Approximability of Vertex Cover in Dense Bipartite Hypergraphs

Jean Cardinal* Marek Karpinski‡
Richard Schmied† Claus Viehmann§

Abstract

We study approximation complexity of the Vertex Cover problem restricted to dense and subdense balanced k-partite k-uniform hypergraphs. The best known approximation algorithm for the general k-partite case achieves an approximation ratio of $\frac{k}{2}$ which is the best possible assuming the Unique Game Conjecture. In this paper, we present approximation algorithms for the dense and the subdense nearly regular instances both with an approximation factor strictly better than $\frac{k}{2}$. On the other hand, we show that the latter approximation upper bound is almost tight under the Unique Games Conjecture.

1 Introduction

The Vertex Cover problem is one of the classical optimization problems proven to be NP-hard in Karp [18]. Given graph G, it consists of finding a minimum cardinality subset of vertices having a nonempty intersection with every edge of G. The problem can be generalized to the minimum vertex cover problem on k-uniform hypergraphs where a k-uniform hypergraph H is a pair $(V(H), E(H))$ with a set of vertices $V(H)$ and a set of hyperedges $E(H)$, in which each hyperedge consists of a set containing exactly k vertices. In addition to it, H is called k-partite if the vertex set $V(H)$ can be partitioned into k nonempty pairwise disjoint sets $(V_1, .., V_k)$ such that each of these sets contains exactly one vertex of each edge and moreover, called balanced if the sizes of the sets V_i are all equal, i.e. $|V_i| = \frac{|V(H)|}{k}$ for all $i \in [k]$.

*Université Libre de Bruxelles (ULB), CP212. Supported by the Communauté française de Belgique (projet ARC). Email: jcardin@ulb.ac.be
†Dept. of Computer Science, University of Bonn. Work supported by Hausdorff Doctoral Fellowship. Email: schmied@cs.uni-bonn.de
‡Dept. of Computer Science and the Hausdorff Center for Mathematics, University of Bonn. Supported in part by DFG grants and the Hausdorff Center grant EXC59-1. Email: marek@cs.uni-bonn.de
§Dept. of Computer Science, University of Bonn. Work partially supported by Hausdorff Center for Mathematics, Bonn. Email: viehmann@cs.uni-bonn.de

The minimum vertex cover problem is k-approximable in k-uniform hypergraphs. This approximation ratio is achieved by a simple approximation algorithm which chooses a maximal set of nonintersecting edges, and then outputs all vertices in them. Interestingly, the best known approximation ratio is $k-(k-1)\ln\ln n/\ln n$ and is due to Halperin [12].

On the lower bound side, Trevisan [27] proved one of the first inapproximability results for the k-uniform vertex cover problem, namely a factor $k^{1/19}$. After an improvement of the inapproximability factor by Holmerin [12] to $k^{1-\epsilon}$, Dinur et. al . [7, 8] gave consecutively two lower bounds, first $(k-3-\epsilon)$ and later on $(k-1-\epsilon)$.

The Unique Games Conjecture (UGC) was introduced by Subhash Khot in 2002 [22]. The conjecture postulates the NP-hardness of determining the value of a optimization problem known as the unique game.

Assuming the unique games conjecture, Khot and Regev [23] proved an inapproximability factor of $k-\epsilon$ for the Vertex Cover problem on k-uniform hypergraphs. Thus, it implies that the currently achieved ratios are the best possible. Only recently, Bansal and Khot [3] showed under the UGC that the same inapproximability factor of $k-\epsilon$ holds even for almost k-partite k-uniform hypergraphs.

While the Vertex Cover problem in graphs and hypergraphs is intractable in general, it is well known that on bipartite graphs, it is solvable in polynomial time. For the Vertex Cover problem on k-partite k-uniform hypergraphs, Lovász [25] achieved a $\left(\frac{k}{2}\right)$-approximation by rounding its natural LP relaxation. In [1], a tight integrality gap of $\frac{k}{2}-o(1)$ was given for the LP relaxation. On the inapproximability side, Ilie, Solis-Oba, and Yu [15] as well as Gottlob and Senellart [10] constructed reductions from 3SAT to it, which imply that the problem is APX-hard. Recently, Guruswami and Saket [11] showed it is NP-hard to approximate the minimum vertex cover problem on k-partite k-uniform hypergraphs to within a factor of $\frac{k}{4}-\epsilon$ for $k\geq 5$, and within a factor of $\frac{k}{2}-\epsilon$ assuming the unique games conjecture for $k\geq 3$.

In order to shed some additional light on lower bounds for general problems, dense instances of many optimization problems has been studied [2, 20, 21, 19]. The vertex Cover problem restricted to dense graphs, where the number of edges is within a constant factor of n^2, was considered by Karpinski and Zelikovsky [21], Eremeev [9], Clementi and Trevisan [6], and later by Imamura and Iwama [16].

The Vertex Cover problem restricted to ϵ-dense hypergraphs, i.e. hypergraphs with $\epsilon\binom{n}{k}$ hyperedges, was introduced and studied by Bar-Yehuda and Kehat [4]. They provided an approximation algorithm with a better approximation ratio than k. Based on Imamura and Iwama's recursive sampling technique [16] for vertex cover in dense graphs, an approximation algorithm with better approximation ratio than k was proposed in [5] for the subdense case.
In this paper, we investigate the approximability of the VC problem restricted to dense and subdense balanced k-partite k-uniform hypergraphs. To the best of our knowledge, this is the only result tackling the dense and the subdense version of this problem.

1.1 Definitions

Let S be a finite set, we use the notation $\binom{S}{k} = \{S' \subseteq S \mid |S'| = k\}$ and $[i] := \{1, 2, \ldots, i\}$. A *$k$-uniform hypergraph* H is a pair $(V(H), E(H))$ with a vertex set $V(H)$ and an edge set $E(H) \subseteq \binom{V}{k}$. A k-uniform hypergraph (V, E) is called *k-partite* if there exists vertex classes $(V_1, V_2, \ldots, V_k)$ such that V is a disjoint union of $V_1, V_2, \ldots, V_k$ with $|V_i \cap e| = 1$ for every $e \in E$ and $i \in [k]$. Furhtermore, we call a k-partite k-uniform hypergraph *balanced* if the vertex partition $(V_1, .., V_k)$ possesses the property $|V_i| = \frac{|V|}{k}$ for all $i \in [k]$. If the k-partition $(V_1, .., V_k)$ of a k-partite k-uniform hypergraph $H = (V(H), E(H))$ is given as a part of the input, we use the notation $H = (V_1, .., V_k, E(H))$. We set $n = |V(H)|$ and $m = |E(H)|$ as usual. In the remainder, we assume that $k = O(1)$.

