Algorithmica (2011) 60: 395–400
DOI 10.1007/s00453-009-9349-5

Approximability of Packing Disjoint Cycles

Zachary Friggstad · Mohammad R. Salavatipour

Received: 28 August 2008 / Accepted: 26 July 2009 / Published online: 11 August 2009
© Springer Science+Business Media, LLC 2009

Abstract Given a graph G, the edge-disjoint cycle packing problem is to find the largest set of cycles of which no two share an edge. For undirected graphs, the best known approximation algorithm has ratio $O(\sqrt{\log n})$ (Krivelevich et al. in ACM Trans. Algorithms, 2009, to appear). In fact, they proved the same upper bound for the integrality gap of this problem by presenting a simple greedy algorithm. Here we show that this is almost best possible. By modifying integrality gap and hardness results for the edge-disjoint paths problem (Andrews et al. in Proc. of 46th IEEE FOCS, pp. 226–244, 2005; Chuzhoy and Khanna in New hardness results for undirected edge disjoint paths. Manuscript, 2005), we show that the undirected edge-disjoint cycle packing problem is quasi-NP-hard to approximate within ratio of $O(\log^{\frac{1}{2}-\epsilon} n)$ for any constant $\epsilon > 0$. The same result holds for the problem of packing vertex-disjoint cycles.

Keywords Approximation · Hardness of approximation · Edge-disjoint cycles · Edge-disjoint paths · Integrality gap

1 Introduction

In the problem of *edge-disjoint cycle packing* (EDC) we are given a graph G and our goal is to find a largest set of edge-disjoint cycles. The vertex analog of the problem,

A preliminary version appeared [9] in Proceedings of ISAAC 2007.

The first author was supported by NSERC and iCore scholarships and the 2nd author was supported by NSERC and an Alberta Ingenuity New Faculty award.

Z. Friggstad · M.R. Salavatipour (✉)
Department of Computing Science, University of Alberta, Edmonton, Alberta T6G 2E8, Canada
e-mail: mreza@cs.ualberta.ca

Z. Friggstad
e-mail: zacharyf@cs.ualberta.ca

vertex-disjoint cycle packing (VDC), is the problem of finding a largest set of vertex-disjoint cycles in the given graph. The EDC problem has been studied extensively in both directed and undirected settings (e.g. see [3, 4, 15]). A discussion on the applications of packing cycles to computational biology and reconstructing evolutionary trees can be found in [3]. Both EDC and VDC are known to be NP-hard even for undirected graphs and for very restricted cases of the problem [8]. This motivates the study of approximation algorithms for these problems. Caprara, Panconesi and Rizzi [4] showed that EDC is APX-hard even when restricted on planar graphs. They also presented a simple greedy algorithm with approximation ratio $O(\log n)$. Recently, Krivelevich et al. [13] showed that a modification of the simple greedy algorithm of [4] with a more careful analysis yields an $O(\sqrt{\log n})$-approximation for EDC on undirected graphs. In fact, the algorithm obtains an integer solution that is within factor $O(\sqrt{\log n})$ of the optimal fractional solution. They showed examples for which the solution obtained by the greedy algorithm was within $\Omega(\sqrt{\log n})$ of the optimal solution but it falls short of proving any super-constant lower bound on the integrality gap or approximability of the problem. They also presented an $O(\sqrt{n})$-approximation for EDC on directed graphs and an $O(\log n)$-approximation for undirected VDC. On the other hand, they provided an integrality gap of $\Omega(\frac{\log n}{\log\log n})$ and a quasi-NP-hardness of $\Omega(\log^{1-\epsilon} n)$, for any constant $\epsilon > 0$ for EDC on directed graphs. However, the best known lower bound on the approximability of EDC on undirected graphs remains APX-hardness and the best lower bound for integrality gap is $O(1)$, prior to our work. For the related problem of edge-disjoint paths (EDP), on directed graphs the best approximation algorithms have ratio $O(\min\{n^{\frac{2}{3}}\log^{\frac{1}{3}} n, \sqrt{m}\})$ [12, 16] and it is known the problem is hard to approximate within $O(m^{\frac{1}{2}-\epsilon})$ for any $\epsilon > 0$ [10]. For undirected graphs, the best known approximation ratio for EDP is $O(\sqrt{n})$ [5] whereas the best known hardness result is only $\Omega(\log^{\frac{1}{2}-\epsilon} n)$ for any $\epsilon > 0$ [1, 7]. The latter result was built on the major advance on the lower bound of EDP (from APX-hardness to $\Omega(\log^{\frac{1}{3}-\epsilon} n)$) by Andrews and Zhang [2]. In this paper, we improve the lower bounds for both EDC and VDC.

