delta} \right) \\
 &\stackrel{\text{Th 1(1)}}{\leq} \frac{\text{Var}(|C|)}{\text{Var}(|C|) + \left(\frac{l\text{Opt}^*}{4\Delta} \right)^2} = \frac{1}{1 + \frac{\left(\frac{l\text{Opt}^*}{4\Delta} \right)^2}{\text{Var}(|C|)}}.
 \end{aligned}$$

By Lemma 3, $\text{Var}(|C|) \leq l\Delta\mathbb{E}(|C|)$, so

$$\begin{aligned}
 \frac{\left(\frac{l\text{Opt}^*}{4\Delta} \right)^2}{\text{Var}(|C|)} &\stackrel{\text{Lem 3(iii)-(2)}}{\geq} \left(\frac{l(\text{Opt}^*)^2}{16\Delta^3\mathbb{E}(|C|)} \right) \\
 &\stackrel{\mathbb{E}(|C|) \leq l\text{Opt}^*}{\geq} \frac{\text{Opt}^*}{16\Delta^3} \\
 &\stackrel{\text{Lem 3(iv)}}{\geq} \frac{k}{16\Delta^4} \stackrel{k \geq 32\Delta^4}{\geq} 2.
 \end{aligned}$$

Therefore we get $\Pr \left(|C| \geq l \left(1 - \frac{l-1}{4\Delta} \right) \text{Opt}^* \right) \leq \frac{1}{1+2} = 1/3$.

□

The proof required $l \leq 4\Delta$. As Δ is assumed to be constant, so l is constant as well. That is the reason why we cannot transfer this approach to the more general setting, where the maximum edge degree D resp. the maximum vertex degree Δ are constant, but l is not constant, see the upcoming Sects. 5 and 6.

5 Analysis for bounded edge degree

In this section we consider hypergraphs with maximum edge size $l \in \mathbb{N}, l \geq 3$, not necessarily assumed to be a constant, but with a constant maximum edge degree D . W.l.o.g we may assume that $D \geq 1$, because otherwise we only have isolated edges and the partial vertex cover problem is trivial. We choose

$$\epsilon := \frac{\text{Opt}^*(1 + \beta)}{k} \quad \text{where} \quad \beta = \frac{1}{3D}. \tag{11}$$

Again we have to ensure that w.l.o.g $\epsilon \in (0, 1)$.

Proposition 4 *We can assume that*

$$\epsilon \leq \frac{1 + \beta}{l - \eta}, \quad \text{with} \quad \eta = \frac{l}{6D}, \tag{12}$$

in particular, $\epsilon \leq 1/2$.

Proof Otherwise if $\epsilon > \frac{1+\beta}{l-\eta}$, it follows from the definition of ϵ in (11) that $\text{Opt}^* > \frac{k}{l-\eta}$, hence $l(1 - \frac{\eta}{l})\text{Opt}^* > k$. Since a partial vertex cover of size k can be trivially found by picking k arbitrary hyperedges and taking one vertex from each of them, pairwise distinct, we get the claimed $l(1 - \frac{\eta}{l})$ approximation ratio. \square

So (12) and $\epsilon \in (0, 1)$ holds w.l.o.g. and we may run the algorithm $k\text{-VC}(\mathcal{H})$ with this ϵ and $\lambda = l(1 - \epsilon)$. Note that $\lambda \geq 1$ due to $l \geq 3$. The main result of this section is:

Theorem 5 *Let \mathcal{H} be a hypergraph with edge size at most $l, l \geq 3$, and constant maximum edge degree D . The algorithm $k\text{-VC}(\mathcal{H})$ with ϵ defined as in (11) and satisfying (12) returns a k -partial vertex cover C such that*

$$|C| \leq l \left(1 - \frac{1}{6D}\right) \text{Opt} \quad \text{with probability at least} \quad \frac{7}{15}.$$

Proof **Claim 1**

$$\Pr \left(W \leq k(1 - \epsilon^2) - 2\sqrt{kD} \right) \leq \frac{1}{3}$$

Proof of Claim 1 We can assume that $k(1 - \epsilon^2) - 2\sqrt{kD}$ is non-negative. Otherwise, since $W \geq 0$ the event " $W \leq k(1 - \epsilon^2) - 2\sqrt{kD}$ " has probability 0.