A *vertex cover* of a k-uniform hypergraph (V, E) is a subset C of V with the property $e \cap C \neq \emptyset$ for all $e \in E$. The Vertex Cover problem consists of finding a vertex cover of minimum size in a given hypergraph.

For a vertex $v \in V$, we define the *degree* $d(v)$ of v to be $|\{e \in E \mid v \in e\}|$. For a subset $S \subseteq V$ on the other hand, we define $d(S)$ as $|\{e \in E \mid S \subseteq e\}|$. We use the abbreviations $\bar{d}$ and Δ for the average degree and maximum degree of a hypergraph, respectively.

We define a balanced k-partite k-uniform hypergraph $H = (V(H), E(H))$ as *ℓ-wise ϵ-dense* for $(\ell + 1) \in [k]$ and $\epsilon \in [0, 1]$ if there exists an $I \in \binom{[k]}{\ell}$ such that the condition $d(S) \geq \epsilon \left(\frac{n}{k}\right)^{k-\ell}$ holds for all subsets $S \subseteq V(H)$ with the restriction $|V_i \cap S| = 1$ and $i \in I$.

An extension of the ϵ-density is the ψ-density. In particular, a balanced k-partite k-uniform hypergraph $H = (V(H), E(H))$ is called $\psi(n)$-dense if the maximum degree Δ and average degree $\bar{d}$ are $\bar{d} = \Theta(\Delta)$ and $\Delta = \frac{n^{k-1}}{\psi(n)}$. Furthermore, a $\psi(n)$-dense hypergraph is called subdense if $\psi(n) = O(\log(n))$ holds.

1.2 Our Results

In this paper, we study the dense and the subdense vertex cover problem in balanced k-partite k-uniform hypergraphs. We prove that a modified version of the approximation algorithms given in [5] for the dense and the subdense vertex cover problem in k-uniform hypergraphs yields improved approximation upper bounds for the balanced k-partite case. On the other hand, we show that the achieved approximation upper bound in the subdense case is almost tight assuming the UGC.

In [5], an approximation algorithm for the vertex cover problem in ℓ-wise ϵ-dense k-uniform hypergraphs was proposed with an approximation ratio of

$$\frac{k}{k - (k-1)(1-\epsilon)^{1/(k-\ell)}}.$$

We improve this approximation ratio in two different ways: In the one hand, the modifications of the algorithm result in an improved analysis and an approximation factor of

$$\frac{k}{k-(k-2)\left(1-\epsilon\right)^{1/(k-\ell)}}.$$

On the other hand, we obtain an improved value for the parameter ϵ, since we use a different definition of ϵ-density in balanced k-partite k-uniform hypergraphs.

For the subdense version of the Vertex Cover problem in k-uniform hypergraphs, the randomized algorithm given in [5] yields an approximation factor

$$\frac{k}{\left(1+(k-1)\frac{\bar{d}}{k\Delta}\right)}.$$

Our modified version achieves an improved approximation upper bound

$$\frac{k}{\left(2+(k-2)\frac{\bar{d}}{\Delta}\right)}$$

on balanced k-partite k-uniform hypergraphs with $\Delta = \Omega\left(\frac{n^{k-1}}{\log n}\right)$.

However, in Section 4, we prove that the approximation ratio of our algorithm for the subdense case is almost optimal under the Unique Games Conjecture [23].

2 Vertex Cover in ϵ-Dense k-Partite Hypergraphs

In this section, we consider the Vertex Cover problem restricted to ℓ-wise ϵ-dense balanced k-partite k-uniform hypergraphs. For this case, we present the following result:

Theorem 1. *The Vertex Cover problem can be approximated in polynomial time with an approximation ratio*

$$\frac{k}{k-(k-2)\left(1-\epsilon\right)^{\frac{1}{k-\ell}}}-o(1)$$

in ℓ-wise ϵ-dense k-partite k-uniform hypergraphs.

Firstly, we provide some lemmas needed to prove Theorem 1. Let us start with a Lemma which is an extension of Lemma 1 in [5].

Lemma 1. *Let $H=(V_1,\ldots,V_k,E(H))$ be a k-partite k-uniform hypergraph with a minimum vertex cover C, and let $W\subseteq V(H)$ such that $|W\cap C|\geq\delta|W|$ for some constant $\delta\in[0,1]$, and we can find a vertex cover R of H in polynomial time. Given W, there is a polynomial time approximation algorithm with an approximation ratio*

$$\frac{k}{2+(\delta k-2)\frac{|W|}{|R|}}.$$

Proof. Let $\tilde{H}$ be the hypergraph induced by the edges that are not covered by W and $\tilde{C}$ an optimal vertex cover of $\tilde{H}$. Clearly, $\tilde{H}$ is a k-partite k-uniform hypergraph. Therefore, we can apply the $\left(\frac{k}{2}\right)$-approximation algorithm due to Lovász [25] in order to generate a vertex cover $\tilde{W}$ of $\tilde{H}$ with $\frac{|\tilde{W}|}{|\tilde{C}|} \leq \frac{k}{2}$. Then, our combined approach returns the solution $S := argmin_{S' \in \{W \cup \tilde{W}, R\}}\{|S'|\}$. Let us analyze the approximation ratio ρ. For this purpose, we consider the worst case solution S'. By definition, we have $|S'| = \rho|C|$. Therefore, we obtain $\frac{\rho|C|}{|R|} = \frac{|S'|}{|R|} \leq 1$ and conclude that

$$
\begin{aligned}
|S| = \rho|C| &\leq |W| + \frac{k}{2}|C'| &(1)\\
&= |W| + \frac{k}{2}\left(|C| - \delta|W|\right) &(2)\\
&= \frac{k}{2}|C| - \left(\delta\frac{k}{2} - 1\right)|W| &(3)\\
&\leq \frac{k}{2}|C| - \left(\delta\frac{k}{2} - 1\right)\frac{|W|}{|R|}\rho|C| &(4)\\
\Rightarrow \rho &\leq \frac{k}{2} - \left(\delta\frac{k}{2} - 1\right)\frac{|W|}{|R|}\rho &(5)\\
\rho &\leq \frac{k}{2 + (\delta k - 2)\frac{|W|}{|R|}}. &(6)
\end{aligned}
$$

