Approximating Vertex Cover in Dense Hypergraphs

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

Abstract

We consider the minimum vertex cover problem in hypergraphs in which every hyperedge has size k (also known as *minimum hitting set* problem, or *minimum set cover* with element frequency k). Simple algorithms exist that provide k-approximations, and this is believed to be the best possible approximation achievable in polynomial time. We show how to exploit density and regularity properties of the input hypergraph to break this barrier. In particular, we provide a randomized polynomial-time algorithm with approximation factor $k/(1+(k-1)\frac{\bar{d}}{k\Delta})$, where $\bar{d}$ and Δ are the average and maximum degree, respectively, and Δ must be $\Omega(n^{k-1}/\log n)$. The proposed algorithm generalizes the recursive sampling technique of Imamura and Iwama (SODA'05) for vertex cover in dense graphs. As a corollary, we obtain an approximation factor $k/(2-1/k)$ for subdense regular hypergraphs, which is shown to be the best possible under the unique games conjecture.

Keywords: Approximation Algorithms, Dense Hypergraphs, Vertex Cover, Inapproximability Bounds, Unique Games

1 Introduction

A *vertex cover* of a graph is a subset of its vertices hitting all edges. The *Minimum Vertex Cover problem* asks to find a minimum size vertex cover, and is one of the most classical NP-complete problems. A *k-uniform hypergraph* is a hypergraph whose edges are k-element subsets of his

*Université Libre de Bruxelles (ULB), CP212. 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

vertex set, and the *k-uniform Vertex Cover problem* is the problem of finding a minimum size vertex cover in a k-uniform hypergraph. This problem is equivalent to the Set Cover problem where each element of the universe occurs in exactly k sets. For $k = 2$, it is the classical Vertex Cover problem in graphs. A simple 2-approximation algorithm exists for this problem by constructing a maximal matching greedily. However, currently best known approximation algorithms can only achieve an approximation ratio of $2 - o(1)$ [8, 12]. The problem is k-approximable in k-uniform hypergraphs, by choosing a maximal set of nonintersecting edges, and picking all vertices in them. The best known approximation factor is $k - (k-1)\ln\ln n/\ln n$ and is due to Halperin [8].

On the inapproximability side, one of the first hardness result for the k-uniform Vertex Cover problem is due to Trevisan [20]. He obtained an inapproximability factor of $k^{1/19}$. Holmerin [9] proved that it is NP-hard to approximate within $k^{1-\epsilon}$, and in addition, in [10], that the 4-uniform Vertex Cover problem is NP-hard to approximate within $2 - \epsilon$. Dinur et al. proved a $(k - 3 - \epsilon)$ lower bound [5], later improved to $(k - 1 - \epsilon)$ [6].

In [16], Khot introduced the Unique Games Conjecture (UGC) as an approach to tackle the inapproximability of NP-hard optimization problems. Assuming the unique games conjecture, an inapproximability factor of $k - \epsilon$ for the k-uniform Vertex Cover problem is due to Khot and Regev [17] which implies that the achieved ratios are the best possible. Recently, Bansal and Khot [2] were able to show in a conceptually different manner that this UGC-based inapproximability bound also holds on k-uniform hypergraphs that are almost k-partite.

In order to break these complexity barriers, we propose to consider dense instances of the Vertex Cover problem in hypergraphs, as was done previously for many other problems [1, 14, 15, 13]. Typically, a graph is said to be dense whenever the number of edges is within a constant factor of n^2, where n is the number of vertices. The Vertex Cover problem in dense hypergraphs, defined as hypergraphs with $\Omega(\binom{n}{k})$ hyperedges, has been considered before by Bar-Yehuda and Kehat [3]. They proposed an approximation algorithm which achieves a better approximation ratio than k and showed that it is best possible under some assumptions similar to the UGC. To the authors' knowledge, this is the only result tackling the dense version of the Vertex Cover problem in hypergraphs.

However, dense instances of the Vertex Cover problem in graphs ($k = 2$) has been considered previously by Karpinski and Zelikovsky [15], Eremeev [7], Clementi and Trevisan [4], and later by Imamura and Iwama [11]. In the latter contribution, the authors give an approximation ratio parameterized by both the average and the maximum degree of the input graphs, that is strictly smaller than 2 whenever the ratio between the two is bounded, and the graph has average degree $\Omega(n/\log\log n)$. Other works tackling special instances of the Vertex Cover problem in hypergraphs include [18, 19].