Let us consider the function

$$f : (D, \infty) \rightarrow \mathbb{R}, f(x) = x - 2\sqrt{Dx}, x \in (D, \infty).$$

Since $f'(x) = 1 - \sqrt{\frac{D}{x}} > 0$, f is strictly monotone increasing. For $k < 4D$ it holds

$$k(1 - \epsilon^2) - 2\sqrt{kD} \leq \underbrace{k - 2\sqrt{kD}}_{=f(k)} < f(4D) = 0,$$

and because $W \geq 0$, $\Pr(W \leq k(1 - \epsilon^2) - 2\sqrt{kD}) = 0$.

Therefore, we can assume $k \geq 4D$. We set $\mu := \mathbb{E}(W)$. Using $\epsilon \leq \frac{1}{2}$ we get

$$k(1 - \epsilon^2) \geq 4D(1 - \epsilon^2) > 4D(1 - \frac{1}{4}) = 3D > D.$$

By Lemma 3(ii), $k(1 - \epsilon^2) \leq \mu$, and as f is monotone increasing, $f(k(1 - \epsilon^2)) \leq f(\mu)$. We continue

$$\begin{aligned} \Pr(W \leq k(1 - \epsilon^2) - 2\sqrt{kD}) &\leq \Pr(W \leq k(1 - \epsilon^2) - 2\sqrt{k(1 - \epsilon^2)D}) \\ &= \Pr(W \leq f(k(1 - \epsilon^2))) \\ &\leq \Pr(W \leq f(\mu)) \\ &= \Pr(W \leq \mu - 2\sqrt{\mu D}) \\ &\stackrel{\text{Th 1(2)}}{\leq} \frac{\text{Var}(W)}{\text{Var}(W) + 4\mu D} = \frac{1}{1 + \frac{4\mu D}{\text{Var}(W)}} \\ &\stackrel{\text{Lem 3(iii-1)}}{\leq} \frac{1}{1 + \frac{4\mu D}{\mu(D+1)}} \leq \frac{1}{3}. \end{aligned}$$

This concludes the proof of Claim 1.

Claim 2 For $\beta = \frac{1}{3D}$ it holds $\Pr(Y \geq l \cdot \text{Opt}^*(1 - \epsilon)(1 + \beta)) < \frac{1}{5}$.

Proof of Claim 2 According to Proposition 1 we may assume that

$$k \geq 16D^5. \tag{13}$$

The random variables X_1, \dots, X_n defined as in (3) are independent and an application of the Angulin-Valiant bound, Theorem 3, shows

$$\begin{aligned} \Pr(Y \geq l(1 - \epsilon)(1 + \beta)\text{Opt}^*) &\stackrel{\text{Lem 3(v)}}{\leq} \Pr(Y \geq \mathbb{E}(Y)(1 + \beta)) \\ &\stackrel{\text{Th 3}}{\leq} \exp\left(-\frac{\beta^2 \mathbb{E}(Y)}{3}\right). \end{aligned}$$

Note that

$$\frac{\mathbb{E}(Y)\beta^2}{3} \stackrel{\text{Lem 3(v)}}{\geq} \frac{\text{Opt}^*}{27D^2} \stackrel{\text{Lem 3(iv)}}{\geq} \frac{k}{27D^3} \geq \frac{16D^2}{27}.$$

Using $D \geq 2$ we continue

$$\Pr(Y \geq l(1 - \epsilon)(1 + \beta)\text{Opt}^*) \leq \exp\left(-\frac{16D^2}{27}\right) \leq \exp\left(-\frac{64}{27}\right) \leq \exp(-2) < \frac{1}{5}.$$

This concludes the proof of Claim 2.

