On the Inapproximability of Vertex Cover on k-Partite k-Uniform Hypergraphs*

Venkatesan Guruswami and Rishi Saket

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213.

Abstract. Computing a minimum vertex cover in graphs and hypergraphs is a well-studied optimizaton problem. While intractable in general, it is well known that on bipartite graphs, vertex cover is polynomial time solvable. In this work, we study the natural extension of bipartite vertex cover to hypergraphs, namely finding a small vertex cover in k-uniform k-partite hypergraphs, when the k-partition is given as input. For this problem Lovász [16] gave a $\frac{k}{2}$ factor LP rounding based approximation, and a matching $\left(\frac{k}{2} - o(1)\right)$ integrality gap instance was constructed by Aharoni *et al.* [1]. We prove the following results, which are the first strong hardness results for this problem (here $\varepsilon > 0$ is an arbitrary constant):

- NP-hardness of approximating within a factor of $\left(\frac{k}{4} - \varepsilon\right)$, and
- Unique Games-hardness of approximating within a factor of $\left(\frac{k}{2} - \varepsilon\right)$, showing optimality of Lovász's algorithm under the Unique Games conjecture.

The NP-hardness result is based on a reduction from minimum vertex cover in r-uniform hypergraphs for which NP-hardness of approximating within $r-1-\varepsilon$ was shown by Dinur *et al.* [5]. The Unique Games-hardness result is obtained by applying the recent results of Kumar *et al.* [15], with a slight modification, to the LP integrality gap due to Aharoni *et al.* [1]. The modification is to ensure that the reduction preserves the desired structural properties of the hypergraph.

1 Introduction

A k-uniform hypergraph $G = (V, E)$ consists of a set of vertices V and hyperedges E where every hyperedge is a set of exactly k vertices. The hypergraph G is said to be m-colorable if there is a coloring of the vertex set V with at most m colors such that no hyperedge in E has all its vertices of the same color. We shall be interested in the stricter condition of *strong* colorability as defined in Aharoni *et al.* [1], wherein G is said to be m-strongly-colorable if there is an m-coloring such of the vertex set V such that every hyperedge E has k distinctly colored vertices. In particular a k-strongly-colorable k-uniform hypergraph is a k-partite

* Research supported in part by a Packard Fellowship and US-Israel BSF-2008293.

k-uniform hypergraph, where the k-partition of the vertex set corresponds to the k color classes.

A *vertex cover* of a hypergraph $G = (V, E)$ is a subset V' of vertices such that every hyperedge in E contains at least one vertex from V'. The problem of computing the vertex cover of minimum size in a (hyper)graph has been deeply studied in combinatorics with applications in various areas of optimization and computer science. This problem is known to be NP-hard. On the other hand, for k-uniform hypergraphs the greedy algorithm of picking a maximal set of disjoint hyperedges and including all the vertices in those hyperedges gives a factor k approximation. More sophisticated algorithmic techniques only marginally improve the approximation factor to $k - o(1)$ [9].

Several inapproximability results have been shown for computing the minimum vertex cover. For general k, an $\Omega(k^{1/19})$ hardness factor was first shown by Trevisan [21], subsequently strengthened to $\Omega(k^{1-\varepsilon})$ by Holmerin [10] and to a $k - 3 - \varepsilon$ hardness factor due to Dinur, Guruswami and Khot [4]. The currently best known $k - 1 - \varepsilon$ hardness factor is due to Dinur, Guruswami, Khot and Regev [5] who build upon [4] and the seminal work of Dinur and Safra [6] who showed the best known 1.36 hardness of approximation for vertex cover in graphs ($k = 2$).

All of the above mentioned results are based on standard complexity assumptions. However, assuming Khot's Unique Games Conjecture (UGC) [12], an essentially optimal $k - \varepsilon$ hardness of approximating the minimum vertex cover on k-uniform hypergraphs was shown by Khot and Regev [14]. In more recent works the UGC has been used to relate the inapproximability of various classes of constraint satisfaction problems (CSPs) to the corresponding semi-definite programming (SDP) integrality gap [19], or the linear programming (LP) integrality gap [17] [15]. The recent work of Kumar *et al.* [15] generalizes the result of [14] and shall be of particular interest in this work.

