Available online at www.sciencedirect.com
ScienceDirect
Discrete Applied Mathematics 156 (2008) 2270–2278
DISCRETE APPLIED MATHEMATICS
www.elsevier.com/locate/dam

Partition into cliques for cubic graphs: Planar case, complexity and approximation ☆

M.R. Cerioli [a,b], L. Faria [c], T.O. Ferreira [b], C.A.J. Martinhon [d], F. Protti [a,e,*], B. Reed [f]

[a] IM - UFRJ, Rio de Janeiro, Brazil
[b] COPPE - UFRJ, Rio de Janeiro, Brazil
[c] DMAT, FFP - UERJ, Rio de Janeiro, Brazil
[d] IC - UFF, Rio de Janeiro, Brazil
[e] NCE - UFRJ, Caixa Postal 2324, 20001-970, Rio de Janeiro, Rj, Brazil
[f] School of Computer Science, McGill University, Montreal, Quebec, Canada

Received 12 July 2005; received in revised form 10 August 2007; accepted 23 October 2007
Available online 20 February 2008

Abstract

Given a graph $G = (V, E)$ and a positive integer k, the PARTITION INTO CLIQUES (PIC) decision problem consists of deciding whether there exists a partition of V into k disjoint subsets $V_1, V_2, \dots, V_k$ such that the subgraph induced by each part V_i is a complete subgraph (clique) of G. In this paper, we establish both the NP-completeness of PIC for planar cubic graphs and the Max SNP-hardness of PIC for cubic graphs. We present a deterministic polynomial time $\frac{5}{4}$-approximation algorithm for finding clique partitions in maximum degree three graphs.

Keywords: NP-complete; Max SNP-hard; Partition into cliques; Coloring; Approximation; Cubic planar graphs

1. Introduction

The PARTITION INTO CLIQUES (PIC) decision problem consists of a graph $G = (V, E)$ and a positive integer k, plus the question of deciding whether there exists a partition of V into k disjoint subsets $V_1, V_2, \dots, V_k$ such that, for every $i \in \{1, 2, \dots, k\}$, the subgraph of G induced by V_i is a complete graph (clique). In case the answer is YES we call $(V_1, \dots, V_k)$ a *partition into k cliques* for the graph $G = (V, E)$.

The nomenclature "PARTITION INTO CLIQUES" is used in the book of Garey and Johnson [6]. Karp [11] named it "VERTEX COVER", and formerly proved that it is NP-complete for general graphs.

A natural question when studying the complexity of a graph-theoretical decision problem is to determine for which special graph classes and upper bounds on the vertex degrees the problem remains hard. In this regard, Garey and Johnson described a proof [6] where PIC has been proved to be NP-complete for K_4-free graphs, and in a paper with

☆ Partially supported by CNPq, CAPES, FAPERJ, FINEP.
* Corresponding author at: NCE - UFRJ, Caixa Postal 2324, 20001-970, Rio de Janeiro, Rj, Brazil.
E-mail addresses: cerioli@cos.ufrj.br (M.R. Cerioli), luerbio@cos.ufrj.br (L. Faria), talita@cos.ufrj.br (T.O. Ferreira), mart@dcc.uff.br (C.A.J. Martinhon), fabiop@nce.ufrj.br (F. Protti), breed@cs.mcgill.ca (B. Reed).

0166-218X/$ - see front matter
doi:10.1016/j.dam.2007.10.015

Table 1
Results on PIC and COLORING

Graph G (instance for PIC)	Graph $\bar{G}$ (instance for COLORING)	Result
General	General	NP-complete [11] $\exists\ n^{\frac{(\log\log n)^2}{(\log n)^3}}$-approx. [8] $\nexists\ n^{\frac{1}{7}-\varepsilon}$-approx. if P $\neq$ NP [2]
$\delta \geq n-5$	$\Delta \leq 4$	NP-complete [7]
$\delta \geq n-4$	$\Delta \leq 3$	poly-time solvable [4]
$\Delta \leq 2$	$\delta \geq n-3$	poly-time solvable [5]
Planar cubic	Co-planar $(n-4)$-regular	NP-complete
Cubic	$(n-4)$-regular	Max SNP-hard
$\Delta \leq 3$	$\delta \geq n-4$	$\exists\ \frac{5}{4}$-approximation

Stockmeyer [7] they proved that PIC is NP-complete for maximum degree four graphs. Recently, Hunt III et al. [10] improved the result establishing that PIC is NP-complete for planar graphs with maximum degree four.

