A (1.999999)-APPROXIMATION RATIO FOR VERTEX COVER PROBLEM

A Preprint

Majid Zohrehbandian
Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran.
zohrebhandian@yahoo.com

May 4, 2024

ABSTRACT

The vertex cover problem is a famous combinatorial problem, and its complexity has been heavily studied over the years. While a 2-approximation for it can be trivially obtained, researchers have not been able to approximate it better than $2-o(1)$. In this paper, by a combination of a new semidefinite programming formulation along with satisfying new proposed properties, we introduce an approximation algorithm for the vertex cover problem with a performance ratio of 1.999999 on arbitrary graphs, en route to answering an open question about the correctness/incorrectness of the unique games conjecture.

Keywords: Combinatorial Optimization · Vertex Cover Problem · Unique Games Conjecture · Complexity Theory

1 Introduction

In complexity theory, the abbreviation $NP$ refers to “nondeterministic polynomial”, where a problem is in $NP$ if we can quickly (in polynomial time) test whether a solution is correct. $P$ and $NP$-complete problems are subsets of $NP$ Problems. We can solve $P$ problems in polynomial time while determining whether or not it is possible to solve $NP$-complete problems quickly (called the $P$ vs $NP$ problem) is one of the principal unsolved problems in Mathematics and Computer science.

Here, we consider the vertex cover problem (VCP) which is a famous $NP$-complete problem. It cannot be approximated within a factor of 1.36 [1], unless $P=NP$, while a 2-approximation factor for it can be trivially obtained by taking all the vertices of a maximal matching in the graph. However, improving this simple 2-approximation algorithm was a hard task [2, 3].

In this paper, we show that there is a $(2-\varepsilon)$-approximation ratio for the vertex cover problem, where the value of $\varepsilon$ is not constant. Then, we fix the $\varepsilon$ value equal to $\varepsilon=0.000001$ and we show that on arbitrary graphs, a 1.999999-approximation ratio can be obtained by solving a new semidefinite programming (SDP) formulation.

The rest of the paper is structured as follows. Section 2 is about the vertex cover problem and introduces new properties about it. In section 3, by using the satisfying properties, we propose a solution algorithm for VCP with a performance ratio of 1.999999 on arbitrary graphs. Finally, Section 4 concludes the paper.

2 Performance ratio based on a VCP feasible solution

In the mathematical discipline of graph theory, a vertex cover of a graph is a set of vertices such that each edge of the graph is incident to at least one vertex of the set. The problem of finding a minimum vertex cover is a typical example of an $NP$-complete optimization problem. In this section, we calculate the performance ratios of VCP feasible solutions which lead to an approximation ratio of $2-\varepsilon$, where the value of $\varepsilon$ is not constant and depends on the produced feasible
A (1.999999)-approximation ratio for vertex cover problem

M. ZOHREHBANDIAN

solution. Then, in the next section, we will fix the value of \( \varepsilon \) equal to \( \varepsilon = 0.000001 \) to produce a 1.999999-approximation ratio for the vertex cover problem.

Let \( G = (V, E) \) be an undirected graph on vertex set \( V \) and edge set \( E \), where \( |V| = n \). Throughout this paper, \( z_{VCP}^* \) is the optimal value for the vertex cover problem and we have produced a feasible solution for the problem with vertex partitioning \( V = V_1 \cup V_0 \) and objective value \( |V_1| \). The integer linear programming (ILP) model for VCP is as follows; i.e. \( z^*_V = z_{VCP}^* \).

\[
(1) \min_{s.t.} z = \sum_{i \in V} x_i
\]

\[
x_i + x_j \geq 1 \quad ij \in E
\]

\[
x_i \in \{0, +1\} \quad i \in V
\]

**Lemma 1.** [4] Let \( x^* \) be an extreme optimal solution to the linear programming (LP) relaxation of the model (1). Then \( x^*_j \in \{0, 0.5, 1\} \quad j \in V \) and if we define \( V^0 = \{ j \in V \mid x^*_j = 0 \} \), \( V^{0.5} = \{ j \in V \mid x^*_j = 0.5 \} \) and \( V^1 = \{ j \in V \mid x^*_j = 1 \} \); then, there exist a VCP optimal solution which includes all of the vertices \( V^1 \) and it is a subset of \( V^{0.5} \cup V^1 \).