In this work we investigate the complexity of computing the minimum vertex cover in hypergraphs that are strongly colorable and where the strong coloring is given as part of the input. Variants of this problem are studied for databases related applications such as distributed data mining [7], schema mapping discovery [8] and in optimizing finite automata [11]. The particular case of computing the minimum vertex cover in k-uniform k-partite (with the partition given) hypergraphs was studied by Lovász [16] who obtained a $k/2$ approximation for it by rounding its natural LP relaxation. Subsequently, Aharoni, Holzman and Krivelevich [1] proved a tight integrality gap of $k/2 - o(1)$ for the LP relaxation. On the hardness side, [11] and [8] give reductions from 3SAT to it, which imply that the problem is APX-hard. However, to the best of our knowledge no better hardness of approximation was known for this problem.

In this work we show a $\left(\frac{k}{4} - \varepsilon\right)$ hardness of approximation factor for computing the minimum vertex cover on k-uniform k-partite hypergraphs. Actually, we prove a more general hardness of approximation factor of $\frac{(m-(k-1))(k-1)}{m} - \varepsilon$ for computing the minimum vertex cover in m-strongly colorable k-uniform hypergraphs. The result for k-uniform k-partite hypergraphs follows by a simple reduction. Our results are based on a reduction from minimum vertex cover in k-uniform hypergraphs for which, as mentioned above, the best known factor $k - 1 - \varepsilon$ hardness of approximation factor was given in [5].

We also study the results of [15] in the context of the problems we consider. In [15], the authors proved that LP integrality gaps for a large class of monotone constraint satisfaction problems, such as vertex cover, can be converted into corresponding UGC based hardness of approximation results. As presented, the reduction in [15] does not guarantee that the structural properties of the integrality gap will be carried through into the final instance. Nevertheless, we observe that the integrality gap instance of [1] can be combined with the work of [15] with only a slight modification to yield an essentially optimal $k/2 - o(1)$ factor hardness of approximation for computing the minimum vertex cover in k-uniform k-partite hypergraphs, i.e. the final instance is also guaranteed to be a k-uniform k-partite hypergraph. Similar tight inapproximability can also be obtained for a larger class of hypergraphs which we shall define later.

Main Results. We summarize the main results of this paper in the following informal statement.

Theorem. *(Informal) For every $\varepsilon > 0$, and integers $k \geqslant 3$ and $m \geqslant 2k$, it is NP-hard to approximate the minimum vertex cover on m-strongly-colorable k-uniform hypergraphs to within a factor of*

$$\frac{(m - (k - 1))(k - 1)}{m} - \varepsilon.$$

In addition, it is NP-hard to approximate the minimum vertex cover on k-uniform k-partite hypergraphs to within a factor of $\frac{k}{4} - \varepsilon$, and within a factor of $\frac{k}{2} - \varepsilon$ assuming the Unique Games conjecture.

We now proceed to formally defining the problems we consider, followed by a discussion of the previous work and a precise statement of our results on these problems.

2 Problem Definitions

We now define the variants of the hypergraph vertex cover problem studied in this paper.

Definition 1. *For any integer $k \geqslant 2$, an instance $G = (V, E)$ of the hypergraph vertex cover problem* HypVC(k), *is a k-uniform hypergraph (possibly weighted) where the goal is to compute a vertex cover $V' \subseteq V$ of minimum weight.*

Definition 2. *For any integers $m \geqslant k \geqslant 2$, an instance of $G = (V, E)$ of* StrongColored-HypVC(m, k) *is a m-strongly-colorable k-uniform hypergraph where the m-strong-coloring of G is given. Formally, a partition of V into m disjoint subsets (color classes) V_i $(1 \leqslant i \leqslant m)$ is given, such that every hyperedge in E has at most one vertex from each color class. In other words, every hyperedge contains k distinctly colored vertices. The goal is to compute the minimum weight vertex cover in G.*

Definition 3. *For any integer $k \geqslant 2$, an instance $G = (V, E)$ of* HypVCPartite(k) *is an k-uniform k-partite hypergraph with the k-partition given as input. The goal is to compute the minimum weight vertex cover in G. Note that* HypVCPartite(k) *is the same as* StrongColored-HypVC(k, k).

The following definition generalizes the class of k-partite hypergraphs and defines the minimum vertex cover problem for that class.