□

Since we can check in polynomial time for every fixed $j \geq 1$, if there is a $p \in \{k-1, .., j\}$, $i \in \{1, .., k\}$ and $R \in \binom{V_i}{|V_i|-p}$ such that R is a vertex cover of H, we can assume that the returned solution is $\leq \left(\min_{i \in \{1,..,k\}}\{|V_i|\}\right) - j \leq \frac{n}{k} - j$. This simple fact combined with the previous Lemma results in the following

Corollary 1. *Let $H = (V(H), E(H))$ be a balanced k-partite k-uniform hypergraph with a minimum vertex cover C. Given set $W \subseteq V(H)$ such that $|W \cap C| \geq \delta|W|$ for some constant $\delta \in [0, 1]$ and every $j \geq 1$, there is a polynomial time approximation algorithm with an approximation ratio*

$$\frac{k}{2 + (\delta k - 2)|W|\left(\frac{n}{k} - j\right)^{-1}}.$$

The following Lemma also plays a key role in our analysis.

Lemma 2. *In a 0-wise ϵ-dense balanced k-partite k-uniform hypergraph H, the first $\left(1 - (1-\epsilon)^{\frac{1}{k}}\right)n$ highest-degree vertices all have degree at least $\left(1 - (1-\epsilon)^{\frac{k-1}{k}}\right)\left(\frac{n}{k}\right)^{(k-1)}$.*

Proof. Let us consider a hypergraph H with $m \geq \epsilon \left(\frac{n}{k}\right)^k$ hyperedges. We define W as the set of the first $\left(1-(1-\epsilon)^{\frac{1}{k}}\right)n$ highest-degree vertices (breaking ties arbitrarily). Our goal is to prove our statement by contradiction. For this purpose, let us assume that the number m of edges in H is strictly smaller than the number of edges in a hypergraph where all vertices of W have degree $\left(\frac{n}{k}\right)^{k-1}$, and all the remaining edges have degree $\left(1-(1-\epsilon)^{\frac{k-1}{k}}\right)\left(\frac{n}{k}\right)^{k-1}$. Therefore, we obtain

$$\begin{aligned}
m &< \frac{1}{k}\left(|W|\left(\frac{n}{k}\right)^{k-1} + (n-|W|)\left(1-(1-\epsilon)^{\frac{k-1}{k}}\right)\left(\frac{n}{k}\right)^{k-1}\right) \\
&= \frac{1}{k}\left(\left(1-(1-\epsilon)^{\frac{1}{k}}\right)n\left(\frac{n}{k}\right)^{k-1}\right. \\
&\quad \left. + \left(n-\left(1-(1-\epsilon)^{\frac{1}{k}}\right)n\right)\left(1-(1-\epsilon)^{\frac{k-1}{k}}\right)\left(\frac{n}{k}\right)^{k-1}\right) \\
&= \left(1-(1-\epsilon)^{\frac{1}{k}}\right)\left(\frac{n}{k}\right)^{k} + (1-\epsilon)^{\frac{1}{k}}\left(1-(1-\epsilon)^{\frac{k-1}{k}}\right)\left(\frac{n}{k}\right)^{k} \\
&= \epsilon\left(\frac{n}{k}\right)^{k}
\end{aligned}$$

Clearly, this contradicts the fact that H is ϵ-dense. □

In order to prove Theorem 1, we first consider the case $\ell = 0$. The recursive algorithm depicted in figure 1 finds a large subset W of a minimum vertex cover. More precisely, it returns a polynomial-sized collection $\mathcal{W}$ of subsets $W_i \subseteq V$, in which at least one W_i is contained in a minimum vertex cover. Finally, we are able to apply Lemma 1.

Next, we prove the following important Lemma:

Lemma 3. *Given a 0-wise ϵ-dense k-partite k-uniform hypergraph G, we can find in polynomial time a set $\mathcal{W} := \{W_i\}_{i=1}^{s}$ of size $s = O\left(n^k\right)$, with $W_i \subseteq V$, and such that*

1. *There exists $i \in [s]$ such that W_i is a subset of a minimum vertex cover of G,*
2. *$|W_i| \geq \left(1-(1-\epsilon)^{\frac{1}{k}}\right)\frac{n}{k}$, $\forall i \in [s]$.*

Proof. Clearly, the algorithm depicted in figure 1 returns a set $\mathcal{W}$ of size $O\left(n^k\right)$ in polynomial time since we assumed k to be $O(1)$.

The first condition will be verified by induction:
If there exists a minimum vertex cover in which all vertices of H are contained, we are done. Otherwise, we obtain a vertex $v \in H$ that does not belong to any minimum vertex cover. Therefore, a minimum vertex cover of H must contain a minimum vertex cover of the $(k-1)$-partite $(k-1)$-uniform hypergraph H' since otherwise some of the edges will not be covered. By induction, the recursive call

Input: a 0-wise ϵ-dense balanced k-partite k-uniform hypergraph
$H = (V_1, .., V_k, E(H))$

1. if $k = 1$ then
   (a) return a minimum vertex cover of H, of size $|E(H)| \geq \epsilon n$
2. else:
   (a) let H be the set of the first $\left(1 - (1-\epsilon)^{\frac{1}{k}}\right) n$ highest-degree vertices (breaking ties arbitrarily)
   (b) add H to $\mathcal{W}$
   (c) for each $v \in H$:
      i. let H' be the $(k-1)$-partite $(k-1)$-uniform hypergraph $(V - V_b, \{e - \{v\} : e \in E, v \in e\})$, where $v \in V_b$.
      ii. let $\epsilon' := 1 - (1-\epsilon)^{\frac{k}{k+1}}$
      iii. call the procedure recursively, with the parameters $H', \epsilon', k-1$; let $\mathcal{W}'$ be its output
      iv. add the sets of $\mathcal{W}'$ to $\mathcal{W}$
   (d) return $\mathcal{W}$

Figure 1: Recursive algorithm for the Vertex Cover problem in 0-wise ϵ-dense k-partite k-uniform hypergraphs

returns one subset contained in a minimum vertex cover of H', hence also in a minimum vertex cover of H. The base case $k = 1$ is trivial.