By Claim 1 and 2 we get for the final partial vertex cover C with probability at least $1 - (\frac{1}{5} + \frac{1}{3}) = \frac{7}{15}$

$$|C| \leq \underbrace{l(1 - \epsilon)(1 + \beta)\text{Opt}^* + k\epsilon^2}_{(*)} + \underbrace{2\sqrt{kD}}_{(**)}.$$

We estimate the terms $(*)$ and $(**)$ in the right hand side separately.

$$\begin{aligned} (*) &= l\left((1 + \beta)(1 - \epsilon) + \frac{\text{Opt}^*(1 + \beta)^2}{lk}\right)\text{Opt}^* \\ &\leq l\left((1 + \beta) - \frac{(l - 1)\text{Opt}^*(1 + \beta)^2}{lk}\right)\text{Opt}^* \\ &\stackrel{\text{Lem 3(iv)}}{\leq} l\left((1 + \beta) - \frac{(l - 1)(1 + \beta)^2}{lD}\right)\text{Opt}^*. \end{aligned}$$

Using $l \geq 3$ we have

$$\frac{(l - 1)(1 + \beta)^2}{lD} - \beta \geq \frac{2(1 + \beta)^2}{3D} - \frac{1}{3D} \geq \frac{1}{3D},$$

therefore

$$l(1 - \epsilon)(1 + \beta)\text{Opt}^* + k\epsilon^2 \leq l\left(1 - \frac{1}{3D}\right)\text{Opt}^*.$$

We proceed to the second term $(**)$. By Lemma 3(iv) we have $\text{Opt}^* \geq \frac{k}{D} \geq 16D^4$, so

$$\begin{aligned} 2\sqrt{kD} &= 2D\sqrt{\frac{k}{D}} \stackrel{\text{Lem 3(iv)}}{\leq} 2D\sqrt{\text{Opt}^*} \\ &= 2\sqrt{\text{Opt}^*}\sqrt{\text{Opt}^*}\frac{D}{\sqrt{\text{Opt}^*}} \\ &\stackrel{\text{Lem 3(iv) and (13)}}{\leq} 2\text{Opt}^*\frac{D}{4D^2} \\ &= \text{Opt}^*\frac{1}{2D}. \end{aligned}$$

Finally,

$$\begin{aligned}
 (*) + (**) &\leq l\text{Opt}^* \left(1 - \frac{1}{3D}\right) + l\text{Opt}^* \frac{1}{2lD} \\
 &\leq_{l \geq 3} l \left(1 - \frac{1}{6D}\right) \text{Opt}^*,
 \end{aligned}$$

holds with probability at least $7/15$. □

6 Analysis for constant vertex degree

In this section we consider hypergraphs with edge size at most $l, l \geq 3$ non-constant, but with constant maximum vertex degree Δ . We may assume $\Delta \geq 2$, because otherwise the hypergraph has only isolated edges and the k -partial vertex cover problem is trivial. One may ask if the approach in the last section is transferable to our present situation, just replacing D by Δ . In fact, everything works, except the proof of Claim 1. There we would get the estimate

$$\frac{1}{1 + \frac{4\mu\Delta}{\text{Var}(W)}} \leq \frac{1}{1 + \frac{4\mu\Delta}{\mu(D+1)}}. \tag{14}$$

With the (trivial) estimate $D \leq l\Delta$, the last term is at most $\frac{1}{1 + \frac{4\Delta}{l\Delta+1}}$ which is smaller than $1/(1+c)$ for a constant $c > 0$, if $l \leq 4/c - 1$, thus $l \leq 2$. We are back to graphs, where such an approximation is known [Gandhi et al. \(2004\)](#), and have gained nothing. In the following we will demonstrate, how an involvement of the Hoeffding–Azuma bound in form of the bounded difference inequality can overcome this restriction and can handle hypergraphs with $l \geq 3$. We wish to run the algorithm k -VC- (\mathcal{H}) and choose

$$\epsilon := \frac{\text{Opt}^*(1 + \beta_1)}{k} \quad \text{for} \quad \beta_1 = \frac{1}{3\Delta}. \tag{15}$$