We notice that as PIC is the same problem as COLORING in the complement graph, coloring complexity results can be converted, using the complement graph, to results on PIC. For instance, given an input graph $G = (V, E)$ with $|V| = n$, it is known that COLORING is a polynomial-time solvable problem [4] when G is a maximum degree three graph, and NP-complete when G is a maximum degree four graph [7], hence PIC is polynomial-time solvable for a minimum degree $n-4$ graph, and NP-complete for a minimum degree $n-5$ graph G. Additionally, Bellare et al. [2] proved that, unless $P = NP$, there is no approximation algorithm for COLORING within $n^{\frac{1}{7}-\varepsilon}$ for any $\varepsilon > 0$, and Halldórsson [8] showed that there is a polynomial-time approximation algorithm for COLORING within $O(n^{\frac{(\log\log n)^2}{(\log n)^3}})$ for general graphs.

In the present work, we prove that PIC remains NP-complete even when restricted to planar cubic graphs. This NP-completeness result is obtained by using the NP-complete problem PLANAR 3SAT WITH EXACTLY THREE OCCURRENCES PER VARIABLE (P3SAT $\underline{3}$). This problem is proved to be NP-complete using the wheel of implications defined in [3], which extends a previous result in [13], where PLANAR 3SAT was shown to be NP-complete. We consider the optimization version MINPIC of PIC. We prove that MINPIC is Max SNP-hard for cubic graphs using the Max SNP-complete [1,15] problem MAX 3-SATISFIABILITY WITH EXACTLY THREE OCCURRENCES PER VARIABLE (MAX3SAT$\underline{3}$). We show an approximation algorithm for MINPIC in the class of maximum degree three graphs achieving a performance guarantee of $\frac{5}{4}$. Our result improves a previous result due to Halldórsson [9], which achieves an approximation ratio $\frac{4}{3}$ for SET-COVERING problem with sets of size at most three. Table 1 summarizes the literature results relating PIC and COLORING in graph classes classified according to the maximum degree Δ and the minimum degree δ, including the results presented in this paper.

2. The results

We shall use the NP-complete problem 3SAT$\underline{3}$ [6] to prove that PIC is NP-complete, where 3SAT$\underline{3}$ has this following special version:

3SAT$\underline{3}$- 3SAT WITH EXACTLY 3 OCCURRENCES PER VARIABLE (decision)

INSTANCE: $I = (U, C)$ in CNF, where U is a set of variables and C a collection of clauses over U such that: (i) each clause $c \in C$ satisfies $|c| = 3$ or $|c| = 2$; (ii) each variable has exactly three occurrences in C being exactly one negative.

QUESTION: Is there a satisfying truth assignment for U satisfying all $c \in C$?

We remark that if for this version of 3SAT$\underline{3}$ every clause has size $|c| = 3$, then the problem is polynomial-time solvable [15]. An instance $I = (U, C)$ of 3SAT$\underline{3}$ is said to be an instance of PLANAR 3SAT WITH EXACTLY 3 OCCURRENCES PER VARIABLE (P3SAT$\underline{3}$) [3] if the bipartite graph H is planar, where H has parts U and C, and there is an edge $uc \in E(H)$ if and only if literal u or $\overline{u}$ belongs to clause c.

We shall consider MINIMUM PARTITION INTO CLIQUES (MINPIC) to refer to the minimization problem of finding the minimum k such that there exists a partition into k cliques for an input graph $G = (V, E)$. Analogously, we

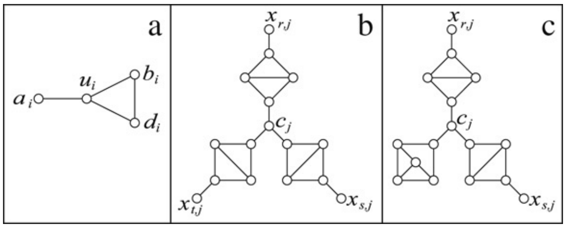

Fig. 1. Truth-Setting subgraph T_i (a) corresponding to variable u_i. Satisfaction-Testing subgraphs (b) $S_j = S_j^3$ and (c) $S_j = S_j^2$ corresponding, respectively, to clauses $c_j = (x_r \vee x_s \vee x_t)$ and $c_j = (x_r \vee x_s)$.

