LETTER *Special Section on Discrete Mathematics and Its Applications*

# Inapproximability of the Edge-Contraction Problem*

**Hideaki OTSUKI**[†a)] ***and*** **Tomio HIRATA**[††], ***Members***

**SUMMARY** For a property $\pi$ on graphs, the edge-contraction problem with respect to $\pi$ is defined as a problem of finding a set of edges of minimum cardinality whose contraction results in a graph satisfying the property $\pi$. This paper gives a lower bound for the approximation ratio for the problem for any property $\pi$ that is hereditary on contractions and determined by biconnected components.

***key words:*** *edge-contraction problem, NP-hard, approximation algorithm, approximability, connected vertex cover problem*

## 1. Introduction

The vertex-deletion and edge-deletion problems are natural graph modification problems. The vertex (edge) deletion problem is defined as a problem of finding a set of vertices (edges) of minimum cardinality whose deletion results in a graph satisfying the class of graph property $\pi$. For these problems, NP-completeness and approximation hardness have been studied [4], [5].

The edge-contraction problem is also a natural graph modification problem, but, to the authors' knowledge, its approximation hardness is not known. For a property $\pi$, the edge-contraction problem (EC) with respect to $\pi$ is defined as that of finding a set of edges of minimum cardinality whose contraction results in a graph satisfying the property $\pi$. If $\pi$ is hereditary on contractions and determined by biconnected components, the corresponding EC is NP-complete [1]. In [1], Asano and Hirata showed the NP-completeness of EC using a reduction from the connected vertex cover problem (CVC). The vertex cover problem is hard to approximate within a ratio $7/6$ [3], and it is easy to see that CVC has the same inapproximability as the vertex cover problem. However, the reduction in [1] does not conclude inapproximability of EC, since it does not have a gap preserving property [7].

Manuscript received August 22, 2005.
Manuscript revised November 4, 2005.
Final manuscript received December 10, 2005.

[†]The author is with the Department of Information and Telecommunicate Engineering, Nanzan University, Seto-shi, 489-0863 Japan.

[††]The author is with the Graduate School of Engineering, Nagoya University, Nagoya-shi, 464-8603 Japan.

[*]A preliminary version has appeared in the proceedings of the 2005 Korean-Japan Joint Workshop on Algorithms and Computation held in Seoul in Aug. 2005 and 4th Forum on Information Technology Letters (2005). A part of this work is supported by 2005 Nanzan University Pache Research Subsidy I-A-2.

a) E-mail: otsuki@nanzan-u.ac.jp

DOI: 10.1093/ietfec/e89–a.5.1425

In this paper, we give a lower bound for the approximation ratio for EC by the following steps. We construct an instance of CVC from that of MAX E3-SAT so that the reduction have a gap preserving property. Further, we reduce a CVC instance to that of EC. Finally, we establish a lower bound for the approximation ratio for EC.

## 2. Construction of an Instance of the Connected Vertex Cover Problem

CVC is a variant of the vertex cover problem which requires the subgraph induced by a cover-set must be connected. In this section we give a gap preserving reduction from MAX E3-SAT to CVC. We show that CVC on a certain class of graphs is hard to approximate within a ratio $41/40$.

### 2.1 Reduction from an Instance of MAX E3-SAT

MAX 3-SAT is the problem of finding a truth assignment which maximizes the number of satisfied clauses for a given 3-CNF $\phi$, and is known to be NP-complete. If each clause has exactly three literals, the problem is called as MAX E3-SAT and is also NP-complete [3]. Under the assumption that $P \neq NP$, it is not possible to approximate MAX E3-SAT within a ratio less than $8/7$ in polynomial time [3]. Here we construct a gap preserving reduction from an instance of MAX E3-SAT to that of CVC.

Let $n$ be the number of variables, and $m$ be the number of clauses. Let $x_i (i = 1, 2, \ldots, n)$ be the variables, and $C_j (j = 1, 2, \ldots, m)$ be the clauses. We assume that $x_i$ appears $t_i$ times in $\phi$. From $\phi$, we construct a graph $G = (V, E)$ as follows.