Definition 4. *For any integer $k \geqslant 2$ and positive integers $p_1, \ldots, p_k$, a hypergraph $G = (V, E)$ is called $(p_1, \ldots, p_k)$-split if there is a k-partitioning $V_1, \ldots, V_k$ of the vertex set V such that for every hyperedge $e \in E$, $|e \cap V_i| = p_i$ for $1 \leqslant i \leqslant k$.* HypVCsplit$(r, k, p_1, \ldots, p_i)$ *denotes the problem of computing the minimum vertex cover in $(p_1, \ldots, p_k)$-split r-uniform hypergraphs where the k-partitioning is given as input. Here $\sum_{i=1}^{k} p_i = r$. Note that* HypVCPartite(k) *is the same as* HypVCsplit$(k, k, 1, \ldots, 1)$.

3 Previous work and our results

3.1 Previous Results

Let LP_0 be the natural "covering" linear programming relaxation for hypergraph vertex cover (see, for example, Section 1 of [1]). The linear program is oblivious to the structure of the hypergraph and can be applied to any of the variants of hypergraph vertex cover defined above. The following theorem, first proved by Lovász [16] gives an upper bound on the integrality gap of the relaxation LP_0 for HypVCPartite(k). All the upper bounds on the integrality gap stated in this section are achieved using polynomial time rounding procedures for LP_0.

Theorem 1. *(Lovász [16]) For any integer $k \geqslant 2$, for any instance G of* HypVCPartite(k),

$$\frac{\mathrm{OPT}_{\mathrm{VC}}(G)}{\mathrm{VAL}_{\mathrm{LP}_0}(G)} \leqslant \frac{k}{2} \tag{1}$$

where $\mathrm{OPT}_{\mathrm{VC}}(G)$ is the weight of the minimum vertex cover in G and $\mathrm{VAL}_{\mathrm{LP}_0}(G)$ is the optimum value of the objective function of the relaxation LP_0 applied to G.

We observe that the relaxation LP_0 does not utilize the k-partiteness property of the input hypergraph. Therefore, the upper bound in Equation (1) holds irrespective of whether the k-partition is given as input. On the other hand, the k-partition is necessary for the efficient rounding algorithm given by the previous theorem. We note that for general k-uniform hypergraphs the gap between the size of the minimum vertex cover and value of the LP solution can be as high as $k - o(1)$. The following theorem states that Equation (1) is essentially tight.

Theorem 2. *(Aharoni* et al. *[1]) The integrality gap of LP_0 on instances of* HypVCPartite(k) *is $k/2 - o(1)$.*

For instances of StrongColored-HypVC(m,k) Aharoni *et al.* [1] proved lower and upper bounds summarized in the following theorem.

Theorem 3. *(Aharoni et al. [1]) Let G be an instance of* StrongColored-HypVC(m,k) *where $m,k \geqslant 2$. If $m \geqslant k(k-1)$ then,*

$$\frac{\mathrm{OPT_{VC}}(G)}{\mathrm{VAL_{LP_0}}(G)} \leqslant \frac{(m-k+1)k}{m} \tag{2}$$

In addition, the integrality gap of LP_0 *when $m \geqslant k(k-1)$ is* $\frac{(m-k+1)k}{m} - o(1)$*. On the other hand, when $k < m < k(k-1)$ then,*

$$\frac{\mathrm{OPT_{VC}}(G)}{\mathrm{VAL_{LP_0}}(G)} \leqslant \frac{mk}{m+k} + \min\left\{\frac{m-k}{2m}a, \frac{k}{m}(1-a)\right\}, \tag{3}$$

where $a = \frac{m^2}{m+r} - \left\lfloor \frac{m^2}{m+r} \right\rfloor$.

Theorems 1 and 2 were generalized by [1] to split hypergraphs as defined in Definition 4. Their general result is stated below.

Theorem 4. *(Aharoni et al. [1]) For any positive integers $r, k, p_1, \ldots, p_k$ such that $\sum_{i=1}^{k} p_i = r \geqslant 2$, and any instance G of* HypVCsplit$(r,k,p_1,\ldots,p_k)$,

$$\frac{\mathrm{OPT_{VC}}(G)}{\mathrm{VAL_{LP_0}}(G)} \leqslant \max\left\{\frac{r}{2}, p_1, \ldots, p_k\right\}. \tag{4}$$

In addition, the integrality gap of LP_0 *on instances of* HypVCsplit$(r,k,p_1,\ldots,p_k)$ *is* $\max\left\{\frac{r}{2}, p_1, \ldots, p_k\right\} - o(1)$.

