Vertex Cover in Bipartite Graphs and k-partite Hypergraphs[1]

- In this note, we describe a randomized rounding algorithm which solves the minimum cost vertex cover problem *exactly* in bipartite graphs. This algorithm is then generalized to give a $\frac{3}{2}$-approximation for minimum cost vertex cover in a *tri-partite* 3-hypergraph. Indeed, the generalization works for k-partite hypergraphs as well, but we stick to 3 for exposition purposes and leave the generalization to 3 as an exercise.

- *Vertex covers in graphs and hypergraphs.* A hypergraph $H = (V, E)$ is a generalization of graphs where a (hyper)-edge $e \in E$ is an arbitrary subset of V (instead of being just a pair). A hypergraph is k-uniform if $|e| = k$ for every $e \in E$. Thus, a graph is a 2-hypergraph. A hypergraph is k-partite if the vertex set can be partitioned into k-parts $V = V_1 \cup V_2 \cup \cdots \cup V_k$ such that for every edge $e \in E$, $|e \cap V_i| \leq 1$ for all $1 \leq i \leq k$. In plain English, every edge has at most one vertex from each part. A k-partite k-uniform hypergraph must satisfy the above inequality with equality for all i. This generalizes bipartite graphs. A vertex cover $C \subseteq V$ is one which hits every edge; for all $e \in E$, $e \cap C \neq \emptyset$. The vertex cover problem in graph/hypergraph is to find the smallest cost vertex cover, when each vertex v is associated with a non-negative cost c_v.

- *LP-relaxation.* The algorithm is a rounding algorithm for the following standard LP.

$$\texttt{opt} \leq \mathsf{lp}(G) := \text{minimize} \quad \sum_{v \in V} c_v z_v \qquad \text{(VC-LP)}$$

$$\sum_{v \in e} z_v \geq 1, \qquad \forall e \in E \qquad (1)$$

$$0 \leq z_v \leq 1, \qquad \forall v \in V \qquad (2)$$

  Note that the above LP is oblivious to the fact that the graph/hypergraph is k-partite or k-uniform. The rounding algorithm will use the k-partition.

- *The Bipartite Graph case.* Before describing the hypergrap case, let's see the randomized algorithm for the graph case. Let $V = V_1 \cup V_2$.

```
1: procedure RANDOMIZED BIPARTITE VC(G = (V, E), c):
2:     Solve (VC-LP) to obtain z_v for every vertex.
3:     Sample r ∈ [0, 1] uniformly at random.
4:     For every vertex v ∈ V_1, add v to C if z_v ≥ r.
5:     For every vertex v ∈ V_2, add v to C if z_v ≥ 1 − r.
6:     return C.
```

[1] *Lecture notes by Deeparnab Chakrabarty. Last modified : 18th Mar, 2022*
These have not gone through scrutiny and may contain errors. If you find any, or have any other comments, please email me at deeparnab@dartmouth.edu. Highly appreciated!

Theorem 1. RANDOMIZED BIPARTITE VC returns a subset C which is a vertex cover with probability 1 and $\mathbf{Exp}[c(C)] = \mathsf{lp}$.

Proof. For any edge (u, w) with $u \in V_1$ and $w \in V_2$, we have $C \cap \{u, w\} = \emptyset$ only if $z_u < r$ and $z_w < 1 - r$. However, this contradicts $z_u + z_w \geq 1$. Therefore, C is a vertex cover with probability 1.