Theorem 1.1 *The EDC problem on undirected graphs is hard to approximate within* $O(\log^{\frac{1}{2}-\epsilon} n)$ *for any* $\epsilon > 0$ *unless* $\mathrm{NP} \subseteq \mathrm{ZPTIME}(n^{\mathrm{polylog}(n)})$.

This shows that the simple greedy algorithm of [13] with approximation ratio $O(\sqrt{\log n})$ is almost best possible for EDC. The same reduction works to prove the same hardness result for VDC. This result is obtained through a modification of the hardness of the edge-disjoint paths problem presented by Chuzhoy and Khanna in [7] (see also [1]). Nevertheless, it shows a rather surprising approximability threshold for a very natural packing problem. In fact there are very few problems known to have a sub-logarithmic approximability threshold [6, 11]. One other important message to be taken from our results is that, in order to improve the hardness of approximation for EDP (from $\Omega(\log^{\frac{1}{2}-\epsilon} n)$ to beyond $\Omega(\sqrt{\log n})$), there has to be substantially new ideas developed that exploit the differences between EDC and EDP problems; since such a hardness result should not be adaptable to work for EDC (because we already have an $O(\sqrt{\log n})$-approximation for EDC).

The reduction is heavily based on the reduction of [7] for EDP, so we do not repeat the details and proofs presented in [7]. This paper is best read with a copy of [7] at hand. The reader may also refer to [9].

2 The Hardness Construction

In this section we prove Theorem 1.1. We show how a modification of the construction used to prove the hardness of approximating edge-disjoint paths by Chuzhoy and Khanna in [7] can be used to show the same hardness for EDC.

They [7] start from a PCP characterization of NP with the following properties. Let Φ be an instance of 3SAT with n variables. For any constant $k > 0$, consider a PCP verifier that uses $r = O(\log n)$ random bits and queries $q = k^2$ bits of a proof Π. Let R be a random string of length r and $b_1(R), \dots, b_q(R)$ be the indices of the bits of the proof given R. Define a *configuration* to be the tuple $(R, a_1, \dots, a_q)$ where $a_i = \Pi_{b_i(R)} \in \{0, 1\}$, for $1 \le i \le q$, are the values of the bits read in the proof given random string R. A configuration $(R, a_1, \dots, a_q)$ is called *accepting* if the PCP verifier accepts upon using random string R and reading proof bits $a_1, \dots, a_q$. It follows ([7]) from the construction of [14] that for every constant $k > 0$ and for sufficiently large constant $\beta >> k^2$ there exists a PCP verifier for Φ such that:

- $\lambda r = O(\log n \log\log n)$ random bits are used with $r = O(\log n)$ and $\lambda = \frac{2\beta \log\log n}{k^2}$.
- Exactly $q = \lambda k^2 = O(\log\log n)$ bits of the proof are queried for each random string.
- If Φ is satisfiable, then there is a proof Π such that the acceptance probability of the verifier upon reading Π is at least $2^{-\lambda}$ and if Φ is not satisfiable, then the acceptance probability of the verifier upon reading Π is at most $2^{-\lambda k^2}$ for all proofs Π.
- Every random string R participates in $2^{\lambda(2k-1)}$ accepting configurations.
- For every random string R and for every $j = 1 \dots q$, the number of accepting configurations with $\Pi_{b_j(R)} = 0$ and the number of accepting configurations with $\Pi_{b_j(R)} = 1$ are equal.
- Let $\mathcal{A}$ be the set of all accepting configurations, and Z_j and O_j be the sets of all accepting configurations with $\Pi_j = 0$ and $\Pi_j = 1$, respectively. Let $n_j = |Z_j| = |O_j|$. Then $n_j \ge 2^{\lambda r/2}$ and $|\mathcal{A}| \le 2^{\lambda r} \cdot 2^{2\lambda k}$.