For each variable $x_i$, we have a set of vertices $X_i = \{x_i^j, \bar{x}_i^j | j = 1, 2, \ldots, t_i\}$ and a set of edges $E(x_i) = \{\{x_i^j, \bar{x}_i^{j'}\} | j, j' = 1, 2, \ldots, t_i\}$, which constructs a bipartite graph $K_{t_i,t_i} = G(x_i)$. We have vertices $c_0$ and $d_0$, an edge $e_0 = \{c_0, d_0\}$ and $E_{0i} = \{\{c_0, x_i^j\}, \{c_0, \bar{x}_i^j\} | j = 1, 2, \ldots, t_i\}$. For each clause $C_j (1 \leq j \leq m)$, we have vertices $c_j$, $d_j$ and an edge $e_j = \{c_j, d_j\}$. Edges between $c_j$ and $G(x_i)$'s vertices correspond to the literals in $C_j$ as follows. Let $l_1, l_2, l_3$ be the three literals in $C_j$. A literal $l_1$ is a variable $x_i$ or its negation $\bar{x}_i$, that appears at $l$th position in $\phi$. If the literal is $x_i$, we add an edge $e_j^1 = \{x_i^l, c_j\}$, otherwise $e_j^1 = \{\bar{x}_i^l, c_j\}$. We add edges $e_j^2, e_j^3$ in the same way for the literals $l_2, l_3$.

From this construction, we define a graph $G = (V, E)$ as

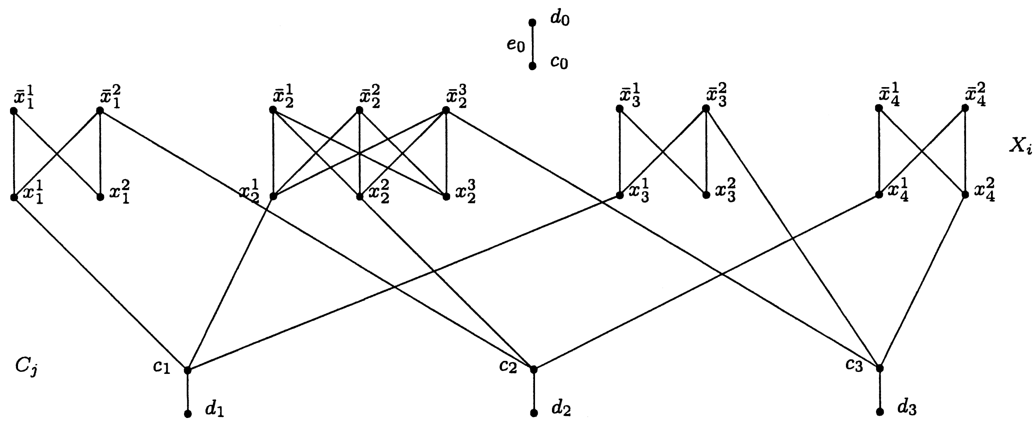


**Fig. 1** A CVC instance constructed from $\phi = (x_1 \vee x_2 \vee x_3) \wedge (\bar{x}_1 \vee x_2 \vee x_4) \wedge (\bar{x}_2 \vee \bar{x}_3 \vee x_4)$.

$$V = \{c_0, d_0\} \cup \bigcup_{i=1}^{n} X_i \cup \bigcup_{j=1}^{m} \{c_j, d_j\}$$

$$E = \{e_0\} \cup \bigcup_{i=1}^{n} (E_{0i} \cup E(x_i)) \cup \bigcup_{j=1}^{m} \{e_j, e_j^1, e_j^2, e_j^3\}.$$

For example, when $\phi = (x_1 \vee x_2 \vee x_3) \wedge (\bar{x}_1 \vee x_2 \vee x_4) \wedge (\bar{x}_2 \vee \bar{x}_3 \vee x_4)$, $G$ is illustrated as Fig. 1 (in which the edges in $E_{0i}$ are omitted).

Let $S_{cvc}$ be an optimal solution of CVC. We have the following lemma.

**Lemma 1:** If $\phi$ is satisfiable

$$|S_{cvc}| = 4m + 1.$$

**Proof.** Let $S$ be a solution of CVC and let $V(x_i) \equiv \{x_i^j | j = 1, 2, \ldots, t_i\}$, $V(\bar{x}_i) \equiv \{\bar{x}_i^j | j = 1, 2, \ldots, t_i\}$. In order to cover all edges of $G(x_i)$, we need

$$V(x_i) \subset S \tag{1}$$

or

$$V(\bar{x}_i) \subset S \tag{2}$$

for each $i$. In order to cover $e_0$, we need $c_0 \in S$ or $d_0 \in S$. As $\sum_{i=1}^{n} t_i = 3m$, in order to cover all edges of $G(x_i)$ and $E_{0i}$, we need at least $3m+1$ vertices. For each $j (1 \le j \le m)$, we need $c_j \in S$ or $d_j \in S$ to cover $e_j$. Hence we need $|S| \ge 4m + 1$.