The following theorem states the best known NP-hardness of approximation for the minimum vertex on general hypergraphs.

Theorem 5. *(Dinur et al. [5]) For any $\varepsilon > 0$ and integer $k \geqslant 3$, it is NP-hard to approximate* HypVC(k) *to a factor of $k - 1 - \varepsilon$.*

The above hardness of approximation for general k is not known to be tight. On the other hand, assuming the Unique Games Conjecture one can obtain optimal inapproximability factors of $k - o(1)$ for HypVC(k). The following formal statement was proved by Khot and Regev [14].

Theorem 6. *(Khot et al. [14]) Assuming the Unique Games Conjecture of Khot [12], For any $\varepsilon > 0$, it is NP-hard to approximate* HypVC(k) *to within a factor of $k - \varepsilon$.*

Remark 1. A recent paper by Bansal and Khot [3] shows a strong hardness result assuming the UGC for distinguishing between a k-uniform hypergraph that is *almost* k-partite and one which has no vertex cover containing at most a $(1-\varepsilon)$ fraction of vertices (for any desired $\varepsilon > 0$). We note that this is very different from our problem where the input is always k-partite with a given k-partition (and in particular has an easily found vertex cover with a $1/k$ fraction of vertices, namely the smallest of the k parts).

3.2 Our Results

NP-hardness results We prove the following theorem on the NP-hardness of approximating the minimum vertex cover on strongly colorable hypergraphs.

Theorem 7. *For every $\varepsilon > 0$ and integer $m \geqslant k \geqslant 3$ (such that $m \geqslant 2k$), it is NP hard to approximate* StrongColored-HypVC(m, k) *to within a factor of*

$$\frac{(m-(k-1))(k-1)}{m} - \varepsilon.$$

The above theorem is proved in Section 4 via a reduction from HypVC(k) to StrongColored-HypVC(m, k). A simple reduction from StrongColored-HypVC(k, k') also shows the following hardness results for HypVCpartite(k) and HypVCsplit$(r, k, p_1, \ldots, p_k)$. We prove Theorem 8 in Section 5 while we omit the proof of Theorem 9 due to lack of space.

Theorem 8. *For every $\varepsilon > 0$ and integer $k > 16$, it is NP-hard to approximate* HypVCpartite(k) *within a factor of $\frac{k}{4} - \varepsilon$.*

Theorem 9. *For every $\varepsilon > 0$, and positive integers $r, k, p_1, \ldots, p_k$ such that $\sum_{i=1}^{k} p_i = r \geqslant 3$ and $t := \max\{p_1, \ldots, p_k\} \geqslant 3$, it is NP-hard to approximate* HypVCsplit$(r, k, p_1, \ldots, p_k)$ *to within a factor of*

$$\max\left\{\frac{r}{4}, t-1\right\} - \varepsilon.$$

The above hardness of approximation results do not quite match the algorithmic results in Theorem 4. The next few paragraphs illustrate how recent results of [15] can be combined with the integrality gaps given in Theorems 1 and 4 to yield tight inapproximability for the corresponding problems.

Unique Games hardness In recent work Kumar, Manokaran, Tulsiani and Vishnoi [15] have shown that for a large class of monotone constraint problems, including hypergraph vertex cover, integrality gaps for a natural LP relaxation can be transformed into corresponding hardness of approximation results based on the Unique Games Conjecture.

The reduction in [15] is analyzed using the general bounds on noise correlation of functions proved by Mossel [18]. For this purpose, the reduction perturbs a "good" solution, say x^*, to the LP relaxation for the integrality gap $G_{\mathcal{I}} = (V_{\mathcal{I}}, E_{\mathcal{I}})$, so that x^* satisfies the property that all variables are integer multiples of some $\varepsilon > 0$. Therefore, the number of distinct values in x^* is $m \approx 1/\varepsilon$. The reduction is based on a "dictatorship test" over the set $[m] \times \{0, 1\}^r$ (for some parameter r) and the hardness of approximation obtained is related to the performance of a certain (efficient) rounding algorithm on x^*, which returns a solution no smaller than the optimum on $G_{\mathcal{I}}$. As described in [15] the reduction is not guaranteed to preserve structural properties of the integrality gap instance $G_{\mathcal{I}}$, such as strong colorability or k-partiteness.