We now prove the second property. This will be done by induction as well. For a fixed value of k, suppose that

$$|W_i| \geq \left(1 - (1-\epsilon)^{\frac{1}{k}}\right) \frac{n}{k}, \ \forall i \in [s]$$

holds for all balanced k-partite k-uniform hypergraphs. We now prove this property for $k+1$. From Lemma 2, the recursive calls are performed on ϵ'-dense hypergraphs with at least $n - \frac{n}{k+1}$ vertices. Thus, by the induction hypothesis, the recursive call returns a collection of sets W_i of size

$$\begin{aligned} |W_i| &\geq \left(1 - (1-\epsilon')^{\frac{1}{k}}\right) \left(\frac{n - \frac{n}{k+1}}{k}\right) \\ &= \left(1 - \left(1 - \left(1 - (1-\epsilon)^{\frac{k}{k+1}}\right)\right)^{\frac{1}{k}}\right) \left(\frac{n\frac{k}{k+1}}{k}\right) \\ &\geq \left(1 - (1-\epsilon)^{\frac{1}{k+1}}\right) \frac{n}{k+1}, \end{aligned}$$

as claimed. The base case $k = 1$ is verified, as in that case the procedure yields at least ϵn vertices. $\square$

Let us now consider the case $\ell > 0$. By definition, we study hypergraphs in which every subset of ℓ vertices is contained in $\epsilon \left(\frac{n}{k}\right)^{k-\ell}$ hyperedges. We prove a similar statement for ℓ-wise ϵ-dense balanced k-partite k-uniform hypergraphs.

Lemma 4. *Given a ℓ-wise ϵ-dense balanced k-partite k-uniform hypergraph G, we can find in polynomial time a set $\mathcal{W} := \{W_i\}_{i=1}^{s}$ of size $s = O\left(n^k\right)$, with $W_i \subseteq V$, and such that*

1. *There exists $i \in [s]$ such that W_i is a subset of a minimum vertex cover of G,*
2. *$|W_i| \geq \left(1 - (1-\epsilon)^{\frac{1}{k-\ell}}\right) \frac{n}{k}\ \forall i \in [s]$.*

Proof. Let $H = (V_1, .., V_k, E(H))$ be the considered hypergraph and S a subset of ℓ vertices that do not belong to a given minimum vertex cover C of H. Let us denote by H' the subhypergraph of H whose vertex set is $V(H') := V(H) \setminus \bigcup_{i\in S} V_i$, and whose hyperedges are the hyperedges of H containing S, restricted to $V(H) \setminus \bigcup_{i\in S} V_i$. By definition, S is not contained in C and therefore, C must contain a vertex cover of H'. Clearly, an ℓ-wise ϵ-dense hypergraph has at least $\epsilon \left(\frac{n}{k}\right)^{k-\ell}$ edges. We conclude that H' is a 0-wise ϵ-dense balanced $(k-\ell)$-partite $(k-\ell)$-uniform hypergraph with at least $n - \ell \cdot \frac{n}{k}$ vertices. We know from Lemma 3 that we can extract $O\left(n^{k-\ell}\right)$ candidates W_i, which are subsets of $V(H')$ of size at least $\left(1 - (1-\epsilon)^{\frac{1}{k-\ell}}\right) \frac{k-\ell}{k(k-\ell)} n$. One of them at least is contained in a minimum vertex cover of H', and therefore, in a minimum vertex cover of H. By enumerating all $O\left(n^\ell\right)$ possibilities for S, we obtain the result in time $O\left(n^k\right)$. □

Now, we are ready to prove Theorem 1. By testing all possible sets $W_i \in \mathcal{W}$ and choosing the one that yields the smallest cover, we obtain from Lemma 1 by setting $\delta = 1$ a polynomial-time approximation algorithm with approximation ratio at most

$$\frac{k}{2 + (k-2)\,|W| \left(\frac{n}{k} - j\right)^{-1}} = \frac{k}{2 + (k-2)\left(1 - (1-\epsilon)^{\frac{1}{k-\ell}}\right)} - o(1)$$

$$= \frac{k}{k - (k-2)\,(1-\epsilon)^{\frac{1}{k-\ell}}} - o(1)$$

in ℓ-wise ϵ-dense k-partite k-uniform hypergraphs.

3 Vertex Cover in Subdense k-Partite Hypergraphs

In this section, we study the Vertex Cover problem restricted to balanced k-partite k-uniform hypergraphs with $\frac{\bar{d}}{\Delta} = \Theta(1)$. For this case, we will explore how low the density condition can be relaxed while still achieving an approximation factor better than $k/2$. We propose a randomized approximation algorithm to tackle this case. In particular, we will prove the following result:

Theorem 2. *For every $\epsilon > 0$ and $k = O(1)$, there is a randomized approximation algorithm which computes with high probability a solution for the Vertex Cover problem in balanced k-partite k-uniform hypergraphs with approximation ratio*

$$\frac{k}{2 + (k-2)\,\frac{\bar{d}}{\Delta}} + \epsilon,$$

where $\bar{d}$ and $\Delta = \frac{n^{k-1}}{\psi(n)}$ denotes the average and maximum degree of the hypergraph, respectively. The running time is $n^{O(1)}2^{O(\psi(n))}$.