Again, w.l.o.g. we can assume $\epsilon \leq 1/2$:

Proposition 5 *We can assume that*

$$\epsilon \leq \frac{1 + \beta_1}{l - \eta}, \quad \text{with} \quad \eta = \frac{l}{6\Delta}. \tag{16}$$

In particular, $\epsilon \leq 1/2$

Proof Otherwise, if $\epsilon > \frac{1+\beta_1}{l-\eta}$, it follows from the definition of ϵ in (15) that $\text{Opt}^* > \frac{k}{l-\eta}$, hence $l(1 - \frac{\eta}{l})\text{Opt}^* > k$. Since a partial vertex cover of size k can be trivially found by picking k arbitrary hyperedges and taking one vertex from each of them, pairwise distinct, we get a $l(1 - \frac{\eta}{l})$ approximation ratio, which is even better than the ratio in Theorem 6. □

So (16) and $\epsilon \in (0, 1)$ hold w.l.o.g. and we may run the algorithm k -VC- (\mathcal{H}) with this ϵ and $\lambda = l(1 - \epsilon)$. Note that $\lambda \geq 1$ due to $l \geq 3$. Its analysis gives

Theorem 6 *Let \mathcal{H} be a hypergraph with edge size at most l , $l \geq 3$ non-constant, and bounded vertex degree Δ . For $k \geq \frac{m}{4}$ the algorithm k -VC (\mathcal{H}) returns a k -partial vertex cover C such that*

$$|C| \leq l \left(1 - \frac{2 - \sqrt{3}}{6\Delta} \right) \text{Opt} \quad \text{with probability at least } \frac{3}{5}.$$

This is an improvement over Theorem 5 because in general $\Delta \leq D$.

Proof of Theorem 6 We need the following claim and shall prove it with the bounded difference inequality, Theorem 2.

Claim 4

$$\Pr \left(W \leq k(1 - \epsilon^2) - 2\sqrt{k l \Delta} \right) \leq \frac{1}{5}.$$

Proof of Claim 4 Because $W \geq 0$, we can assume that the upper bound on W in Claim 4 is non-negative. W itself is a function

$$W : \{0, 1\}^n \rightarrow \mathbb{N}, \quad W(X_1, \dots, X_n) = \sum_{j=1}^m Z_j.$$

W is component-wise Lipschitz bounded: Let $k \in [n]$, $X = (X_1, \dots, X_{k-1}, X_k, X_{k+1}, \dots, X_n)$ and

$X' = (X_1, \dots, X_{k-1}, X'_k, X_{k+1}, \dots, X_n)$. Then

$$|W(X) - W(X')| \leq d(v_k) \text{ for all } k.$$

This is true, because as X_k, X'_k differ, the two associated covers C resp. C' differ by the vertex v_k . But then the edges hit by C resp. C' differ by at most $d(v_k)$ edges. By Theorem 2 we get for any $t > 0$,

$$\Pr(W - \mathbb{E}(W) \leq -t) \leq \exp \left(\frac{-2t^2}{\sum_{v \in V} d(v)^2} \right). \tag{17}$$

We choose $t = 2\sqrt{\Delta l k}$.

$$\begin{aligned} \Pr \left(W \leq k(1 - \epsilon^2) - 2\sqrt{\Delta l k} \right) &\stackrel{\text{Lem 3 (ii)}}{\leq} \Pr \left(W \leq \mathbb{E}(W) - 2\sqrt{\Delta l k} \right) \\ &\stackrel{\text{Ineq (17)}}{\leq} \exp \left(\frac{-8\Delta l k}{\sum_{v \in V} d(v)^2} \right) \\ &\stackrel{\text{Lem 3 (vi) and } k \geq \frac{m}{4}}{\leq} e^{-2} < \frac{1}{5}, \end{aligned}$$

and Claim 4 is proved.