We make the simple observation that the dictatorship test in the above reduction can analogously be defined over $V_{\mathcal{I}} \times \{0,1\}^r$ which then preserves strong colorability and partiteness properties of $G_{\mathcal{I}}$ into the final instance. The gap obtained depends directly on the optimum in $G_{\mathcal{I}}$. This observation, combined with the result of [15] and the integrality gap for HypVCpartite(k) stated in Theorem 1 yields the following optimal UGC based hardness result.

Theorem 10. *Assuming the Unique Games Conjecture, it is NP-hard to approximate* HypVCpartite(k) *to within a factor of* $\frac{k}{2} - \varepsilon$ *for any* $\varepsilon > 0$.

Due to lack of space we omit the proof and refer the reader to the full version of this paper and [15] for details of the analysis. Similar inapproximability results can be obtained for StrongColored-HypVC(m,k) and HypVCsplit($r,k,p_1,\ldots,p_k$) using the corresponding integrality gaps given in Theorems 3 and 4.

4 Reduction from HypVC(k) to StrongColored-HypVC(m,k) and Proof of Theorem 7

Let k and m be two positive integers such that $m \geqslant k \geqslant 2$. In this section we give a reduction from an instance of HypVC(k) to an instance of StrongColored-HypVC(m,k).

Reduction. Let the $H = (U,F)$ be an instance of HypVC(k), i.e. H is a k-uniform hypergraph with vertex set U, and a set F of hyperedges. The reduction constructs an instance $G = (V,E)$ of StrongColored-HypVC(m,k) where G is an k-uniform, m-strongly colorable hypergraph, i.e. $V = \cup_{i=1}^m V_i$, where V_i are m disjoint subsets (color classes) such that every hyperedge in E has exactly one vertex from each subset. The main idea of the reduction is to let new vertex set V be the union of m copies of U, and for every hyperedge $e' \in F$, add all hyperedges which contain exactly one copy (in V) of every vertex in e', and at most one vertex from any of the m copies of U (in V). Clearly every hyperedge 'hits' any of the m copies of U in V at most once which naturally gives an m-strong coloring of V. It also ensures that if there is a vertex cover in G which is the union of a subset of the copies of U, then it must contain at least $m-k+1$ of the copies. Our analysis shall essentially build upon this idea.

To formalize the reduction we first need to define a useful notation.

Definition 5. *Given a hyperedge* $e' = \{u_1,\ldots,u_k\}$ *in* F*, and a subset* $I \subseteq [m]$ *where* $|I| = k$*, a mapping* $\sigma : I \mapsto \{u_1,\ldots,u_k\}$ *is said to be a "*(I,e')*-matching" if* σ *is a one-to-one map. Let* $\Gamma_{I,e'}$ *be the set of all* (I,e')*-matchings. Clearly,* $|\Gamma_{I,e'}| = k!$ *, for all* $I \subseteq [m], |I| = k$ *and* $e' \in F$.

The steps of the reduction are as follows.

1. For $i = 1,\ldots,m$, let $V_i = U \times \{i\}$

2. For every hyperedge e' in F, for every subset $I \subseteq [m]$ such that $|I| = k$, for every (I, e')-matching $\sigma \in \Gamma_{I,e'}$ we add the hyperedge $e = e(e', I, \sigma)$ which is defined as follows.

$$\forall i \in [m], \qquad V_i \cap e = \begin{cases} (\sigma(i), i) & \text{if } i \in I \\ \emptyset & \text{otherwise.} \end{cases} \tag{5}$$

The above reduction outputs the instance $G = (V, E)$ of StrongColored-HypVC(m, k). Note that the vertex set V is of size $m|U|$ and for every hyperedge $e' \in F$ the number of hyperedges added in E is $\binom{m}{k} \cdot k!$. Therefore the reduction is polynomial time. In the next section we present the analysis of this reduction.

Analyzing the reduction

Theorem 11. *Let C be the size of the optimal vertex cover in $H = (U, F)$, and let C' be the size of the optimal vertex cover in $G = (V, E)$. Then,*

$$(m - (k-1))C \leqslant C' \leqslant mC$$

Using the above theorem we can complete the proof of Theorem 7 as follows.