3.1 Overview of the algorithm

The algorithm iteratively collects disjoint sets V' of vertices from a hypergraph H. It removes the actual considered set V' and proceeds to collect vertices in the remaining graph $H[V(H) - V']$, until a sufficiently small set of vertices remain. Finally, it applies the $\left(\frac{k}{2}\right)$-approximation algorithm on the residual instance. The next subsection deals with an efficient sample algorithm, which performs the extraction of the vertex set $\mathcal{W}$. It will be a randomized version of the algorithm introduced in section 2. The union of the collected sets will define the set W allowing us to use Lemma 1. We aim at extracting such a set W of size approximately $\beta\frac{n}{k}$, with $\beta := \frac{\bar{d}}{\Delta}$.

Firstly, we introduce some notation which will be used. Let $H_i = (V^i, E^i)$ be the hypergraph considered at the ith step, by n_i we denote its number of vertices, by $\epsilon_i := \frac{|E^i|}{\left(\frac{n_i}{k}\right)^k}$ its density (in the $\ell = 0$ sense), by V_j^i the j-th partition of the vertex set V^i and by $\bar{d}_i := \frac{|E^i|k}{n_i}$ its average degree. We also define $\psi(n) := \frac{\left(\frac{n}{k}\right)^{(k-1)}}{\Delta}$. Let $s_i := \frac{n_i}{k} - (1-\beta)\,\frac{n}{k}$. Note that $s_i = 0 \Rightarrow \frac{n_i}{k} = (1-\beta)\,\frac{n}{k} \Rightarrow \frac{n-n_i}{k} = \beta\frac{n}{k}$. Since $\frac{n-n_i}{k}$ is the size of the extracted set W, s_i can serve as a measure of progress of the procedure. At every step, we remove $\frac{c^2 n}{k\psi(n)}$ vertices, until $s_i \le c\frac{n}{k}$, for a small constant $c \in (0, 1]$. Thus, at the end of the procedure, we will have

$$s_i \le c\frac{n}{k} \Rightarrow |W| \ge (\beta - c)\,\frac{n}{k}$$

In the remainder of this subsection, we show that we can always find a set of this size contained in a minimum vertex cover. From Lemma 3, we know that we can extract in every iteration i a set of size at least r_i, where

$$r_i := \left(1 - (1-\epsilon_i)^{\frac{1}{k}}\right)\frac{n_i}{k}.$$

We assume that we can efficiently guess this subset, and prove that it is large enough.

Lemma 5. *Provided $s_i \ge c\frac{n}{k}$, the following inequality holds:*

$$r_i \ge c^2\frac{n}{k\psi(n)}. \tag{7}$$

In order to prove the previous Lemma, we need the fact that $|E_i| \geq \Delta s_i k$ for all $i \geq 1$ $(+)$ and

$$1 - (1-\epsilon)^{\frac{1}{k}} \geq \frac{\epsilon}{k} \text{ for all } \epsilon \in [0,1] \text{ and } k \geq 1. \ (*)$$

The proofs of the two facts can be found in [5].

Proof of Lemma 5. Combining the two previous facts, we obtain:

$$\begin{aligned}
\frac{r_i}{s_i} &\geq \frac{\epsilon_i \cdot n_i}{k \cdot s_i \cdot k} \text{ from } (*) &\quad (8)\\
&= \frac{|E_i| \cdot n_i}{k^2 s_i \left(\frac{n_i}{k}\right)^k} &\quad (9)\\
&\geq \frac{\Delta s_i k \cdot n_i}{k^2 s_i \left(\frac{n_i}{k}\right)^k} \text{ from } (+) &\quad (10)\\
&= \frac{c\Delta}{\left(\frac{n_i}{k}\right)^{k-1}} \geq \frac{c\Delta}{\left(\frac{n}{k}\right)^{k-1}} = \frac{c}{\psi(n)} &\quad (11)\\
\Rightarrow r_i &\geq \frac{c \cdot s_i}{\psi(n)} \geq c^2 \frac{n}{k\psi(n)}. &\quad (12)
\end{aligned}$$

□

At this point, we know that we can extract $c^2 \frac{n}{k\psi(n)}$ vertices at each step. Hence, the number t of required steps is

$$t := \frac{(\beta - c)\frac{n}{k}}{c^2 \frac{n}{k\psi(n)}} = \frac{\psi(n)}{c}\left(\frac{\beta}{c} - 1\right). \qquad (13)$$

3.2 The recursive sampling procedure

In this subsection, we formulate a recursive sampling procedure called IR that guesses in every iteration i a small collection of candidate sets. The procedure IR is a sampling version of the procedure given in the previous section and forms the inner recursion of the whole algorithm. The outer recursion defined by the procedure ER iterates this extraction until $s_i \leq c\frac{n}{k}$. More precisely, it will perform exactly t external iterations. Initially, ER is called with $i = 0$, $t = \frac{\psi(n)}{c}\left(\frac{\beta}{c} - 1\right)$ and also uses a variable l, which sets the size of the sample.

We will fix a constant $p \in (0,1)$, and define the sample size l as

$$l := \left\lceil \frac{\log\left(1 - p^{\frac{1}{k}}\right)}{\log p} \right\rceil.$$

With this value of l, the procedure IR has the following property:

Input: a balanced k-partite k-uniform hypergraph $H = (V_1, .., V_k, E(H))$ and $l \in \mathbb{N}$

1. $\mathcal{W} \leftarrow \emptyset$
2. if $k = 1$ then
   (a) return $\{C\}$, where C is set of $\frac{c^2 n}{k\psi(n)}$ arbitrary vertices of $V(H)$
3. else:
   (a) let C be the set of the first $\frac{c^2 n}{k\psi(n)}$ highest-degree vertices (breaking ties arbitrarily)
   (b) add C to $\mathcal{W}$
   (c) let $C' \subseteq C$ be a random subset of l vertices
   (d) for each $v \in C'$:
      i. let H' be the $(k-1)$-partite $(k-1)$-uniform hypergraph $(V(H) \setminus V_b, \{e \setminus \{v\} : e \in E(H), v \in e\})$, where $v \in V_b$
      ii. call the procedure recursively with the parameters $k-1, H', l$; let $\mathcal{W}'$ be its output
      iii. add the sets of $\mathcal{W}'$ to $\mathcal{W}$
   (e) return $\mathcal{W}$