Claim 5 For $\beta_1 = \frac{1}{3\Delta}$ it holds $\Pr(Y \geq l \cdot \text{Opt}^*(1 - \epsilon)(1 + \beta_1)) < \frac{1}{5}$.

Proof of Claim 5 Again, Proposition 1 allows to assume w.l.o.g

$$k \geq 16\Delta^5. \tag{18}$$

The Angulin-Valiant inequality yields

$$\begin{aligned} \Pr(Y \geq l(1 - \epsilon)(1 + \beta_1)\text{Opt}^*) &\stackrel{\text{Lem 3(v)}}{\leq} \Pr(Y \geq \mathbb{E}(Y)(1 + \beta_1)) \\ &\stackrel{\text{Th 3}}{\leq} \exp\left(-\frac{\beta_1^2 \mathbb{E}(Y)}{3}\right). \end{aligned}$$

With Lemma 3 (iv) resp. (v), $k \geq 16\Delta^5$ and $\Delta \geq 2$ we get

$$\frac{\mathbb{E}(Y)\beta_1^2}{3} \geq \frac{k}{27\Delta^3} \geq \frac{16\Delta^2}{27} \geq \frac{64}{27} \geq 2.$$

So

$$\Pr(Y \geq l(1 - \epsilon)(1 + \beta_1)\text{Opt}^*) \leq \exp(-2) < \frac{1}{5}.$$

This concludes the proof of Claim 5.

By Claim 4 and 5 we get an upper bound for the final cover with probability at least $1 - (\frac{1}{5} + \frac{1}{5}) \geq \frac{3}{5}$:

$$|C| \leq \underbrace{l(1 - \epsilon)(1 + \beta_1)\text{Opt}^* + k\epsilon^2}_{(*)} + \underbrace{2\sqrt{kl\Delta}}_{(**)}.$$

As in the proof of Theorem 5 we can show

$$(*) = l(1 - \epsilon)(1 + \beta_1)\text{Opt}^* + k\epsilon^2 \leq l\left(1 - \frac{1}{3\Delta}\right)\text{Opt}^*.$$

By Lemma 3(iv) and (18) it holds:

$$\text{Opt}^* \geq \frac{k}{\Delta} \geq 16\Delta^4,$$

so

$$\begin{aligned} 2\sqrt{kl\Delta} &= (2\sqrt{l})\sqrt{\frac{k}{\Delta}}\Delta \leq (2\sqrt{l})\sqrt{\text{Opt}^*}\Delta = \frac{2l\text{Opt}^*\Delta}{\sqrt{l}\sqrt{\text{Opt}^*}} \\ &\stackrel{l \geq 3}{\leq} l\text{Opt}^* \frac{2}{\Delta\sqrt{48}}. \end{aligned}$$

Finally the sum of (*) and (**) is

$$\begin{aligned}
 (*) + (**) &\leq l\text{Opt}^* \left(1 - \frac{1}{3\Delta}\right) + l\text{Opt}^* \frac{2}{\Delta\sqrt{48}} \\
 &\leq l \left(1 - \frac{1}{3\Delta} + \frac{2}{\Delta\sqrt{48}}\right) \text{Opt}^* \\
 &= l \left(1 - \frac{2 - \sqrt{3}}{6\Delta}\right) \text{Opt}^*,
 \end{aligned}$$

and this holds with probability at least $3/5$. □

This result shows again a hypergraph characteristic. An approximation ratio of the form $l \left(1 - \frac{c}{\Delta}\right)$ for some constant $c > 0$ is valid for hypergraphs and the k -partial vertex cover problem with $k \geq m/4$, while in graphs it is valid for any k . By examining our probabilistic arguments, there seems to be no obvious way to release the restriction on k . We leave it as an open problem to prove or to disprove an approximation ratio of the form $l \left(1 - \frac{c}{\Delta}\right)$ for constant Δ and $k < m/4$. In view of Theorem 4, the first interesting hypergraph for the investigation of this problem would be a Δ -regular and $(4\Delta + 1)$ -uniform hypergraph. For such hypergraphs, none of our results apply, if we consider $k < m/4$.