Proof. (of Theorem 7) Theorem 11 combined with the $k-1-\varepsilon$ inapproximability for HypVC(k) given by [5] and stated in Theorem 5, implies an inapproximability of,

$$\frac{(m - (k-1))C_1}{mC_2},$$

for some integers C_1, C_2 (depending on H) such that $C_1 \geqslant C_2(k - 1 - \varepsilon)$ for some $\varepsilon > 0$. It is easy to see that the above expression can be simplified to yield

$$\frac{(m - (k-1))(k-1)}{m} - \varepsilon' \tag{6}$$

as the inapproximability factor for StrongColored-HypVC(m, k). This proves Theorem 7. □

Proof. (of Theorem 11) We first show that there is a vertex cover of size at most mC in G, where C is the size of an optimal vertex cover U^* in H. To see this consider the set $V^* \subseteq V$, where $V^* = U^* \times [m]$. For every hyperedge $e' \in F$, $e' \cap U^* \neq \emptyset$, and therefore $e \cap U^* \times \{i\} \neq \emptyset$, for some $i \in [m]$, for all $e = e(e', I, \sigma)$. Therefore, $V^* \cap e \neq \emptyset$ for all $e \in E$. The size of V^* is mC which proves the upper bound in Theorem 11. In the rest of the proof we shall prove the lower bound in Theorem 11.

Let S be the optimal vertex cover in G. Our analysis shall prove a lower bound on the size S in terms of the size of the optimal vertex cover in H. Let $S_i := V_i \cap S$ for $i \in [m]$. Before proceeding we introduce the following useful quantity. For every $Y \subseteq [m]$, we let $A_Y \subseteq U$ be the set of all vertices which have a copy in S_i for some $i \in Y$. Formally,

$$A_Y := \{u \in U \mid \exists i \in Y \text{ s.t. } (u, i) \in S_i\}.$$

The following simple lemma follows from the construction of the edges E in G.

Lemma 1. *Let $I \subseteq [m]$ be any subset such that $|I| = k$. Then A_I is a vertex cover of the hypergraph H.*

Proof. Fix any subset I as in the statement of the lemma. Let $e' \in F$ be any hyperedge in H. For a contradiction assume that $A_I \cap e' = \emptyset$. This implies that the sets S_i $(i \in I)$ do not have a copy of any vertex in e'. Now choose any $\sigma \in \Gamma_{I,e'}$ and consider the edge $e(e', I, \sigma) \in E$. This edge can be covered only by vertices in V_i for $i \in I$. However, since S_i does not contain a copy of any vertex in e' for $i \in I$ the edge $e(e', I, \sigma)$ is not covered by S which is a contradiction. This completes the proof. □

The next lemma combines the previous lemma with the minimality of S to show a strong structural statement for S, that any S_i is "contained" in the union of any other k sets S_j. It shall enable us to prove that most of the sets S_i are large.

Lemma 2. *Let $I \subseteq [m]$ be any set of indices such that $|I| = k$. Then, for any $j' \in [m]$, $S_{j'} \subseteq A_I \times \{j'\}$.*

Proof. Let I be any choice of a set of k indices in $[m]$ as in the statement of the lemma. From Lemma 1 we know that A_I is a vertex cover in H and is therefore non-empty. Let $j' \in [m]$ be an arbitrary index for which we shall verify the lemma for the above choice of I. If $j' \in I$, then the lemma is trivially true. Therefore, we may assume that $j' \notin I$. For a contradiction we assume that,

$$(u, j') \in S_{j'} \setminus (A_I \times \{j'\}) \tag{7}$$

From the minimality of S, we deduce that there must be a hyperedge, say $e \in E$ such that e is covered by (u, j') and by no other vertex in S; otherwise $S \setminus \{(u, j')\}$ would be a smaller vertex cover in G. Let $e = e(e', I', \sigma)$ for some $e' \in F$, $I' \subseteq [m]$ $(|I'| = k)$ and $\sigma \in \Gamma_{I',e'}$. Now, since (u, j') covers e, we obtain that $j' \in I'$ and $\sigma(j') = u \in e'$. Combining this with the fact that $j' \notin I$, and that $|I| = |I'| = k$, we obtain that $I \setminus I' \neq \emptyset$.

Let $j \in I \setminus I'$. We claim that $(u, j) \notin S_j$. To see this, observe that if $(u, j) \in S_j$ then $u \in A_I$ which would contradict our assumption in Equation (7).