Figure 2: **Procedure** $IR(H, l)$ (Inner Recursion)

Lemma 6. *Let C be an arbitrary vertex cover of the input hypergraph, the set $\mathcal{W}$ returned by the procedure IR contains a subset W' such that $|W' \cap C| \geq p|W'|$ with probability at least p.*

Proof. Let B be the set with the first $\frac{c^2 n}{k\psi(n)}$ highest-degree vertices. If $|B \cap C| \geq p|B|$ holds, the statement is true. For the remainder we will assume that $|B \cap C| < p|B|$. In this case, the probability that a random vertex of H belongs to C is at most p. Consequently, the probability of sampling a vertex $v \notin C$ is at least $1 - p^l$. But then, C must contain a vertex cover of the hypergraph H' induced by v. By iterating the described process, we conclude that the probability is at least $\left(1 - p^l\right)^k$, which from the definition of l is $\geq p$. $\square$

The procedure ER performs a recursive exploration of a search tree, branching on every subset W' in the set of candidates $\mathcal{W}$. A root-to-leaf path in this tree yields a set W, defined as the union of all the candidates W' selected along the path. We now prove that with high probability, this search tree contains a path yielding a suitable set W.

Lemma 7. *For any $\delta > 0$, the procedure ER returns a set W of $(\beta - c)\frac{n}{k}$ vertices,* Input: a balanced k-partite k-uniform hypergraph H with $i \in [t]$ and $l \in \mathbb{N}$

1. $\mathcal{W} \leftarrow \emptyset$
2. if $i < t$ then:
   (a) $\mathcal{W} \leftarrow IR(H, l)$
   (b) return $\min\{W' \cup ER(H[V(H) \setminus W'], i+1) \mid W' \in \mathcal{W}\}$
3. else ($i = t$)
   (a) apply a $\left(\frac{k}{2}\right)$-approximation algorithm to H and let C be the resulting vertex cover
   (b) return C

Figure 3: **Algorithm** $ER(H, i, t)$ (Outer Recursion)

such that $|W \cap C| \geq (1-\delta)\, p^2 |W|$, *with probability at least*

$$1 - e^{-\frac{\psi(n)}{c}\left(\frac{\beta}{c}-1\right)p\frac{\delta^2}{2}}.$$

Proof. Let X_i be a $0/1$ random variable, which denotes the success in the ith step. We set $p(X_i = 1) = p$ and $p(X_i = 0) = 1 - p$. Furthermore, we introduce another random variable $X = \sum_{i \in [t]} X_i$. Now, we want to lower bound the expectation of X and we obtain:

$$E[X] \geq tp = \frac{\psi(n)}{c}\left(\frac{\beta}{c} - 1\right)p.$$

We know that exactly $\frac{c^2 n}{k\psi(n)}$ vertices are chosen at every step. Therefore, the expected number of vertices of W that are contained in C is

$$E[X]pc^2\frac{n}{k\psi(n)} \geq \left(\frac{\beta}{c} - 1\right)p^2 c\frac{n}{k} \tag{14}$$

$$= (\beta - c)\, p^2 \frac{n}{k}. \tag{15}$$

The claimed result is obtained by using Chernoff bounds. □

3.3 The approximation ratio

Let us analyze the achieved approximation ratio and establish the proof of Theorem 2 which follows directly from the previous lemmas.

Proof of Theorem 2. Firstly, we describe how to initialize the procedure ER to obtain the specified approximation upper bound:
Select a constant $c \in (0, 1)$, which can be arbitrarily small. Then, choose a $p \in (0, 1)$, that can be arbitrarily close to 1. Compute the predefined value of

the sample size l and the number of steps t. Finally, run the procedure ER with these parameters.
We know from Lemmas 7 and 1 that with probability at least $1 - e^{-\frac{\psi(n)}{c}\left(\frac{\beta}{c}-1\right)p\frac{\delta^2}{2}}$ the procedure ER achieves an approximation ratio

$$\frac{k}{2 + ((1-\delta)\,p^2k - 2)\,(\beta - c)}.$$

Clearly, the approximation ratio is arbitrarily close to $\frac{k}{(2+(k-2)\beta)}$ if we let $c \to 0$ and $p \to 1$.
Now, let us analyze the running time of procedure ER:
The procedure ER generates a search tree of height t and fan-out less than $(l+1)^k$. The procedure IR needs $O\left(n^{O(1)}\right) + O\left(l^k\right)$ time at every node of the tree. Hence, the overall running time of the algorithm is $O\left(n^{O(1)} \cdot l^{kt}\right)$. For the remainder of the consideration, we assume $\beta = \Theta(1)$ since it is the only way to obtain a better approximation ratio better than $\frac{k}{2}$. Then, we obtain $kt = k\frac{\psi(n)}{c}\left(\frac{\beta}{c} - 1\right) = \Theta(\psi(n))$ and $l = \Theta(\log k) = \Theta(1)$. Hence, the running time is $n^{O(1)}2^{O(\psi(n))}$, as claimed. $\square$

4 Inapproximability Result

In this section, we prove that achieved approximation upper bound of the former section is almost optimal in a specified range of Δ assuming the Unique Games Conjecture [22].

For this reason, we start with a reformulation of a theorem given in [11] and deals with the approximability of balanced k-partite k-uniform hypergraphs assuming the UGC.

Theorem 3. *[11] Given a balanced k-partite k-uniform hypergraph $H = (V, E)$, let OPT denote an optimal vertex cover of H. For every $\delta > 0$, the following is UGC-hard to decide:*

$$|V|\left(\frac{1}{2(k-1)} - \delta\right) \leq |OPT| \text{ or } |OPT| \leq |V|\left(\frac{1}{k(k-1)} + \delta\right)$$

As the starting point of our reduction, we will use Theorem 3 and prove the following inapproximability result:

Theorem 4. *Assuming the UGC, for every $c \in \mathbb{N}$, there no polynomial time approximation algorithm with an approximation ratio better than*

$$\frac{k}{2 + \frac{2(k-1)(k-2)\frac{\bar{d}}{\Delta}}{k+(k-2)\frac{\bar{d}}{\Delta}}}$$

by a constant for the Vertex Cover problem in balanced k-partite k-uniform hypergraphs with average degree $\bar{d}$, maximum degree $\Delta = \Omega\left(n^{\frac{k-1}{c}}\right)$ and $\Delta = o(n^{k-1})$.