**Theorem 1.** Let \( x^* \) be an extreme optimal solution to the LP relaxation of the model (1) and \( V^0 = \{ j \in V \mid x^*_j = 0 \} \), \( V^{0.5} = \{ j \in V \mid x^*_j = 0.5 \} \), \( V^1 = \{ j \in V \mid x^*_j = 1 \} \) and \( G_{0.5} \) be the induced graph on the vertices \( V^{0.5} \). If we can introduce a vertex cover feasible partitioning \( V^{0.5} = V^{0.5}_1 \cup V^{0.5}_0 \) with an approximation ratio of \( 1 \leq \rho < 2 \), for the VCP on \( G_{0.5} \), then, the vertex cover feasible partitioning \( V = (V_1 \cup V_0) = (V^{0.5}_1 \cup V^1) \cup (V^{0.5}_0 \cup V^0) \), has an approximation ratio of \( 1 \leq \rho < 2 \), for the VCP on \( G \).

**Proof.** We have \( \frac{\|V^{0.5}_1\| + |V^1|}{z_{VCP}(G_{0.5})} \leq \rho \) Therefore,

\[
|V^{0.5}_1| + (1 - \rho) |V^1| \leq \rho \cdot z_{VCP}(G_{0.5})
\]

We know that we can efficiently solve the following SDP formulation as a relaxation of the VCP model (1).

\[
(2) \min_{s.t.} z = \sum_{i \in V} X_{oi}
\]

\[
X_{oi} + X_{oj} \geq 1 \quad ij \in E
\]

\[
0 \leq X_{oi} \leq +1 \quad i \in V
\]

\[
X \geq 0
\]

This model can be written as follows:

\[
(3) \min_{s.t.} z = \sum_{i \in V} X_{oi}
\]

\[
X_{oi} + X_{oj} - X_{ij} = 1 \quad ij \in E
\]

\[
X_{ii} = 1, \quad 0 \leq X_{ij} \leq +1 \quad i, j \in V \cup \{o\}
\]

\[
X \geq 0
\]

Moreover, by introducing the vector set \( v_o, v_1, ..., v_n \) for which \( V_1 = \{ i \in V \mid v_i = v_o \} \) is a feasible vertex cover, and \( V_o = V - V_1 \) is the set of the perpendicular vectors to \( v_o \) and \( v_i, v_j \) = \( X_{ij} \), see Figure 1, SDP (3) can be written as follows:

\[
(4) \min_{s.t.} z = \sum_{i \in V} v_o \cdot v_i
\]

\[
v_o \cdot v_i + v_o \cdot v_j - v_i \cdot v_j = 1 \quad ij \in E
\]

\[
v_i \cdot v_j = 1, \quad 0 \leq v_i \cdot v_j \leq +1 \quad i, j \in V \cup \{o\}
\]

**Theorem 2.** Although it is hard to produce the exact VCP optimal value, let’s assume that we have a lower bound on the VCP optimal value and we know \( z_{VCP}^* \geq \frac{n}{2} + \frac{n}{2} = \frac{(k+2)n}{2k} \). Then, for all vertex cover feasible partitioning \( V = V_1 \cup V_0 \), we have the approximation ratio \( \frac{|V_1|}{z_{VCP}^*} \leq \frac{2k}{k+2} < 2 \).
1.999999-approximation ratio on arbitrary graphs. To do this, we introduce the following property on a solution value $v$ with $\varepsilon$.