1.1 Definitions

We use the notation $[i] := \{1, 2, \ldots, i\}$. Let S be a finite set and $k \in [|S|]$, we introduce the abbreviation $\binom{S}{k}$ for the set of all subsets $S' \subseteq S$ such that $|S'| = k$. A *k-uniform hypergraph* is a pair (V, E), where V is the vertex set, and E is a subset of $\binom{V}{k}$. We will usually set $n := |V|$ and $m := |E|$. In the remainder, unless stated explicitly, we will suppose that $k = O(1)$.

A *Vertex Cover* of a k-uniform hypergraph (V, E) is a set $C \subseteq V$ such that $e \cap C \neq \emptyset \ \forall e \in E$.

The minimum Vertex Cover problem consists of finding a vertex cover of minimum size in a given hypergraph.

The *degree* $d(v)$ of a vertex v is equal to $|\{e \in E : v \in e\}|$. Analogously, we define the degree $d(S)$ of a $S \subseteq V$ to be $|\{e \in E : S \subseteq e\}|$. We refer to the average and maximum degree of a vertex as $\bar{d}$ and Δ, respectively. A hypergraph is said to be *d-regular* whenever $d(v) = d \ \forall v \in V$, and *regular* whenever there exists a d such that it is d-regular.

We say that a k-uniform hypergraph is *ℓ-wise ϵ-dense* for $\ell + 1 \in [k]$ and $\epsilon \in [0, 1]$ whenever for every subset $S \in \binom{V}{\ell}$, we have $d(S) \geq \epsilon\binom{n-\ell}{k-\ell}$. Thus, for instance a 0-wise ϵ-dense k-uniform hypergraph is a hypergraph with at least $\epsilon\binom{n}{k}$ hyperedges, while a 1-wise ϵ-dense k-uniform hypergraph is such that every vertex is contained in at least $\epsilon\binom{n-1}{k-1}$ hyperedges. This definition naturally generalizes the notion of *weak* and *strong density* in graphs defined in previous works [15]. We call a k-uniform hypergraph *subdense* if $\bar{d} = \Omega\left(\frac{n^{k-1}}{\log n}\right)$ holds.

1.2 Our Results

We propose new approximation algorithms for the k-uniform Vertex Cover problem with approximation factors parameterized by the density and regularity parameters of the input hypergraph.

In 2004, Bar-Yehuda and Kehat [3] proved that the minimum vertex cover in 0-wise ϵ-dense k-uniform hypergraphs was approximable within a factor $k/(k - (k-1)(1-\epsilon)^{\frac{1}{k}})$. In the next section, we generalize this result to ℓ-wise ϵ-dense hypergraphs with $\ell > 0$. We also provide a shorter proof of their result.

In Section 3, we propose a randomized algorithm yielding an approximation factor $k/(1 + (k-1)\frac{\bar{d}}{k\Delta})$ on k-uniform hypergraphs with $\bar{d} = \Omega(n^{k-1}/\log n)$. This implies a $k/(2 - 1/k)$-approximation algorithm for subdense regular hypergraphs. The proposed algorithm is based on Imamura and Iwama's recursive sampling technique [11] for vertex cover in dense graphs. We generalize the technique to hypergraphs, and also improve its analysis.

Finally, in Section 4, we prove the optimality of our bounds for a specified range of Δ under the Unique Games Conjecture [17]. Essentially, we show that any improvement on the approximation factor of our algorithms would yield an approximation factor asymptotically smaller than k for the Vertex Cover problem in arbitrary k-uniform hypergraphs.

2 Approximating Vertex Cover in Dense Hypergraphs

We prove the following result:

Theorem 1. *The Vertex Cover problem is approximable in polynomial time within a factor*

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

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

In order to attain theorem 1, we provide several lemmas which are crucial to the proof of the main result in this section. We start with the following simple extension of a Lemma from Bar-Yehuda and Kehat [3].

Lemma 1. *Let $G = (V, E)$ be a k-uniform hypergraph with a minimum vertex cover C, and let $W \subseteq V$ such that $|W \cap C| \geq \delta|W|$ for some constant $\delta \in [0, 1]$. For every constant $j \in \mathbb{N}$, given W, we can compute an approximate solution to the Vertex Cover problem in G, with an approximation factor*

$$\frac{k}{1 + (\delta k - 1)\frac{|W|}{n-j}},$$

in polynomial time.