We now consider the following hyperedge $\tilde{e} = \tilde{e}(e', \tilde{I}, \tilde{\sigma}) \in E$ where the quantities are defined as follows. The set $\tilde{I}$ simply replaces the index j' in I' with the index j, i.e.

$$\tilde{I} = (I' \setminus \{j'\}) \cup \{j\}. \tag{8}$$

Analogously, $\tilde{\sigma} \in \Gamma_{\tilde{I},e'}$ is identical to σ except that it is defined on j instead of j' where $\tilde{\sigma}(j) = \sigma(j') = u$. Formally,

$$\tilde{\sigma}(i) = \begin{cases} \sigma(i) & \text{if } i \in \tilde{I} \setminus \{j\} \\ u & \text{if } i = j \end{cases} \tag{9}$$

Equations (8) and (9) imply the following,

$$V_i \cap \tilde{e} = V_i \cap e \ \forall i \in [m] \setminus \{j, j'\} \tag{10}$$

$$V_j \cap \tilde{e} = (u, j) \tag{11}$$

$$V_{j'} \cap \tilde{e} = \emptyset \tag{12}$$

Since $(u, j') \in S$ uniquely covers e, Equation (10) implies that $\tilde{e}$ is not covered by any vertex in S_i for all $i \in [m] \setminus \{j, j'\}$. Moreover, since $j' \notin \tilde{I}$ no vertex in $S_{j'}$ covers $\tilde{e}$. On the other hand, by our assumption in Equation (7) $(u, j) \notin S_j$, which along with Equation (11) implies that no vertex in S_j covers $\tilde{e}$. Therefore, $\tilde{e}$ is not covered by S. This is a contradiction to the fact that S is a vertex cover in G and therefore our assumption in Equation (7) is incorrect. This implies that $S_{j'} \subseteq A_I \times \{j'\}$. This holds for every j', thus proving the lemma. □

Note that the above lemma immediately implies the following corollary.

Corollary 1. *For every $I \subseteq [m], |I| = k$, we have $A_{[m]} = A_I$.*

It is easy to see the following simple lemma.

Lemma 3. *For any vertex $u \in A_{[m]}$, let $I_u \subseteq [m]$ be the largest set such that $u \notin A_{I_u}$. Then, $|I_u| < k$.*

Proof. Suppose the above does not hold. Then I_u (or any subset of I_u of size k) would violate Corollary 1, which is a contradiction. This completes the proof. □

The above lemma immediately implies the desired lower bound on the size of S.

Lemma 4. *Let C be the size of the optimal vertex cover in H. Then,*

$$|S| \geqslant (m - (k-1))C.$$

Proof. For convenience, let $q = |A_{[m]}|$. Note that, by Lemma 1 $A_{[m]}$ is a vertex cover in H. Therefore, $q \geqslant C$. From Lemma 3 we deduce that every vertex $u \in A_{[m]}$ has a copy (u, i) in at least $m - (k-1)$ of the sets S_i. Therefore, S contains $m - (k-1)$ copies of every vertex in $A_{[m]}$ which yields,

$$|S| \geqslant (m - (k-1))q \geqslant (m - (k-1))C,$$

thus completing the proof. □

The above also completes the proof of the lower bound of Theorem 11. □

5 Reduction from StrongColored-HypVC(k, k') to HypVCpartite(k) and Proof of Theorem 8

We prove Theorem 8 by giving a simple reduction from an instance $G = (V, E)$ of StrongColored-HypVC(k, k') to an instance $G' = (V', E')$ of HypVC-partite(k) where the parameters will be chosen later.

Given $G = (V, E)$ construct V' by adding k "dummy" vertices $b_1, \ldots, b_k$ to V, i.e. $V' = V \cup \{b_1, \ldots, b_k\}$. Let $V_1, \ldots, V_k$ be the k color classes of V. For any hyperedge $e \in E$, construct a corresponding hyperedge $e' \in E'$ which contains all the vertices in e in addition to b_i if $e \cap V_i = \emptyset$ for all $i \in [k]$. It is easy to see that G' is a k-partite hypergraph with the k-partition given by the subsets $V_i \cup \{b_i\}$. As a final step, set the weight of the dummy vertices $b_1, \ldots, b_k$ to be much larger than $|V'|$ so that no dummy vertex is chosen in any optimal vertex cover in G'. This is because V is always a vertex cover in G. Note that the hypergraph can be made unweighted by the (standard) technique of replicating each dummy vertex many times and multiplying the hyperedges appropriately.