**Theorem 4.** If $z_{VCP}^* \geq \frac{(k+2)n}{2k}$ and we have produced a VCP feasible partitioning $V = V_1 \cup V_0$, where $|V_1| \leq \frac{kn}{k+1}$ and $|V_0| \geq \frac{n}{k+1}$ (or $|V_1| \leq k |V_0|$), then, based on such a solution we have an approximation ratio $\frac{|V_1|}{z_{VCP}^*} \leq \frac{2k}{k+1} < 2$.

**Proof.** If $|V_1| \leq \frac{kn}{k+1}$ then $n \geq \frac{k+1}{k} |V_1|$. Hence, $z_{VCP}^* \geq \frac{n}{2k} \geq \frac{k+1}{2k} |V_1|$ which concludes that $\frac{|V_1|}{z_{VCP}^*} \leq \frac{2k}{k+1} < 2$.

3 A (1.999999)-approximation algorithm on arbitrary graphs

In section 2, based on a feasible solution for the vertex cover problem, we introduced a $(2-\varepsilon)$-approximation algorithm on arbitrary graphs $\varepsilon$ value was not a constant value. In this section, we fix the value of $\varepsilon$ equal to $\varepsilon=0.000001$ to produce a 1.999999-approximation ratio on arbitrary graphs. To do this, we introduce the following property on a solution value of the SDP (4) formulation.

**Property 1.** For some vertex cover problems, after solving the SDP (4), both of the following conditions occur:

(a) For less than 0.000001n of vertices $j \in V$ and corresponding vectors we have $v_j^* v_j^* < 0.5$.

(b) For less than 0.01n of vertices $j \in V$ and corresponding vectors we have $v_j^* v_j^* > 0.5004$.

**Theorem 4.** If $z_{VCP}^* \geq \frac{n}{2k}$ and the solution of the SDP (4) does not meet the Property (1) then we can produce a VCP solution with a performance ratio of 1.999999.

**Proof.** If the solution of the SDP (4) does not meet the Property (1.a) then we can introduce $V_0 = \{ j \in V \mid v_j^* v_j^* < 0.5 \}$ and $V_1 = V - V_0$, to have a VCP feasible solution with $|V_0| \geq 0.000001n$ and $|V_1| \leq 0.999999n \leq 0.999999 |V_0|$. Then, for such a solution and based on Theorem (3), we have an approximation ratio $\frac{|V_1|}{z_{VCP}^*} < \frac{0.999999}{1.999999+0.000001} = 1.999999 < 1.999999$.

Otherwise, if the solution of the SDP (4) meets the Property (1.a) but it does not meet the Property (1.b) then we have

$$z_{VCP}^* \geq z_{SDP(4)}^* \geq (0)(0.000001n) + (0.5)(0.989999n) + (0)(0.01n) = \frac{n}{2} + 0.0000035n.$$  

Note that, due to the correctness of Property (1.a) we have less than 0.000001n of vertices $j \in V$ with $v_j^* v_j^* < 0.5$ and due to the incorrectness of Property (1.b) we have more than 0.01n of vertices $j \in V$ with $v_j^* v_j^* > 0.5004$. Therefore, in the above inequality, the first summation is the lower bound on the vertices $j \in V$ with $v_j^* v_j^* < 0.5$, and the third summation is the lower bound on only 0.01n of the vertices $j \in V$ with $v_j^* v_j^* > 0.5004$ (only 0.01n of the vertices with $v_j^* v_j^* > 0.5004$ are considered in third summation and beyond the 0.01n of such vertices are considered in second summation). Moreover, the second summation is the lower bound on the other vertices; i.e. the vertices $j \in V$ with $0.5 \leq v_j^* v_j^* \leq 0.5004$ or the vertices $j \in V$ with $v_j^* v_j^* > 0.5004$ and beyond the 0.01n of such vertices which have been considered in third summation.