Proof. For a fixed $j \geq k$, we can check for all $i \in \{k-1, .., j\}$ and for all $S \in \binom{V}{n-i}$ whether S is a vertex cover of G in polynomial time. Thus, our solution can be assumed to be smaller than $n - j$. Then, the algorithm computes a k-approximation in the hypergraph G' induced by the edges that are not covered by W, and returns the union of this with W. Let us denote by S the returned solution, by C the considered optimal solution, by ρ the approximation factor $|S|/|C|$, and by C' the optimal solution of the Vertex Cover problem on G'. Clearly, we can assume $\frac{\rho|C|}{n-j} \leq 1$. Hence, we obtain:

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

□

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

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

Proof. We consider a hypergraph G containing $m \geq \epsilon\binom{n}{k}$ hyperedges. Let us denote by H the set of the first $(1 - (1 - \epsilon)^{\frac{1}{k}})n$ highest-degree vertices (breaking ties arbitrarily). Suppose the statement is not true. Then the number m of edges in G is strictly smaller than the number of edges in a hypergraph in which all vertices of H have degree $\binom{n-1}{k-1}$ (the maximum possible), and all the

remaining edges have degree exactly $(1-(1-\epsilon)^{\frac{k-1}{k}})\binom{n-1}{k-1}$. Thus

$$\begin{aligned} m &< \frac{1}{k}\left(|H|\binom{n-1}{k-1} + (n-|H|)(1-(1-\epsilon)^{\frac{k-1}{k}})\binom{n-1}{k-1}\right) && (7)\\ &= \frac{1}{k}\left((1-(1-\epsilon)^{\frac{1}{k}})n\binom{n-1}{k-1} + (n-(1-(1-\epsilon)^{\frac{1}{k}})n)(1-(1-\epsilon)^{\frac{k-1}{k}})\binom{n-1}{k-1}\right) && (8)\\ &= (1-(1-\epsilon)^{\frac{1}{k}})\binom{n}{k} + (1-\epsilon)^{\frac{1}{k}}(1-(1-\epsilon)^{\frac{k-1}{k}})\binom{n}{k} && (9)\\ &= \epsilon\binom{n}{k} && (10) \end{aligned}$$

which is a contradiction, since G is weakly ϵ-dense. □

We now prove Theorem 1, and first consider the case $\ell = 0$. In order to apply Lemma 1, we need to find a large subset W of a minimum vertex cover. The following recursive algorithm returns a polynomial-size collection $\mathcal{W}$ of subsets $W_i \subseteq V$, such that at least one of them is contained in a minimum vertex cover.

input: a 0-wise ϵ-dense k-uniform hypergraph $G = (V, E)$

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

Lemma 3. *Given a 0-wise ϵ-dense k-uniform hypergraph G, we can find in polynomial time a set $\mathcal{W} := \{W_i\}_{i=1}^{s}$ of size $s = O(n^k)$, 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. $|W_i| \geq (1-(1-\epsilon)^{\frac{1}{k}})(n-k+1),\ \forall i \in [s]$.

Proof. The above algorithm outputs of set $\mathcal{W}$ of size $O(n^k)$ in time $O(n^k)$, which is polynomial since we assumed $k = O(1)$.

The first condition is verified by induction. If all vertices in H belong to a minimum vertex cover, we are done. Otherwise, there exists $v \in H$ that does not belong to any minimum vertex cover. But then a minimum vertex cover of G must contain a minimum vertex cover of the $(k-1)$-uniform hypergraph G', as otherwise some edges will not be covered. By induction, the recursive call returns one subset contained in a minimum vertex cover of G', hence also in a minimum vertex cover of G. The base case $k = 1$ is trivial.

We prove the second property by induction as well. Suppose that $|W_i| \geq (1-(1-\epsilon)^{\frac{1}{k}})(n-k+1),\ \forall i \in [s]$ holds for k-uniform hypergraphs, for some fixed value of k. We now prove the property for $k+1$. From Lemma 2, the recursive calls are performed on weakly ϵ'-dense hypergraphs with $n-1$ vertices. Thus by the induction hypothesis, the recursive call returns a collection of sets W_i of size

$$\begin{aligned} |W_i| &\geq (1-(1-\epsilon')^{\frac{1}{k}})((n-1)-k+1) && (11)\\ &= (1-(1-(1-(1-\epsilon)^{\frac{k}{k+1}}))^{\frac{1}{k}})(n-(k+1)-1) && (12)\\ &= (1-(1-\epsilon)^{\frac{1}{k+1}})(n-(k+1)-1), && (13) \end{aligned}$$

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

We now tackle the case $\ell > 0$, i.e. hypergraphs in which every subset of ℓ vertices is contained in $\epsilon\binom{n-\ell}{k-\ell}$ hyperedges.