Proof. We use a construction similar to that in Theorem 6 in [5]. In the remainder, we set $\epsilon = \frac{\bar{d}}{\Delta}$. We consider the graph $H = (V(H), E(H))$ from Theorem 3 with vertex partition $(V_1, .., V_k)$. Then, we construct a new hypergraph $H' = (V', E')$, which consists of $\left(1 - \frac{\bar{d}}{\Delta}\right) n$ disjoint copies of H and of $n\frac{\bar{d}}{\Delta}$ disjoint complete balanced k-partite k-uniform hypergraphs C_j of size n (k-partite cliques). Clearly, we have $n' := |V'| = n^2$. Let $V_i(C_j)$ be the i-th vertex partition of the j-th k-partite clique C_j and $V_i(H_j)$ be the i-th vertex partition of the j-th copy of H. By defining $V'_i := \bigcup_j V_i(C_j) \cup V_i(H_j)$ as the i-th partition of H', we obtain a balanced k-partite hypergraph. Furthermore, we join an edge e to E' only if e has an nonempty intersection with the first partition of a clique, i.e. $e \cap \left[\bigcup_j V_1(C_j)\right] \neq \emptyset$. By adding as many hyperedges as needed, we can make H' to have asymptotically the average degree $\bar{d} = \omega(n'^{\frac{k-1}{2}})$ and maximum degree $\Delta = \omega(n'^{\frac{k-1}{2}})$. Notice that the maximal degree which can be obtained in this way is $o(n^{2(k-1)})$. We see that a vertex cover of H' must include at least $\frac{n}{k}$ vertices of each clique. Let us now consider the two cases in the decision problem above. If a vertex cover of H requires $n\left(\frac{1}{2(k-1)} - \delta\right)$ vertices, we will need $\left(1 - \frac{\bar{d}}{\Delta}\right) n \cdot n \left(\frac{1}{2(k-1)} - \delta\right) + \left(\frac{n}{k}\right) \cdot n\frac{\bar{d}}{\Delta}$ vertices to obtain a vertex cover for each copy of H and for each clique C_j. In the other case, $\left(1 - \frac{\bar{d}}{\Delta}\right) n^2 \left(\frac{1}{k(k-1)} + \delta\right) + \left(\frac{n}{k}\right) n\frac{\bar{d}}{\Delta}$ vertices will suffice to cover H'. By denoting OPT' as an optimal vertex cover of H', the UGC-hard decision question from Theorem 3 becomes the following:

$$
\begin{aligned}
\left(1 - \frac{\bar{d}}{\Delta}\right) n^2 \left(\frac{1}{2(k-1)} - \delta\right) + \frac{n^2 \frac{\bar{d}}{\Delta}}{k} &\leq |OPT'| \\
&\text{or} \\
\left(1 - \frac{\bar{d}}{\Delta}\right) n^2 \left(\frac{1}{k(k-1)} + \delta\right) + \frac{n^2 \frac{\bar{d}}{\Delta}}{k} &\geq |OPT'|
\end{aligned}
$$

Hence, assuming the UGC, the above decision problem results in the hardness of approximating within a factor of:

$$
\begin{aligned}
\frac{\left(1 - \frac{\bar{d}}{\Delta}\right) n^2 \left(\frac{1}{2(k-1)} - \delta\right) + \frac{n^2 \frac{\bar{d}}{\Delta}}{k}}{\left(1 - \frac{\bar{d}}{\Delta}\right) n^2 \left(\frac{1}{k(k-1)} + \delta\right) + \frac{n^2 \frac{\bar{d}}{\Delta}}{k}} &= \frac{\frac{1-\frac{\bar{d}}{\Delta}}{2(k-1)} - \delta\left(1 - \frac{\bar{d}}{\Delta}\right) + \frac{\frac{\bar{d}}{\Delta}}{k}}{\frac{1-\frac{\bar{d}}{\Delta}}{k(k-1)} + \delta\left(1 - \frac{\bar{d}}{\Delta}\right) + \frac{\frac{\bar{d}}{\Delta}}{k}} \\
&= \frac{\frac{\left(1-\frac{\bar{d}}{\Delta}\right)k}{2(k-1)k} + \frac{2\frac{\bar{d}}{\Delta}(k-1)}{2k(k-1)}}{\frac{1-\frac{\bar{d}}{\Delta}}{(k-1)k} + \frac{\frac{\bar{d}}{\Delta}(k-1)}{k(k-1)}} - \delta' \\
&= \frac{\frac{k-\frac{\bar{d}}{\Delta}k+2\frac{\bar{d}}{\Delta}k-2\frac{\bar{d}}{\Delta}}{2(k-1)k}}{\frac{1-\frac{\bar{d}}{\Delta}+\frac{\bar{d}}{\Delta}k-\frac{\bar{d}}{\Delta}}{(k-1)k}} - \delta'
\end{aligned}
$$

$$\begin{aligned}
\frac{\frac{k-\frac{\bar{d}}{\Delta}k+2\frac{\bar{d}}{\Delta}k-2\frac{\bar{d}}{\Delta}}{2(k-1)k}}{\frac{1-\frac{\bar{d}}{\Delta}+\frac{\bar{d}}{\Delta}k-\frac{\bar{d}}{\Delta}}{(k-1)k}}-\delta' &= \frac{k+(k-2)\frac{\bar{d}}{\Delta}}{2\left(1+(k-2)\frac{\bar{d}}{\Delta}\right)}-\delta' \\
&= \frac{k}{\frac{2k\left(1+(k-2)\frac{\bar{d}}{\Delta}\right)}{k+(k-2)\frac{\bar{d}}{\Delta}}}-\delta' \\
&= \frac{k}{\frac{2k+2(k-2)\frac{\bar{d}}{\Delta}+(2k-2)(k-2)\frac{\bar{d}}{\Delta}}{k+(k-2)\frac{\bar{d}}{\Delta}}}-\delta' \\
&= \frac{k}{2+\frac{(2k-2)(k-2)\frac{\bar{d}}{\Delta}}{k+(k-2)\frac{\bar{d}}{\Delta}}}-\delta' \\
&= \frac{k}{2+\frac{2(k-1)(k-2)\frac{\bar{d}}{\Delta}}{k+(k-2)\frac{\bar{d}}{\Delta}}}-\delta'
\end{aligned}$$

If we use $O\left(n^c\right)$ copies of H with $c = O\left(1\right)$ in the construction of H', we can lower the maximum degree of the resulting hypergraph H' and the result follows.