Therefore, based on the above lower bound on $z_{VCP}^*$ value and based on Theorem (2), for all VCP feasible solutions $V = V_1 \cup V_0$, we have the approximation ratio $\frac{|V_1|}{z_{VCP}^*} \leq \frac{2(1.999999)}{0.0000035+2} < 1.999999$. 

Figure 1. A VCP feasible solution
Theorem 5. For 4 normalized vectors $v_1, v_2, v_3, v_4$ which are perpendicular to each other, there exists exactly one normalized vector $v$ where $v.v_i = 0.5$ for $i = 1, 2, 3, 4$. Such a vector $v$ satisfies the equation $v = 0.5(v_1 + v_2 + v_3 + v_4)$.

Proof.
The following LP model has a unique optimal solution that corresponds to an extreme optimal solution for the LP relaxation of the model (1).


Moreover, we have a normalized vector \( v \) where \( v.v_i \approx 0.5 \) for all \( i = 1, 2, 3, 4 \). satisfies the equation \( v.v_i \approx 0.5(v_1 + v_2 + v_3 + v_4) \).

Corollary 1. For 4 normalized vectors \( v_1, v_2, v_3, v_4 \) which are almost perpendicular to each other, a normalized vector \( v \) where \( v.v_i \approx 0.5 \) for all \( i = 1, 2, 3, 4 \), satisfies the equation \( v.v_i \approx 0.5(v_1 + v_2 + v_3 + v_4) \).

Theorem 6. By solving SDP (6) on \( G_2 \), it is not possible to have an optimal solution that meets the Property (1) on both graphs \( G' \) and \( G'' \) unless the induced graph on \( V'_2 \) is bipartite and the induced graph on \( V''_2 \) is bipartite.

Proof. Suppose that we have an optimal solution that meets the Property (1) on both graphs \( G' \) and \( G'' \). Therefore, for an edge \( ab \) in \( E'_2 \) and an edge \( cd \) in \( E''_2 \) (a complete subgraph of \( G_2 \) on four vertices \( a, b, c, d \)) we have four normalized vectors \( v_a, v_b, v_c, v_d \) which are almost perpendicular to each other.

Moreover, we have a normalized vector \( v_0 \) for which \( v_0.v_h \approx 0.5 \) for all \( h = a, b, c, d \). Hence, based on Corollary (1) we have \( v_0.v_i \approx 0.5(v_0 + v_i + v_i + v_i) \). This means that for each edge \( ij \) in \( E'_2 \) we have \( v_0.v_i \approx 0.5(v_0 + v_j + v_i + v_i) \), and for each edge \( ij \) in \( E''_2 \) we have \( v_0.v_i \approx 0.5(v_0 + v_i + v_i + v_i) \).

Therefore, for each edge \( ij \) in \( E'_2 \) we have \( v_i + v_j \approx 2v_0 - v_i - v_j = U \), and for each edge \( ij \) in \( E''_2 \) we have \( v_i + v_j \approx 2v_0 - v_i - v_j = U \), where, due to almost perpendicular property of the vectors \( v_i \) and \( v_j \), we have \( |U| \approx W \approx \sqrt{v_i^2 + v_j^2} = \sqrt{2} \).

Now, suppose that we have an odd cycle on \( t \) vertices \( 1, 2, ..., t \), in \( G'_c = (V'_c, E'_c) \), where \( t = 2k + 1 \) is an odd number. Then, by addition of the vectors in this cycle, we have \( S = (v_1 + v_2) + (v_2 + v_3) + ... + (v_{t-2} + v_{t-1} + v_1) \approx tU \).