In order to prove Lemma 1, it is sufficient to show the existence of a solution $S$ with $|S| = 4m + 1$. We construct $S$ from $\phi$ as follows. If $\phi$ assigns TRUE to $x_i$, we set $V(x_i)$ into $S$. Otherwise we set $V(\bar{x}_i)$ into $S$. We also include all $c_j (j = 1, 2, \ldots, m)$ to cover all $e_j$ and $e_j^i$. Since $\phi$ is satisfiable, each clause has at least one literal which is TRUE and thus each $c_j (j = 1, 2, \ldots, m)$ is connected with a vertex of $V(x_i)$ or $V(\bar{x}_i)$ in $S$. Now $|S| = 3m + m$, and all vertices in $S$ are connected. Further, we choose $c_0$ in $S$ so that $S$ covers $e_0$ and $E_{0i} (i = 1, 2, \ldots, n)$. $S$ induces a connected subgraph of $G$, and covers all of edges of $G$. $S$ is optimal since $|S| = 4m + 1$. □

We have another lemma.

**Lemma 2:** If no assignment satisfies more than $(1 - \epsilon)m$ clauses of $\phi$,

$$|S_{cvc}| \ge 4m + 1 + \epsilon m.$$

**Proof.** A solution $S$ of CVC induces an assignment $A$ of variables of $\phi$ as follows. If (1) holds and (2) does not, $A$ gives $x_i$ TRUE. If (2) holds and (1) does not, $A$ gives $x_i$ FALSE. If both (1) and (2) hold, $A$ gives $x_i$ either TRUE or FALSE. We say that this solution is consistent with the corresponding assignment $A$.

From the proof of Lemma 1, $|S| \ge 4m + 1$. Recall that $A$ does not satisfy at least $\epsilon m$ clauses. If $A$ does not satisfy a clause $C_j$, in order to connect $c_j$ with $S(Gx_i)$, $S$ must include a vertex of $G(x_i)$ corresponding to a literal to which $A$ assigns FALSE. So for any solution, in order to connect all $c_j (j = 1, 2, \ldots, m)$ with $S(Gx_i)$, additional $\epsilon m$ vertices of $S(Gx_i)$ must be included in $S$ and thus we have $|S| \ge 4m + \epsilon m + 1$. □

Now We have the following theorem.

**Theorem 1:** CVC for $G$ constructed above is NP-hard to approximate within a ratio 41/40.

**Proof.** From Lemma 1, Lemma 2 and $\epsilon = 1/8$, $m \ge 1$

$$\frac{4m + 1 + \epsilon m}{4m + 1} = 1 + \frac{\epsilon}{4 + \frac{1}{m}} \ge 1 + \frac{1}{40} = \frac{41}{40}.$$ □

## 3. Inapproximability of the Edge-Contraction Problem

From $G$ of the previous section, we construct an instance of the edge-contraction problem as follows. Let $G(2)$ be the graph obtained from $G$ by introducing a new vertex in the middle of each edge of $G$. That is, we replace each edge of $G$ with a path of length 2. We denote by $A(2)$ the set of newly introduced vertices. Let $M$ be a graph with the minimum number of vertices that violates $\pi$. Since $\pi$ is determined by biconnected components, $M$ is biconnected. Let $M - e$ be the graph obtained by deleting an edge $e$ from $M$. We construct $G_1$ from $G(2)$ as follows. For every pair $a$ and $a'$ of vertices in $A(2)$ which are adjacent to a common vertex in $V(G)$, we attach, to $a$ and $a'$, $k_1 + 1$ copies of $M - e$ through the node of $e$, where $k_1$ is an integer defined in the following proposition. Further, we denote by $S_{ec}$ an optimal solution of the edge-contraction problem of $G_1$.