Furthermore, for any $u \in V_1$, the probability $u \in C$ is precisely the probability $r \in [0, z_u]$. Thus $\mathbf{Pr}[u \in C] = z_v$ since r is uniformly at random chosen in $[0, 1]$ and thus the probability is the ratio of the lengths. Similarly, for any $w \in V_2$, the probability $w \in C$ is precisely the probability $r \in [1 - z_w, 1]$. Thus $\mathbf{Pr}[w \in C] = z_w$. Thus, $\mathbf{Exp}[\mathtt{cost}(C)] = \sum_{v \in V_1 \cup V_2} c_v \mathbf{Pr}[v \in C] = \sum_{v \in V_1 \cup V_2} c_v z_v = \mathsf{lp}$. □

Remark: *Note that the cost of C can never be* less *than* lp *otherwise we would get a vertex cover of cost* $< \mathsf{lp} \leq \mathtt{opt}$. *This means that* ***any*** *solution returned, irrespective of r, must be of cost* lp, *and* lp *must equal* $\mathtt{opt}$. *Thus, randomness isn't necessary at all! And furthermore,* (VC-LP) *has integrality gap 1 when G is a bipartite graph. This can be proved in many ways, but the above is a really slick proof.*

- *Generalizing to hypergraphs.* Before we proceed to the hypergraph case, let's understand what happened above. We wanted to choose *two* (since the graph was *bi*-partite) coupled random variables (r_1, r_2) with the following plan : (a) for *every* vertex v, we put it in the cover only if and only if $z_v \geq r_i$, (b) the variable r_i is uniform in some interval $[0, \alpha]$ such that the probability of v being in the solution can then be analyzed to be $1/\alpha$, and (c) for every edge to be covered, we wanted $r_1 + r_2 \leq 1$ with probability 1; this would imply the random set we pick is a cover with probability 1 since the z's sum to at least 1. We obtained this by selecting $r \in [0, 1]$ and setting $r_1 = r$ and $r_2 = 1 - r$, both of which are uniform in $[0, 1]$.

  We follow the same scheme for hypergraphs. We limit our discussion to tri-partite 3-uniform hypergraphs and we give a $\frac{3}{2}$-approximation. However, the same ideas give a $k/2$-approximation for k-partite k-uniform hypergraphs. To make the scheme precise, here is a definition.

  Definition 1 (Nice Distribuition)**.** *A distribution (r_1, r_2, r_3) of is called a* nice *distribution $\mathcal{D}$ if (a) the marginal distribution of each r_i is* **uniform** *in* $\left[0, \frac{2}{3}\right]$, *and (b)* $r_1 + r_2 + r_3 = 1$ *with probability* 1.

  It is not immediately clear nice distributions exist. In the next bullet point we will show one explicit nice distribution. Before doing so, let us see that they imply a 3/2-approximation for the minimum cost vertex cover problem in tri-partite 3-hypergraphs.

procedure RANDOMIZED TRIPARTITE 3HYPVC(Tripartite 3-regular hypergraph $G = (V_1 \cup V_2 \cup V_3, E), c_v$):
    ▷ *Assume access to nice distribution $\mathcal{D}$.*
    Solve (VC-LP) to obtain z_v for every vertex.
    Sample $(r_1, r_2, r_3) \sim \mathcal{D}$.
    For every vertex $v \in V_i$, add v to C if $z_v \geq r_i$.
    return C.

Theorem 2. RANDOMIZED TRIPARTITE 3HYPVC returns a subset C which is a vertex cover with probability 1 and $\mathbf{Exp}[c(C)] \leq \frac{3}{2} \cdot \mathsf{lp}$.

Proof. For any edge $e = (v_1, v_2, v_3)$, $C \cap e = \emptyset$ implies $\sum_{i=1}^{3} z_{v_i} < \sum_{i=1}^{3} r_i$, but since the latter is $= 1$ we get a contradiction that z's formed a feasible solution to the LP. Thus, C is a vertex cover with probability 1. Furthemore, for any $1 \leq i \leq 3$ and any $v \in V_i$, we get that $\mathbf{Pr}[v \in C] = \mathbf{Pr}[r_i \in [0, z_v]]$. Since the marginal of r_i is uniform in $[0, \frac{2}{3}]$, this probability is at most $\frac{z_v}{2/3} = \frac{3z_v}{2}$. Note that if $z_v > 2/3$, then we get a strict inequality, otherwise, we get an equality. Therefore, Thus, $\mathbf{Exp}[c(C)] \leq \sum_{v \in V_1 \cup V_2 \cup V_3} c_v \mathbf{Pr}[v \in C] \leq \frac{3}{2} \cdot \sum_{v \in V_1 \cup V_2 \cup V_3} c_v z_v = \frac{3\mathsf{lp}}{2}$. □