The authors of [7] construct a bit gadget $G(i)$ for each bit of Π_i of the proof. For each Π_i and for each configuration $\alpha \in Z_i \cup O_i$, for a parameter X to be defined soon, $G(i)$ has X starting vertices and X end vertices and a collection of paths, called *canonical paths*, each connecting one start and end vertex. Then using this bit gadget as a building block, they [7] build an instance $\mathcal{I}$ of the EDP problem. Starting from instance Φ of 3SAT with n variables and using the aforementioned PCP, with parameters $X = 2^{2^{2\lambda(k^2+4k)}}$ and $M = 2^{\lambda(k^2+k)}$, the final instance $\mathcal{I}$ for EDP has $X \cdot 2^{\lambda r} \cdot 2^{\lambda(2k-1)}$ pairs and has size $N \le X \cdot 2^{\lambda r} \cdot M \cdot 2^{2\lambda k} \le X \cdot 2^{O(\log n \log\log n)}$. We use the same bit gadget construction and the same set of parameters as in [7]. For each index i of proof Π, let $\P_0(i)$ be the set of canonical paths (of $G(i)$) corresponding to a configuration in Z_i and $\P_1(i)$ be the set of canonical paths corresponding to a configuration in O_i. Following their definition, a bit gadget $G(i)$ is bad if there exist

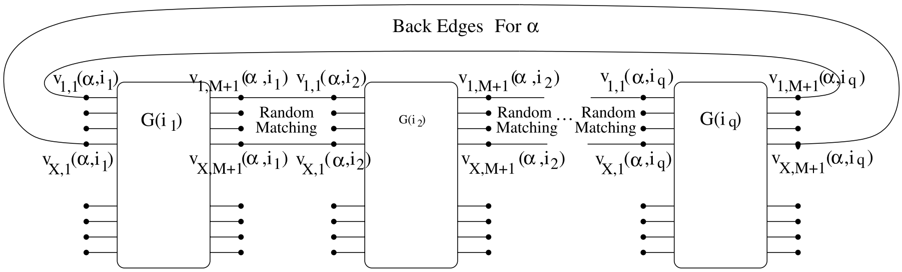

Fig. 1 The final instance for configuration α

$A \subseteq P_0(i)$, $B \subseteq P_1(i)$ with $|A| = |B| = \frac{8 \log MXn_i}{M}$ such that all paths in $A \cup B$ are edge disjoint. Define $\mathcal{B}_1$ to be the event that there is some bit gadget that is bad. The following lemma is proved in [7].

Lemma 2.1 [7] *The probability that bad event $\mathcal{B}_1$ happens is at most* $\frac{1}{\mathrm{poly}(n)}$.

For each accepting configuration α of the PCP and each $1 \le x \le X$, the authors of [7] define a canonical path $P_x(\alpha)$ that passes through canonical paths of all bit gadgets corresponding to bits read in α. If $s_x(\alpha)$ and $t_{x'}(\alpha)$ are the endpoints of $P_x(\alpha)$, then we add a *back* edge connecting $t_{x'}(\alpha)$ and $s_x(\alpha)$. Finally, for each accepting configuration α and each $1 \le x \le X$, we define a canonical cycle $C_x(\alpha)$ which is simply the canonical path $P_x(\alpha)$ followed by the back edge from $t_{x'}(\alpha)$ to $s_x(\alpha)$. This is our only modification to the reduction. Note that all vertices are of degree at most 3. Figure 1 illustrates this modified construction.

Following [7], let C_{YI} denote the maximum number of edge-disjoint cycles that can be packed in G if Φ is satisfiable. The following lemmas hold by essentially the same arguments as analogous results in [7]. We skip repeating the proofs here. Let $\mathcal{C}$ be the set of all canonical cycles.

Lemma 2.2 $C_{YI} \ge \frac{|\mathcal{C}|}{2^{2\lambda k}}$.