7 Further work

We find it interesting to give better approximations for hypergraphs with other kind of sparseness conditions, like uncrowdedness or small VC-dimension. Another challenge is the derandomization of this and other hybrid algorithms combining randomized rounding and greedy heuristic in order to obtain feasible solutions.

References

Alon N, Moshkovitz D, Safra S (2006) Algorithmic construction of sets for k -restrictions. *ACM Trans Algorithms* 2:153–177

Angluin D, Valiant LG (1979) Fast probabilistic algorithms for Hamiltonian circuits and matchings. *J Comput Syst Sci* 18:155–193

Alon N, Spencer J (2000) *The probabilistic method*, 2nd edn. Wiley Interscience, Hoboken, NJ

Berge C (1989) *Hypergraphs- combinatorics of finite sets*. North Holland Mathematical Library

Chvátal V (1979) A greedy heuristic for the set covering problem. *OMEGA* 4(3):233–235

Duh R, Fürer M (1997) Approximating k -set cover by semi-local optimization. In: *Proceedings of the 29th annual ACM symposium on theory of computing (STOC '97)*, pp 256–264, May

Feige U (1998) A threshold of $\ln n$ for approximating set cover. *J ACM* 45(4):634–652

Feige U, Langberg M (2001) Approximation algorithms for maximization problems arising in graph partitioning. *J Algorithms* 41(2):174–201

Frieze A, Jerrum M (1997) Improved approximation algorithms for max k -cut and max bisection. *Algorithmica* 18:67–81

Gandhi R, Khuller S, Srinivasan A (2004) Approximation algorithms for partial covering problems. *J Algorithms* 53(1):55–84

Halperin E (2000) Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs. *ACM-SIAM Symp Discret Algorithms* 11:329–337

- Halperin E, Srinivasan A (2002) Improved approximation algorithms for the partial vertex cover problem. In: 5th international workshop on approximation algorithms for combinatorial optimization problems, LNCS, vol 2462, pp 161–174
- Hochbaum DS (1982) Approximation algorithms for the set covering and vertex cover problems. *SIAM J Comput* 11(3):555–556
- Jäger G, Srivastav A (2005) Improved approximation algorithms for maximum graph partitioning problems. *J Comb Optim* 10(2):133–167
- Karmarkar N (1984) A new polynomial time algorithm for linear programming. *Combinatorica* 4(4):373–395
- Khachian LG (1979) A polynomial algorithm in linear programming [in Russian]. *Doklady Akademii Nauk SSSR* 244: 1093–1096, 1979. English translation: *Soviet Mathematics Doklady* 20: 191–194
- Khot S, Regev O (2008) Vertex cover might be hard to approximate to within 2-epsilon. *J Comput Syst Sci* 74(3):335–349
- Krivelevich J (1997) Approximate set covering in uniform hypergraphs. *J Algorithms* 25(1):118–143
- Lovász L (1975) On the ratio of optimal integral and fractional covers. *Discrete Math* 13:383–390
- Lund C, Yannakakis M (1994) On the hardness of approximating minimization problems. *J Assoc Comput Mach* 41:960–981
- Matousek J (2010) *Geometric discrepancy. Algorithms and combinatorics*. Springer, Berlin
- Matousek J, Wagner U (2004) New constructions of weak epsilon-nets. *Discret Comput Geom* 32(2):195–206
- McDiarmid C (1989) On the method of bounded differences. *Surveys in combinatorics* (Norwich, 1989), pp 148–188. Cambridge University Press, Cambridge
- Motwani R, Raghavan P (1995) *Randomized algorithms*. Cambridge University Press, Cambridge
- Peleg D, Schechtman G, Wool A (1997) Randomized approximation of bounded multicovering problems. *Algorithmica* 18(1):44–66
- Raz R, Safra S (1997) A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In: *Proceedings 29th ACM symposium on theory of computing*, pp 475–484