Lemma 4. *Given a ℓ-wise ϵ-dense k-uniform hypergraph G, we can find in polynomial time a set $\mathcal{W} := \{W_i\}_{i=1}^{s}$ of size $s = O(n^k)$, 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. *$|W_i| \geq (1-(1-\epsilon)^{\frac{1}{k-\ell}})(n-k+1),\ \forall i \in [s]$.*

Proof. Let V be the vertex set of G. Consider a subset S of ℓ vertices that do not belong to a given minimum vertex cover C of G. We define the hypergraph G' as the subhypergraph of G whose vertex set is $V' := V \setminus S$, and whose hyperedges are the hyperedges of G containing S, restricted to $V \setminus S$. Since S is not contained in C, C must contain a vertex cover of G'. From the definition of ℓ-wise ϵ-density, G' has at least $\epsilon\binom{n-\ell}{k-\ell}$ edges. Hence G' is a 0-wise ϵ-dense $(k-\ell)$-uniform hypergraph with $n-\ell$ vertices. From Lemma 3, we can extract $O(n^{k-\ell})$ candidates W_i, which are subsets of V' of size at least $(1-(1-\epsilon)^{\frac{1}{k-\ell}})(n-k+1)$. One of them at least is contained in a minimum vertex cover of G', and hence of G. By enumerating all $O(n^\ell)$ possibilities for S, we get the result in time $O(n^k)$. □

The proof of Theorem 1 is now straightforward. By testing all possible sets $W_i \in \mathcal{W}$ and choosing the one that yields the smallest cover, we get from Lemma 1 (with $\delta = 1$) a polynomial-time approximation algorithm with approximation factor

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

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

3 A Randomized Approximation Algorithm for Near-Regular Hypergraphs

We now prove a complementary result, yielding an approximation factor for the problem that is parameterized by both the average and the maximum degree of the input hypergraph. Our algorithm and analysis extend the work of Imamura and Iwama [11].

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

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

where $\bar{d}$ and Δ denote the average and maximum degree of the hypergraph, respectively. The running time is $n^{O(1)}2^{O(k\cdot\psi(n)\log\log k)}$, which is polynomial if $k\cdot\psi(n)\log\log k = O(\log n)$ and quasi-polynomial if $\psi(n) = \operatorname{polylog} n$.

By letting $\bar{d}/\Delta \to 1$ and $\bar{d} = \Omega\left(\frac{n^{k-1}}{\log n}\right)$, we get the following corollary.

Corollary 1. *The Vertex Cover problem is approximable in polynomial time within a factor $\frac{k}{2-1/k}$ in subdense regular k-uniform hypergraphs with $k = O(1)$.*

3.1 Outline

The algorithm first iteratively removes vertex subsets, until a sufficiently small set of vertices remain, then applies the trivial k-approximation algorithm on the remaining induced hypergraph. We first suppose that at every step i of the algorithm, we are able to guess a sufficiently large subset of an optimal solution of the current hypergraph G_i. This subset of vertices is removed, together with the edges that they cover, to form G_{i+1}. We will see in the next subsection how we can sample the set $\mathcal{W}$ computed in the algorithm of section 2 to perform this guessing step efficiently. The union of the removed sets will form the set W allowing us to use Lemma 1. We aim at obtaining such a set W of size approximately βn, with $\beta := \bar{d}/(k\Delta)$.