- ***Nice distributions.*** We now describe one nice distribution. This, by any means, is not the only way to design one.
  - Sample $r \in [0, \frac{2}{3}]$ and set $r_1 \leftarrow r$.
  - If $r \leq \frac{1}{3}$, set $(r_2, r_3) \leftarrow \left(\frac{1}{3} + r, \frac{2}{3} - 2r\right)$.
  - Else, if $\frac{1}{3} < r \leq \frac{2}{3}$, set $(r_2, r_3) \leftarrow \left(r - \frac{1}{3}, \frac{4}{3} - 2r\right)$.
  - Return (r_1, r_2, r_3).

Lemma 1. *The above distribution (r_1, r_2, r_3) is a nice distribution.*

Proof. By design, $r_1 + r_2 + r_3 = 1$ with probability 1. Also, by design, r_1 is distributed uniformly in $[0, \frac{2}{3}]$. We need argue that r_2 and r_3 are uniformly distributed in $[0, \frac{2}{3}]$.

Fix an $x \in [0, 2/3]$. We claim that $\mathbf{Pr}[r_i \leq x] = \frac{x}{2/3}$ for $i = 2, 3$. Let's take $i = 2$ first. Case 1: $x \leq \frac{1}{3}$. Then $r_2 \leq x$ if and only if $r > 1/3$ (otherwise $r_2 > 1/3$) and $r_2 = r - \frac{1}{3} \leq x$. That is, $\frac{1}{3} < r \leq \frac{1}{3} + x$. And this probability is precisely $\frac{x}{2/3}$. Case 2: $x > \frac{1}{3}$. Then $r_2 \leq x$ if $r > 1/3$ (for then $r_2 = r - 1/3 < 1/3 < x$) and if $r \leq x - \frac{1}{3}$. This probability is $\frac{x - \frac{1}{3} + \frac{1}{3}}{2/3} = \frac{x}{2/3}$.

We leave the similar calculation about the cdf of r_3 as an exercise for the reader. □

Exercise: *Show a $\frac{k}{2}$-approximation for the minimum cost vertex cover problem in k-partite k-uniform hypergraphs. To do so, first generalize the notion of "nice distributions". Then use the fact that nice distributions exist for $k = 2$ and $k = 3$ to generalize for all k. Hint: for even k, just the $k = 2$ case is enough.*

Notes

The above result is an old theorem of Lovasz [3] from his doctoral thesis. This thesis, unfortunately, is hard to find and presumably in Hungarian. The presentation here follows from the paper [1] by Aharoni, Holzman, and Krivelevich. The latter paper also gives an integrality gap example; more precisely, for any $\varepsilon > 0$, they describe a hypergraph for which the smallest vertex cover is of size $\geq \left(\frac{k}{2} - \varepsilon\right) \mathsf{lp}$. Inspired by this example, the paper [2] by Guruswami, Sachdeva, and Saket showed it is UGC-hard to obtain an $(k/2 - \varepsilon)$-approximation, and indeed NP-hard to obtain an $(\frac{k}{2} - 1 + \frac{1}{2k} - \varepsilon)$-approximation, for any $\varepsilon > 0$.

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] V. Guruswami, S. Sachdeva, and R. Saket. Inapproximability of minimum vertex cover on k-uniform k-partite hypergraphs. *SIAM Journal on Discrete Mathematics (SIDMA)*, 29(1):36–58, 2015.

[3] L. Lovász. On minimax theorems of combinatorics. *Doctoral Thesis, Math. Lapok*, 26:209–264, 1975. In Hungarian.