But, the above summation can do as \( S = 2(v_1 + v_2 + v_3 + ... + v_{t-2} + v_{t-1} + v_1) \) to produce the following results, which all of them must be \( \approx tU \).

\[
S = 2(v_1 + v_2 + v_3 + v_4 + ... + v_{t-2} + v_{t-1} + v_1) \approx 2(kU + v_1) = (t-1)U + 2v_1
\]
\[
S = 2(v_2 + v_3 + v_4 + v_5 + ... + v_{t-1} + v_1) \approx (t-1)U + 2v_1
\]
\[
S = 2(v_3 + v_4 + v_5 + v_6 + ... + v_1 + v_2) \approx (t-1)U + 2v_2
\]

... 

\[
v_1 \approx v_2 \approx ... \approx v_{t-1} \approx v_t \approx 0.5U
\]

Hence \( |U| \approx 2 |v_1| \approx 2 \neq \sqrt{2} \) and this is a contradiction; e.g. \( v_1 v_2 \approx 0.5U \), \( 0.5U \neq 0 \). Therefore, there is not any odd cycle in \( G'_c \), and similarly, there is not any odd cycle in \( G''_c \). Therefore, if the optimal solution of SDP (6) on \( G_2 \) meets the Property (1) on both graphs \( G' \) and \( G'' \), then both of the subgraphs \( G'_c \) and \( G''_c \) are bipartite.

Corollary 2. To produce a performance ratio of 1.999999 for problems with \( z_{VCP} \geq \frac{n}{2} \), we should solve SDP (6) on \( G_2 \). Then, if the solution of SDP (6) does not meet the Property (1), we have a performance ratio of 1.999999. Otherwise, the VCP problem on the bipartite graph \( G'_c \) is simple, and because \( |V_c| \geq 0.989999n \), solving such a simple problem produces a performance ratio of 1.999999.

Moreover, based on Theorem (1) and Corollary (2), to produce a performance ratio of 1.999999 for problems with \( z_{VCP} < \frac{n}{2} \), it is sufficient to produce an extreme optimal solution for the LP relaxation of the model (1).

Theorem 7. The following LP model has a unique optimal solution that corresponds to an extreme optimal solution for the LP relaxation of the model (1).

\[
(7) \min_{x \in \mathbf{R}_+} z = \sum_{i=1}^{n} (0.1)^{i} x_i
\]
\[ x_i + x_j \geq 1 \quad ij \in E \]
\[
\sum_{i \in V} x_i = z^* \]
\[ 0 \leq x_i \leq +1 \quad i \in V \]

**Proof.** The feasible region of the model (7) is an optimal face of the feasible region of the LP relaxation of the model (1). Therefore, its extreme optimal points correspond to the extreme optimal points of the LP relaxation of the model (1). Due to the properties of these extreme points, introduced in Lemma (1), and the objective coefficients of model (7), it is not possible to have more than one optimal extreme point. In other words, based on the priority weights on the decision variables of the model (7), its optimal solution corresponds to the unique extreme point solution of the following algorithm.

**Step 0.** Let \(k=1\) and \(z^*\) be the optimal value of the LP relaxation of the model (1).

**Step k.** Solve the following LP model.

\[
\begin{align*}
(8) \quad & \min_{x,t} \ z(k) = x_k \\
& x_i + x_j \geq 1 \quad ij \in E \\
& \sum_{i \in V} x_i = z^* \\
& x_i = x_i^* = z(k)^* \quad i = 1, \ldots, k-1 \\
& 0 \leq x_i \leq +1 \quad i \in V
\end{align*}
\]

Let \(k=k+1\). If \(k < n\) repeat this step, otherwise, the solution \(x^*\) is an extreme optimal solution of the LP relaxation of the model (1).

Therefore, our algorithm to produce an approximation ratio 1.999999 for arbitrary vertex cover problems is as follows:

**Mahdis Algorithm** (To produce a vertex cover solution on graph \(G\) with a ratio factor \(\rho = 1.999999\))

**Step 1.** Let \(V^1 = V^0 = \{\}\) and solve the LP relaxation of the model (1) on \(G\).

**Step 2.** If \(z_{LP \_relaxation}^1 < \frac{\rho}{2}\) then solve the model (7) to produce an extreme optimal solution of the LP relaxation of the model (1). Based on such a solution \((x_j^* \in \{0, 0.5, 1\} \quad j \in V\), introduce \(V^0 = \{j \in V \mid x_j^* = 0\}, V^{0.5} = \{j \in V \mid x_j^* = 0.5\}, V^1 = \{j \in V \mid x_j^* = 1\}\), and let \(G = G_{0.5}\) as the induced graph on the vertex set \(V^{0.5}\).