Letting G_i be the hypergraph considered at the ith step, we denote by n_i its number of vertices, by E_i its edge set, by $\epsilon_i := |E_i|/\binom{n_i}{k}$ its density (in the $\ell = 0$ sense), and by $\bar{d}_i := |E_i|k/n_i$ its average degree. We also let $\psi(n) := \binom{n}{k-1}/\Delta$. Let $s_i := n_i - (1-\beta)n$. Note that $s_i = 0 \Rightarrow n_i = (1-\beta)n \Rightarrow n - n_i = \beta n$. Since $n - n_i$ 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 $cn/\psi(n)$ vertices, until $s_i \le cn$, for a small constant $c \in [0,1]$. Thus at the end of the procedure, we will have $s_i \le cn \Rightarrow |W| \ge (\beta - c)n$. We now show we can always find a set of this size contained in a minimum vertex cover. From Lemma 3, we know there exists such a set of size at least r_i, where

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

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

Lemma 5. *provided $s_i \geq cn$, we have the following inequality:*

$$r_i \geq c\frac{n}{\psi(n)}. \tag{14}$$

We need three simple intermediate results.

Lemma 6. *Let H be a k-uniform hypergraph with n vertices, maximum degree Δ and maximum independent set of size at least x. Then H has at most $\Delta(n-x)$ edges.*

Proof. Every vertex covers at most Δ edges, hence the size of a minimum vertex cover of H, say τ, satisfies $|E(H)| \leq \Delta\tau$. Also, by definition, the largest independent set of H has size $n - \tau \geq x$, hence $\tau \leq n - x$. □

Lemma 7. $|E_i| \geq \Delta s_i$

Proof. The hypergraph $G' := G - E_i$ has an independent set of size at least $n_i = s_i + (1-\beta)n$, as by definition all the vertices of G_i form an independent set in G'. Thus from Lemma 6, G' can have at most $\Delta(n - (s_i + (1-\beta)n)) = \Delta(\beta n - s_i)$ edges. So we obtain

$$|E_i| = |E| - |E(G')| \geq \frac{\bar{d}n}{k} - \Delta(\beta n - s_i) = \Delta\beta n - \Delta(\beta n - s_i) = \Delta s_i.$$

□

Lemma 8. $1 - (1-\epsilon)^{\frac{1}{k}} \geq \frac{\epsilon}{k}\ \forall \epsilon \in [0,1], k \geq 1.$

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

$$\begin{aligned}
\frac{r_i}{s_i} &\geq \frac{\epsilon_i(n_i - k + 1)}{ks_i} \text{ (from Lemma 8)} && (15)\\
&= \frac{|E_i|(n_i - k + 1)}{ks_i\binom{n_i}{k}} && (16)\\
&\geq \frac{\Delta s_i(n_i - k + 1)}{ks_i\binom{n_i}{k}} \text{ (from Lemma 7)} && (17)\\
&= \frac{\Delta}{\binom{n_i}{k-1}} \geq \frac{\Delta}{\binom{n}{k-1}} = \frac{1}{\psi(n)} && (18)\\
r_i &\geq \frac{s_i}{\psi(n)} \geq c\frac{n}{\psi(n)}. && (19)
\end{aligned}$$

□

Thus we know we can extract this number of vertices at each step. The number t of required steps is therefore

$$t := \frac{(\beta - c)n}{c\frac{n}{\psi(n)}} = \psi(n)(\beta/c - 1). \tag{20}$$

3.2 The recursive sampling procedure

We now define a recursive sampling procedure that will allow us to efficiently guess the subsets. The procedure IR returns a small set of candidate subsets. It is a sampling version of the procedure given in the previous section.

Procedure $IR(G, l)$ (Inner Recursion)

Input: a k-uniform hypergraph $G = (V, E)$ and $l \in \mathbb{N}$

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

The procedure ER below iterates this extraction until $s_i \leq cn$, using t recursion levels. It is initially called with $i = 0$ and $t = \psi(n)(\beta/c - 1)$. It also uses a variable l, which sets the size of the sample used.

Algorithm $ER(G, i)$ (Outer Recursion)

Input: a k-uniform hypergraph G with $i \in \mathbb{N}$, $i \leq t$

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

Let us choose a constant $p \in (0,1)$, and define the sample size l as

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

Note that $1 - p^{\frac{1}{k}} = \Theta(\frac{1}{k})$, hence $l = \Theta(\log k)$. With this value of l, the procedure IR has the following property. We denote by C an arbitrary vertex cover of the input hypergraph.