**Proposition 1** (Asano and Hirata [1]): There is a subset $S$ of $E(G_1)$ with $|S| \leq k_1$ such that the contraction $G_1/S$ satisfies $\pi$ if and only if $G$ has a connected vertex cover of size $\leq k$, where $k_1 = k + |E(G)| - 1$.

We denote $S_{cvc}$ as an optimal solution of CVC in case that $\phi$ has a satisfiable assignment, and denote $S'_{cvc}$ otherwise. From the proposition, the size of the optimal solution of EC is $|S_{cvc}| + |E(G)| - 1$ if $\phi$ is satisfiable, and it is at least $|S'_{cvc}| + |E(G)| - 1$ if $\phi$ is unsatisfiable. So it is NP-hard for EC with respect to the property $\pi$ to approximate within a ratio

$$r_{ec} = \frac{|S'_{cvc}| + |E(G)| - 1}{|S_{cvc}| + |E(G)| - 1}.$$

From an instance of CVC which is reduced from an instance of MAX E3-SAT, we have

$$|E(G)| = m + 3m + 6m + 1 + \sum_{i=1}^{n} t_i^2 = 10m + 1 + \sum_{i=1}^{n} t_i^2.$$

Further, if the number of appearance of all variables in $\phi$ is constant(= $l$), $\sum_{i=1}^{n} t_i^2 = nl^2 = 3ml$. We use $\epsilon_l$ instead of $\epsilon$ in this case. By Lemma 1 and Lemma 2, $|S| = 4m + 1, |S'| \geq 4m + 1 + \epsilon_l m$. We conclude

$$r_{ec} = 1 + \frac{\epsilon_l}{14 + 3l + 1/m} > 1 + \frac{\epsilon_l}{15 + 3l}.$$

Now we have the following theorem.

**Theorem 2:** There is a constant $r$ so that $r$-approximation of the edge-contraction problem of $G(2)$ is NP-hard.

Papadimitriou and Yannakakis [6] showed that in case of $l = 29$, $\epsilon_l = 1/(8 \cdot 43) = 0.0029069767$. Hence we have $r = \epsilon_l/102 = 1.00002849977$.

Replacing all edges in $M$ with a path of length 2, we can make $G_1$ bipartite. Since $\pi$ is hereditary on contraction, Proposition 1 still holds. In this case, we need $\pi$ to be "determined by 3-connected components". We omit details. See Corollary 4 of [1]. We have another theorem.

**Theorem 3:** There is a constant $r$ so that $r$-approximation of the edge-contraction problem for $\pi$, restricted to bipartite graphs is NP-hard, where $\pi$ is hereditary on contractions, and determined by 3-connected components.

## 4. Conclusions

We have shown that when a graph property $\pi$ is hereditary on contractions and determined by biconnected components, the edge-contraction problem with respect to $\pi$ is hard to approximate within a ratio $1 + \epsilon_l/(15 + 3l)$, where $l$ is the number of appearance of each variable in MAX-E3 SAT, and $\epsilon_l$ is a ratio with which the approximation of MAX-E3 SAT is NP-hard. Furthermore, we have the same result for bipartite graphs when $\pi$ is hereditary on contractions and determined by 3-connected components. Our future work is to seek a larger lower bound of the approximation ratio for EC with respect to $\pi$ and inapproximability results of EC with respect to properties other than $\pi$ considered here.

### References

[1] T. Asano and T. Hirata, "Edge-contraction problems," J. Comput. Syst. Sci., vol.26, pp.197–208, 1983.
[2] M.R. Garey and D.S. Johnson, "The rectilinear Steiner tree problem is *NP*-complete," SIAM J. Appl. Math., vol.32, pp.826–834, 1977.
[3] J. Håstad, "Some optimal inapproximability results," J. ACM, vol.48, pp.798–859, 2001.
[4] P.G. Kolaitis and M.N. Thakur, "Approximation properties of NP minimization classes," J. Comput. Syst. Sci., vol.50, pp.391–411, 1995.
[5] C. Lund and M. Yannakakis, "The approximation of maximum subgraph problems," Proc. 20th Int. Colloquium on Automata, Languages and Programming, vol.700, pp.40–51, 1993.
[6] C.H. Papadimitriou and M. Yannakakis, "Optimization, approximation, and complexity classes," J. Comput. Syst. Sci., vol.43, pp.425–440, 1991.
[7] V.V. Vazirani, Approximation Algorithms, Springer Verlag, 2001.