If Φ is not satisfiable, then let $\mathcal{C}'$ be a collection of edge-disjoint cycles in G. Define $g = 2^{2\lambda(k^2+k)}$. say a cycle is short if its length is less than g; otherwise say the cycle is long. Partition $\mathcal{C}$ into sets $\mathcal{C}_1, \mathcal{C}_2$, and $\mathcal{C}_3$ where $\mathcal{C}_1$ is the set of all canonical cycles in $\mathcal{C}$, $\mathcal{C}_2$ is the set of long non-canonical cycles, and $\mathcal{C}_3$ is the set of short non-canonical cycles.

Lemma 2.3 *If bad event $\mathcal{B}_1$ does not happen, then* $|\mathcal{C}_1| \le \frac{2C_{YI}}{2^{\lambda k^2 - 2\lambda k - \lambda}}$.

Lemma 2.4 $|\mathcal{C}_2| \le \frac{|\mathcal{C}|}{2^{\lambda k^2}} \le \frac{C_{YI}}{2^{\lambda k^2 - 2\lambda k}}$.

To bound the number of short non-canonical cycles, we must make a modification to the proof in [7] because we have introduced back edges. For that we first define bad event $\mathcal{B}_2$ as the event $|\mathcal{C}_3| > \frac{C_{YI}}{2^{\lambda k^2}}$.

Lemma 2.5 *Event $\mathcal{B}_2$ happens with probability at most $\frac{1}{3}$.*

Proof Let G' be the resultant graph when all of the special edges of the bit gadgets in G are contracted. The lemma follows by the same arguments in [7] if we can establish the following claim.

Claim 2.6 *The probability of each edge $e = uv$ appearing in the graph G' given the existence of $g' - 1$ other edges that do not form a canonical path from u to v, is at most $\frac{1}{X-g'+1}$.*

This is easy to see for the case of a non-back-edge (*i.e.* random matching edge) as each matching edge exists with probability at most $\frac{1}{X-g'+1}$ given the existence of $g' - 1$ other edges. The case of a potential back-edge is different as the back-edges are not completely random (each is created between the source and sink of a canonical path; but the path is created randomly). Consider a potential back edge $e = uv$ between a source node u and a sink node v (note that u and v are not necessarily the end points of a canonical path) and suppose we are given the existence of up to $g' - 1$ other edges that do not form a canonical path from u to v. Moreover, consider the partial canonical paths from u and from v using the other at most $g' - 1$ other edges. Since there is currently no canonical path from u to v (otherwise we have a canonical cycle with e), then the probability that u and v are endpoints of the same canonical path is exactly the probability that they will be connected with a new random-matching edge. Thus, the probability that e exists is at most $\frac{1}{X-g'+1}$. □

Therefore, if event $\mathcal{B}_2$ does not happen, then: $|\mathcal{C}_3| \leq 2^{4\lambda r g} \leq 2^{2^{2\lambda(k^2+3k)+\log\log n}}$, because $r = O(\log n)$. Also, since $\lambda = \beta \log\log n / k^2$ for $\beta >> k^2$, then $\lambda k \geq \log\log n$ resulting in $|\mathcal{C}_3| \leq 2^{2^{2\lambda(k^2+4k)}} \leq X \leq \frac{C_{YI}}{2^{\lambda(r-1)}} \leq \frac{C_{YI}}{2^{\lambda k^2}}$.

If neither of bad events $\mathcal{B}_1$ nor $\mathcal{B}_2$ happens, then $|\mathcal{C}'| = |\mathcal{C}_1| + |\mathcal{C}_2| + |\mathcal{C}_3| \leq \frac{C_{YI}}{2^{\lambda(k^2-3k)}}$. Following the arguments in [7], we see that this gap is at least $\log^{\frac{1}{2}-\epsilon} N$ by choosing an appropriately large $k = k(\epsilon)$. The probability of either of events $\mathcal{B}_1$ or $\mathcal{B}_2$ occurring is at most $1/(\text{poly}(n)) + 1/3 \leq 1/2$. So, if an $O(\log^{\frac{1}{2}-\epsilon} n)$-approximation algorithm exists for undirected EDC for any $\epsilon > 0$, then a co-RPTIME($n^{\text{polylog}(n)}$) algorithm for 3SAT exists, which in turn implies the existence of a ZPTIME($n^{\text{polylog}(n)}$) algorithm for 3SAT by a standard result. Thus, for any $\epsilon > 0$, it is hard to approximate the edge-disjoint cycle packing problem within a factor of $\Omega(\log^{\frac{1}{2}-\epsilon} n)$ unless NP $\subseteq$ ZPTIME($n^{\text{polylog}(n)}$).