Lemma 9. *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 H be the first $cn/\psi(n)$ highest-degree vertices. If $|H \cap C| \geq p|H|$ then we are done. Thus we can suppose that $|H \cap C| < p|H|$, and the probability that a random vertex of H belongs to C is at most p. Thus with probability at least $1 - p^l$ we get a vertex $v \notin C$ in the selected sample, and C must contain a vertex cover of the hypergraph G' defined by v. By iterating, we eventually get that the probability is at least $(1 - p^l)^k \geq p$, from the definition of l. □

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 10. *For any $\delta > 0$, the procedure ER constructs a set W of $(\beta - c)n$ vertices, such that $|W \cap C| \geq (1 - \delta)p^2|W|$, with probability at least $1 - e^{-\psi(n)(\beta/c-1)p\frac{\delta^2}{2}}$.*

Proof. We introduce the random variable X_i denoting the success in the ith step. Therefore, we have $p(X_i = 1) = p$ and $p(X_i = 0) = 1 - p$. We let $X = \sum_{i \in [t]} X_i$. Clearly, we have

$$E[X] \geq tp = \psi(n)(\beta/c - 1)p.$$

Since $cn/\psi(n)$ vertices are chosen at every step, the expected number of vertices of W that are contained in C is

$$E[X]pc\frac{n}{\psi(n)} \geq (\beta/c - 1)p^2 cn \quad (21)$$

$$= (\beta - c)p^2 n. \quad (22)$$

The claimed statement is obtained using Chernoff bounds. □

3.3 Assembling the pieces

The proof of Theorem 2 follows directly from the previous lemmas.

Proof of Theorem 2. The algorithm is as follows. First, select a constant $c < 1$, arbitrarily small, and a probability p, that can be arbitrarily close to 1. Then compute the corresponding value of the sample size l and the number of steps t, and run the procedure ER with these parameters. From Lemmas 10 and 1, with probability at least $1 - e^{-\psi(n)(\beta/c-1)p\frac{\delta^2}{2}}$, the procedure ER achieves an approximation ratio

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

When $c \to 0$ and $p \to 1$, this ratio is arbitrarily close to $k/(1+(k-1)\beta)$.

The procedure ER generates a search tree of height t and fan-out less than l^k. At every node of the tree, the procedure IR is called, taking $O(n^{O(1)}) + O(l^k)$ time. The overall running time of the procedure is therefore $O(n^{O(1)} \cdot l^{kt})$. We have $kt = k\psi(n)(\beta/c - 1) = \Theta(k\psi(n))$ and $l = \Theta(\log k)$. Hence the running time is $n^{O(1)}2^{O(k\psi(n)\log\log k)}$, as claimed. □

4 Lower bounds

In this section, we provide several inapproximability results based on two well-known conjectures, namely the UGC and $P \neq NP$. In particular, we show that the achieved approximation ratios in section 2 and 3 are optimal in a specified range of Δ assuming the Unique Game Conjecture.

We introduce two hardness results on which our inapproximability bounds are based. The following hardness result is due to Khot and Regev [17] and is underlying the UGC.

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

$$|V|(1-\delta) \leq |OPT| \text{ or } |OPT| \leq |V|(1/k + \delta)$$

On the other hand, Dinur et al. [6] proved the following NP-hardness result.

Theorem 4. *Given a k-uniform hypergraph $H = (V, E)$ with $k \geq 3$, let OPT denote an optimal vertex cover of H. For every $\delta > 0$, the following is NP-hard to decide:*

$$|V|(1-\delta) \leq |OPT| \text{ or } |OPT| \leq |V|(1/(k-1) + \delta)$$

The next theorem shows that the approximation algorithm from theorem 1 is optimal assuming the UGC.

Theorem 5. *Assuming the UGC and $P \neq NP$, respectively, there is no polynomial time algorithm with an approximation ratio better than*

$$\frac{k}{k + (k-1)(1-\epsilon)^{\frac{1}{k-\ell}}} \quad \text{and for any } k \geq 3 \quad \frac{k-1}{k-1 + (k-2)(1-\epsilon)^{\frac{1}{k-\ell}}},$$

respectively, by a constant for the Vertex Cover problem in ℓ-wise ϵ-dense k-uniform hypergraphs.