**Step 3.** Produce \(G^2\) based on \(G\) and solve the SDP (6) model.

**Step 4.** If for more than 0.000001n of vertices \(j \in V^n\) and corresponding vectors we have \(v_j^*v_j^* < 0.5\), then produce a suitable solution \(V_1 \cup V_0\), correspondingly, where \(V_0 = \{j \in V^n \mid v_j^*v_j^* < 0.5\}\) and \(V_1 = V^n - V_0\) and go to Step 9.

Hence, the solution does not meet the Property (1.a) and we have \(\frac{|V_1|}{|V_{VCP}|} \leq 1.999999\). Otherwise, go to Step 5.

**Step 5.** If for more than 0.000001n of vertices \(j \in V^n\) and corresponding vectors we have \(v_j^*v_j^* < 0.5\), then produce a suitable solution \(V_1 \cup V_0\), correspondingly, where \(V_0 = \{j \in V^n \mid v_j^*v_j^* < 0.5\}\) and \(V_1 = V^n - V_0\) and go to Step 9.

Hence, the solution does not meet the Property (1.a) and we have \(\frac{|V_1|}{|V_{VCP}|} \leq 1.999999\). Otherwise, go to Step 6.

**Step 6.** If for more than 0.01n of vertices \(j \in V^n\) and corresponding vectors, we have \(v_j^*v_j^* > 0.5004\), then it is sufficient to produce an arbitrary VCP feasible solution \(V = V_1 \cup V_0\) to have \(\frac{|V_1|}{|V_{VCP}|} \leq 1.999999\) and go to Step 9.

Otherwise, go to Step 7.

**Step 7.** If for more than 0.01n of vertices \(j \in V^n\) and corresponding vectors, we have \(v_j^*v_j^* > 0.5004\), then it is sufficient to produce an arbitrary VCP feasible solution \(V = V_1 \cup V_0\) to have \(\frac{|V_1|}{|V_{VCP}|} \leq 1.999999\) and go to Step 9.

Otherwise, go to Step 8.

**Step 8.** The solution meets the Property (1) and based on Theorem (6), the VCP problem on \(G'_x\) is simple and \(|V'_x| \geq 0.989999n\). Therefore, solve the VCP problem on bipartite subgraph \(G'_x\) to produce a feasible solution \(V_1 \cup V_0\) for which we have \(\frac{|V_1|}{|V_{VCP}|} \leq 1.999999\). Then, go to Step 9.

**Step 9.** The partitioning \((V_1 \cup V^1) \cup (V_0 \cup V^0)\) produces a VCP feasible solution on the original graph \(G\) with an approximation ratio factor \(\rho = 1.999999\).

**Corollary 3.** Based on the proposed 1.999999-approximation algorithm for the vertex cover problem, the unique games conjecture is not true.
4 Conclusions

One of the open problems about the vertex cover problem is the possibility of introducing an approximation algorithm within any constant factor better than 2. Here, we proposed a new algorithm to introduce a 1.999999-approximation ratio for the vertex cover problem on arbitrary graphs, and this lead to the conclusion that the unique games conjecture is not true.

Acknowledgment. I would like to extend my sincere appreciation to Dr. Flavia Bonomo for her useful comments. Moreover, I could publish in arXiv due to her endorsement.

Competing Interest and Data Availability
The authors have no relevant financial or non-financial interests to declare that are relevant to the content of this article. Data sharing is not applicable to this article as no data-sets were generated or analyzed during the current study.

References