3 Concluding Remarks

Since each vertex has degree at most 3 in the construction of G, it is easy to see that we get a similar hardness for the vertex-disjoint cycle packing problem. Theorem 1.1 together with the results of [13] yield an almost tight ratio for approximability of EDC in the undirected setting ($O(\sqrt{\log n})$ v.s. $\Omega(\log^{\frac{1}{2}-\epsilon} n)$ for any $\epsilon > 0$). However, the gap between the best approximation ratio and hardness lower bounds for undirected VDC as well as directed EDC (and VDC) are pretty wide.

References

1. Andrews, M., Chuzhoy, J., Khanna, S., Zhang, L.: Hardness of the undirected edge-disjoint paths problem with congestion. In: Proc. of 46th IEEE FOCS, pp. 226–244 (2005)
2. Andrews, M., Zhang, L.: Hardness of the undirected edge-disjoint paths problem. In: Proc. of 37th ACM STOC, pp. 276–283. ACM Press, New York (2005)
3. Balister, P.: Packing digraphs with directed closed trials. Comb. Probab. Comput. **12**, 1–15 (2003)
4. Caprara, A., Panconesi, A., Rizzi, R.: Packing cycles in undirected graphs. J. Algorithms **48**, 239–256 (2003)
5. Chekuri, C., Khanna, S., Shepherd, B.: An $O(\sqrt{n})$ approximation and integrality gap for disjoint paths and UFP. Theory Comput. **2**, 137–146 (2006)
6. Chuzhoy, J., Guha, S., Halperin, E., Khanna, S., Kortsarz, G., Krauthgamer, R., Naor, S.: Tight lower bounds for the asymmetric k-center problem. J. ACM **52**(4), 538–551 (2005)
7. Chuzhoy, J., Khanna, S.: New hardness results for undirected edge disjoint paths. Manuscript (2005)
8. Dor, D., Tarsi, M.: Graph decomposition is NPC—A complete proof of Holyer's conjecture. In: Proc. of 20th ACM STOC, pp. 252–263. ACM Press, New York (1992)
9. Friggstad, Z., Salavatipour, M.R.: Approximability of packing disjoint cycles. In: Proc. of the 18th Int. Symp. on Algorithms and Computation (ISAAC) 2007. LNCS, pp. 304–315. Springer, Berlin (2007)
10. Guruswami, V., Khanna, S., Rajaraman, R., Shepherd, B., Yannakakis, M.: Near-optimal hardness results and approximation algorithms for edge-disjoint paths and related problems. J. Comput. Syst. Sci. **67**(3), 473–496 (2003). Earlier version in STOC'99
11. Halldorsson, M.M., Kortsarz, G., Radhakrishnan, J., Sivasubramanian, S.: Complete partitions of graphs. Combinatorica **27**(5), 519–550 (2007)
12. Kleinberg, J.: Approximation algorithms for disjoint paths problems. PhD. Thesis, MIT, Cambridge, MA, May 1996
13. Krivelevich, M., Nutov, Z., Salavatipour, M.R., Verstraete, J., Yuster, R.: Approximation algorithms and hardness results for cycle packing problems. ACM Trans. Algorithms **3**(4), 48 (2007)
14. Samorodnitsky, A., Trevisan, L.: A PCP characterization of NP with optimal amortized query complexity. In: Proc. of 32nd ACM STOC, pp. 191–199. ACM Press, New York (2000)
15. Seymour, P.D.: Packing directed circuits fractionally. Combinatorica **15**, 281–288 (1995)
16. Varadarajan, K.R., Venkataraman, G.: Graph decomposition and a greedy algorithm for edge-disjoint paths. In: Proc. of 15 ACM-SIAM SODA, pp. 379–380 (2004)