Proof. As a starting point of the reduction, we use the hypergraph $H = (V, E)$ from theorem 3 and construct a ℓ-wise ϵ-dense k-uniform hypergraph $H' = (V', E')$. We introduce the abbreviations $N := |V'|$ and $n := |V|$. First, we join a clique $C = (V(C), \binom{V(C)}{k})$ of size $c \cdot N$ (that is, $c \cdot N$ vertices with all possible hyperedges of size k) to H. Furthermore, we add all hyperedges $e \in \binom{C \cup V}{k}$, such that at least one vertex in e is from C. Thus, we obtain $N = cN + n$ and $N = n/(1-c)$. We denote by OPT' an optimal vertex cover of H'. The UG-hard decision question from theorem 3 transforms into the following:

$$n(1-\delta) + \frac{cn}{1-c} \leq |OPT'| \text{ or } |OPT'| \leq n(\frac{1}{k} + \delta) + \frac{cn}{1-c}$$

Assuming the UGC, this implies the hardness of approximating the Vertex Cover problem in ℓ wise ϵ-dense hypergraphs for every $\delta' > 0$ to within:

$$\frac{n(1-\delta)+\frac{cn}{1-c}}{n(1/k+\delta)+\frac{cn}{1-c}} = \frac{n(1-\delta)(1-c)+nc}{n(1/k+\delta)(1-c)+cn} = \frac{1-\delta(1-c)}{1/k-c/k+c+\delta/k(1-c)} \quad (23)$$

$$= \frac{k}{1+(k-1)c} - \delta' \quad (24)$$

By using Theorem 4, an analog calculation leads to the fact that it is NP-hard to approximate the Vertex Cover problem in ℓ wise ϵ-dense hypergraphs for every $\delta' > 0$ to within:

$$\frac{k-1}{1+(k-2)c} - \delta' \quad (25)$$

In order to obtain a ℓ-wise ϵ-dense k-uniform hypergraph, we have to determine the right value for c. Let $S \in \binom{V}{\ell}$ be a subset of V, and $d(S)$ denote its degree in H'. We can assume that $\min\{d(S') \mid S' \in \binom{V'}{\ell}\} = \min\{d(S) \mid S \in \binom{V}{\ell}\}$. Hence, it suffices to ensure that $d(S) \geq \epsilon\binom{N-\ell}{k-\ell} = \epsilon\frac{(N-\ell)^{k-\ell}}{(k-\ell)!} - o\left(\frac{(N-\ell)^{k-\ell}}{(k-\ell)!}\right)$ holds. We obtain

$$\begin{aligned} d(S) &\geq \binom{N-\ell}{k-\ell} - \binom{N-\ell-cN}{k-\ell} \\ &\geq \frac{(N-k+1)^{k-\ell}}{(k-\ell)!} - \frac{(N-\ell-cN)^{k-\ell}}{(k-\ell)!} \\ &= \frac{N^{k-\ell}[(1-(k+1)/N)^{k-\ell} - (1-\ell/N-c)^{k-\ell}]}{(k-\ell)!} \end{aligned}$$

By comparing the factors, we get $(1-(k+1)/N)^{k-\ell} - (1-\ell/n'-c)^{k-\ell} = \epsilon - o(1)$. Therefore, we deduce that H' is ℓ-wise ϵ-dense for

$$c = 1 - [(1-k/n')^{k-\ell} - \epsilon]^{\frac{1}{k-\ell}} + o(1) = 1 - (1-\epsilon)^{\frac{1}{k-\ell}} + o(1).$$

Plugging c into equation (24), we obtain the following inapproximability factor R assuming the UGC.

$$\begin{aligned} R &= \frac{k}{1+(k-1)c} - \delta' \\ &= \frac{k}{k+(k-1)(1-\epsilon)^{\frac{1}{k-\ell}}} - o(1) - \delta' \end{aligned}$$

Analogously, by using equation (25) instead, we get for every $\delta' > 0$ the inapproximability factor

$$\frac{k-1}{1+(k-2)(1-\epsilon)^{\frac{1}{k-\ell}}} - o(1) - \delta'$$

assuming $P \neq NP$.

□

The next theorem generalizes the former result in the sense that the inapproximability factor is parameterized by the maximum degree Δ of the hypergraph. In addition, it shows that the approximation algorithm from theorem 2 is optimal in a specified range of Δ assuming the UGC.