Since no optimal vertex cover in G' contains a dummy vertex we deduce that an optimal vertex cover in G' is an optimal vertex cover in G and vice versa. From Theorem 7, for any $\varepsilon > 0$, we obtain a hardness factor of,

$$\frac{(k - (k' - 1))(k' - 1)}{k} - \varepsilon,$$

for approximating HypVCpartite(k). Let $\alpha := \frac{(k'-1)}{k}$. The above expression is maximized in terms of k when $\frac{(1-\alpha)\alpha}{2}$ attains a maximum where $\alpha \in [0, 1]$. Clearly, the maximum is obtained when $\alpha = \frac{(k'-1)}{k} = \frac{1}{2}$, thus yielding as the hardness of approximation factor:

$$\left(\frac{k' - 1}{2}\right) - \varepsilon = \left(\frac{k}{4}\right) - \varepsilon,$$

which proves Theorem 8.

Acknowledgment The authors would like to thank Anupam Gupta for suggestions which helped simplify the analysis.

References

1. R. Aharoni, R. Holzman, and M. Krivelevich. On a theorem of Lovász on covers in r-partite hypergraphs. *Combinatorica*, 16(2):149–174, 1996.
2. P. Austrin and E. Mossel. Approximation resistant predicates from pairwise independence. In *IEEE Conference on Computational Complexity*, pages 249–258, 2008.
3. N. Bansal and S. Khot. Inapproximability of hypergraph vertex cover and applications to scheduling problems. In *Proceedings of ICALP*, July 2009.
4. I. Dinur, V. Guruswami, and S. Khot. Vertex cover on k-uniform hypergraphs is hard to approximate within factor (k-3-ε). ECCC Technical Report TR02-027, 2002.
5. I. Dinur, V. Guruswami, S. Khot, and O. Regev. A new multilayered PCP and the hardness of hypergraph vertex cover. In *Proc.* 35^{th} *ACM STOC*, pages 595–601, 2003.
6. I. Dinur and S. Safra. The importance of being biased. In *Proc.* 34^{th} *ACM STOC*, pages 33–42, 2002.
7. T. Feder, R. Motwani, L. O'Callaghan, R. Panigrahy, and D. Thomas. Online distributed predicate evaluation. Technical Report, Stanford University, 2003.
8. G. Gottlob and P. Senellart. Schema mapping discovery from data instances. *J. ACM*, 57(2), Jan. 2010.
9. E. Halperin. Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs. *SIAM J. Comput.*, 31(5):1608–1623, 2002.

10. J. Holmerin. Improved inapproximability results for vertex cover on k -uniform hypergraphs. In *Proc. ICALP*, pages 1005–1016, 2002.
11. L. Ilie, R. Solis-Oba, and S. Yu. Reducing the size of NFAs by using equivalences and preorders. In *Proc. CPM*, pages 310–321, 2005.
12. S. Khot. On the power of unique 2-prover 1-round games. In *Proc.* 34^{th} *ACM STOC*, pages 767–775, 2002.
13. S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007.
14. S. Khot and O. Regev. Vertex cover might be hard to approximate to within 2-epsilon. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008.
15. A. Kumar, R. Manokaran, M. Tulsiani, and N. K. Vishnoi. On the optimality of a class of LP-based algorithms. Manuscript, 2009.
16. L. Lovász. On minimax theorems of combinatorics. *Doctoral Thesis, Mathematiki Lapok*, 26:209–264, 1975.
17. R. Manokaran, J. Naor, P. Raghavendra, and R. Schwartz. SDP gaps and UGC hardness for multiway cut, 0-extension, and metric labeling. In *Proc.* 40^{th} *ACM STOC*, pages 11–20, 2008.
18. E. Mossel. Gaussian bounds for noise correlation of functions and tight analysis of long codes. In *Proc.* 49^{th} *IEEE FOCS*, pages 156–165, 2008.
19. P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc.* 40^{th} *ACM STOC*, pages 245–254, 2008.
20. A. Samorodnitsky and L. Trevisan. Gowers uniformity, influence of variables, and PCPs. *SIAM J. Comput.*, 39(1):323–360, 2009.
21. L. Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proc.* 33^{rd} *ACM STOC*, pages 453–461, 2001.