□

References

[1] R. Aharoni, R. Holzman, and M. Krivelevich, *On a Theorem of Lovász on Covers in r-Partite Hypergraphs*, Combinatorica 16, pp. 149–174, 1996.

[2] S. Arora, D. Karger, and M. Karpinski, *Polynomial Time Approximation Schemes for Dense Instances of NP-Hard Problems*, Proc. 27th ACM STOC (1995), pp. 284–293; see also in J. Comput. Syst. Sci. 58, pp. 193–210, 1999.

[3] N. Bansal and S. Khot, *Inapproximability of Hypergraph Vertex Cover and Applications to Scheduling Problems*, Proc. 37th ICALP (2010), LNCS 6198, Springer, 2010, pp. 250–261.

[4] R. Bar-Yehuda and Z. Kehat, *Approximating the Dense Set-Cover Problem*, J. Comput. Syst. Sci. 69, pp. 547–561, 2004.

[5] J. Cardinal, M. Karpinski, R. Schmied, and C. Viehmann, *Approximating Vertex Cover in Dense Hypergraphs*, CoRR abs/1012.2573, 2010.

[6] A. Clementi and L. Trevisan, *Improved Non-Approximability Results for Minimum Vertex Cover with Density Constraints*, Theor. Comput. Sci. 225, pp. 113–128, 1999.

[7] I. Dinur, V. Guruswami, and S. Khot *Vertex Cover on k-Uniform Hypergraphs is Hard to Approximate within Factor* $(k-3-\epsilon)$, ECCC TR02-027, 2002.

[8] I. Dinur, V. Guruswami, S. Khot, and O. Regev, *A New Multilayered PCP and the Hardness of Hypergraph Vertex Cover*, SIAM J. Comput. 34, pp. 1129–1146, 2005.

[9] A. Eremeev, *On some Approximation Algorithms for Dense Vertex Cover Problem*, Proc. Symp. on Operations Research 1999, pp. 48–52.

[10] G. Gottlob and P. Senellart, *Schema Mapping Discovery from Data Instances*, J. ACM 57, 2010.

[11] V. Guruswami and R. Saket, *On the Inapproximability of Vertex Cover on k-Partite k-Uniform Hypergraphs*, Proc. 37th ICALP (2010), LNCS 6198, pp. 360–371, Springer, 2010.

[12] E. Halperin, *Improved Approximation Algorithms for the Vertex Cover Problem in Graphs and Hypergraphs* SIAM J. Comput. 31, pp. 1608–1623, 2002.

[13] J. Holmerin, *Improved Inapproximability Results for Vertex Cover on k-Uniform Hypergraphs*, Proc. 29th ICALP (2002), LNCS 2380, Springer, 2002, pp. 1005–1016.

[14] J. Holmerin, *Vertex Cover on 4-Regular Hyper-Graphs is Hard to Approximate within* $2-\epsilon$, Proc. 34th ACM STOC (2002), pp. 544–552.

[15] L. Ilie, R. Solis-Oba, and S. Yu, *Reducing the Size of NFAs by Using Equivalences and Preorders*, Proc. 16th CPM (2005), LNCS 3537, pp. 310–321, Springer, 2005.

[16] T. Imamura and K. Iwama, *Approximating Vertex Cover on Dense Graphs*, Proc. 16th ACM-SIAM SODA (2005), pp. 582–589.

[17] G. Karakostas, *A Better Approximation Ratio for the Vertex Cover Problem*, ACM Transactions on Algorithms 5, 2009.

[18] R. Karp, *On the Computational Complexity of Combinatorial Problems*, Networks 5, pp 45–68, 1975.

[19] M. Karpinski, *Polynomial Time Approximation Schemes for Some Dense Instances of NP-Hard Optimization Problems*, Algorithmica 30, pp. 386–397, 2001.

[20] M. Karpinski, A. Rucinski, and E. Szymanska, *The Complexity of Perfect Matching Problems on Dense Hypergraphs*, Proc. 20th ISAAC (2009), LNCS 5878, Springer, 2009, pp. 626–636.

[21] M. Karpinski and A. Zelikovsky, *Approximating Dense Cases of Covering Problems*, Proc. DIMACS Workshop on Network Design: Connectivity and Facilities Location (1997), pp. 169–178; also published in ECCC TR97-004, 1997.

[22] S. Khot, *On the Power of Unique 2-Prover 1-Round Games*, Proc. 34th ACM STOC (2002), pp. 767–775.

[23] S. Khot and O. Regev, *Vertex Cover Might be Hard to Approximate to within* $2 - \epsilon$, J. Comput. Syst. Sci. 74, pp. 335-349, 2008.

[24] M. Krivelevich, *Approximate Set Cover in Uniform Hypergraphs*, J. Alg. 25 (1), pp. 118–143, 1997.

[25] L. Lovász, *On Minimax Theorems of Combinatorics*, Doctoral Thesis, Mathematiki Lapok 26, pp. 209–264, 1975.

[26] M. Okun, *On Approximation of the Vertex Cover Problem in Hypergraphs*, Disc. Opt. 2(1), pp. 101–111, 2005.

[27] L. Trevisan, *Non-Approximability Results for Optimization Problems on Bounded Degree Instances*, Proc. 33rd ACM STOC (2001), pp. 453–461.