Theorem 6. *Assuming the UGC and $P \neq NP$, respectively, for every $c \in \mathbb{N}$, there is a $c_2 \in [0,1]$ such that no polynomial time algorithm can find a solution with an approximation ratio better than*

$$\frac{k}{1+(k-1)\,\frac{\bar{d}}{k\Delta}} \text{ and for any } k \geq 3 \quad \frac{k-1}{1+(k-2)\,\frac{\bar{d}}{k\Delta}},$$

respectively, by a constant for the Vertex Cover problem k-uniform hypergraphs with average degree $\bar{d}$, maximum degree $\Delta = \Omega(n^{\frac{k-1}{c}})$ and $\Delta \leq c_2 \frac{n^{k-1}}{k^{k-1}(k-1)!}$.

Proof. We proceed as before, by considering the hypergraph $H = (V, E)$ of Theorem 3, with $n := |V|$. Since the only way to achieve a better factor than k is in case of $\frac{\bar{d}}{\Delta} \in (0,1]$, we will set $\frac{\bar{d}}{\Delta} := \epsilon \in (0,1]$ in the remainder. We construct a new hypergraph $H' = (V', E')$ consisting of $(1-\epsilon/k)n$ disjoint copies of H, together with $n\epsilon/k$ disjoint complete k-uniform hypergraphs of size $n+k$ (cliques). We let V_1 be the set of vertices in the copies of H, and V_2 the set of vertices of the cliques, and $n' := |V'| = |V_1| + |V_2| = n^2 + o(n^2)$. Note that the degrees of the vertices restricted to neighbors in V_1 or V_2 are at most $\binom{n+k-1}{k-1} = O(n^{k-1}) = O(\sqrt{n'^{k-1}})$. Hence, we can make H' to have asymptotically the average degree $\bar{d} = \omega(n'^{\frac{k-1}{2}})$ and maximum degree $\Delta = \omega(n'^{\frac{k-1}{2}})$ by adding as many hyperedges as needed with vertices in both V_1 and V_2, thus defining E'. By construction, the maximal degree which can be obtained is proportional to $\binom{n^2\epsilon/k}{k-1}$.

By definition, a vertex cover of H' must contain at least $n+k-k = n$ vertices of each clique. Thus, we need to include at least $n^2\epsilon/k$ vertices. We now consider the two cases in the decision problem above. If a vertex cover of H requires $n(1-\delta)$ vertices, then we need $(1-\epsilon/k)n \cdot n(1-\delta) = (1-\epsilon/k)n^2(1-\delta)$ additional vertices to cover all the copies. In the other case, $(1-\epsilon/k)n^2(1/(k-1)+\delta)$ vertices suffice. Up to a $O(n)$ term, those vertices are sufficient to cover H'. Denoting by OPT' an optimal vertex cover of H', the UG-hard decision question from theorem 4 therefore becomes:

$$(1-\epsilon/k)\,n^2(1-\delta) + \frac{n^2\epsilon}{k} + o(n^2) \leq |OPT'| \text{ or } (1-\epsilon/k)\,n^2(1/(k-1)+\delta) + \frac{n^2\epsilon}{k} + o(n^2) \geq |OPT'|$$

Under the UGC, this implies the hardness of approximating within a factor:

$$\begin{aligned}
\frac{(1-\epsilon/k)\,n^2(1-\delta) + \frac{n^2\epsilon}{k}}{(1-\epsilon/k)\,n^2(1/(k-1)+\delta) + \frac{n^2\epsilon}{k}} &= \frac{(k-1)\,(1-\epsilon/k)\,(1-\delta) + (k-1)\frac{\epsilon}{k}}{(k-1)\,(1-\epsilon/k)\,(1/(k-1)+\delta) + (k-1)\frac{\epsilon}{k}} && (26)\\
&= \frac{k-1}{1-\epsilon/k+(k-1)\epsilon/k} - \delta' && (27)\\
&= \frac{k-1}{1+(k-2)\epsilon/k} - \delta'. && (28)
\end{aligned}$$

Since we can add $O(n^c)$ copies of H with an arbitrary $c = O(1)$ to construct H', the result follows. □

A similar construction with $\epsilon \to 1$ yields the following.

Theorem 7. *Assuming the UGC and $P \neq NP$, respectively there is no polynomial time algorithm yielding an approximation ratio better than $\frac{k}{2-1/k}$ and for $k \geq 3$ than $\frac{k}{2}$, respectively, by a constant for the Vertex Cover problem in k-uniform regular hypergraphs.*

References

[1] 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.

[2] 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.

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

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

[5] 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.

[6] 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.

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

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

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

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

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

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

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

[14] 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.

[15] 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.

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

[17] 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.

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

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

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