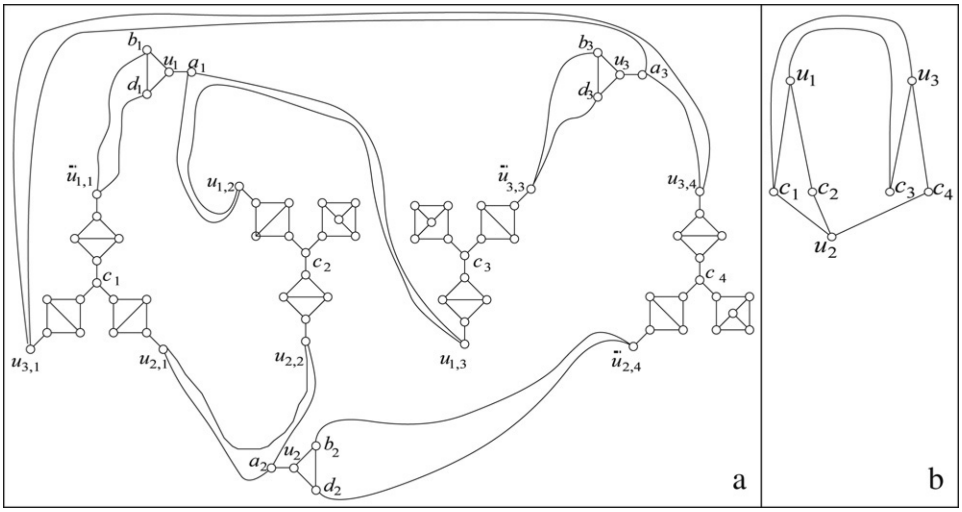

Fig. 2. Graph G obtained from the $\text{MAX3SAT}_{\underline{3}}$ instance $I = (U, C) = (\{u_1, u_2, u_3\}, \{(\overline{u}_1 \vee u_2 \vee u_3), (u_1 \vee u_2), (u_1 \vee \overline{u}_3), (\overline{u}_2 \vee u_3)\})$, $I = (U, C)$ is also a $\text{P3SAT}_{\underline{3}}$ instance with bipartite plane graph depicted in (b).

consider $\text{MAX3SAT}_{\underline{3}}$ and $\text{MAXP3SAT}_{\underline{3}}$ to refer to the maximization problems of finding the maximum k such that there exists a truth assignment for U satisfying k clauses of the instance $I = (U, C)$.

Given a $\text{MAX3SAT}_{\underline{3}}$ instance $I = (U, C)$, we construct in polynomial time in the size of I a special instance $G = (V, E)$ for MINPIC. This construction is used in order to obtain both the NP-completeness and the Max SNP-hardness results.

For each variable $u_i \in U$ there is one *Truth Setting subgraph* T_i defined in Fig. 1(a). For each clause $c_j \in C$ there is one *Satisfaction Testing subgraph* S_j, which can be of two types defined in Fig. 1(b) and (c), according to either $S_j = S_j^3$ corresponds to a clause $c_j = (x_r \vee x_s \vee x_t)$ with three literals, or $S_j = S_j^2$ corresponds to a clause $c_j = (x_r \vee x_s)$ with two literals.

The only part in the construction of G that depends on which literals occur in which clauses consists of the following sets of edges. For each variable $u_i \in U$ occurring twice positively in clauses c_j and c_k, and once negatively in the clause c_ℓ, we add to $E(G)$ the set of edges $E_i = \{a_i u_{i,j}, a_i u_{i,k}, u_{i,j} u_{i,k}, b_i \overline{u}_{i,\ell}, d_i \overline{u}_{i,\ell}\}$, where $u_{i,j}, u_{i,k}$ and $\overline{u}_{i,\ell}$ denote, respectively, the vertices corresponding to the occurrences of literal u_i in clauses c_j and c_k, and the occurrence of literal $\overline{u}_i$ in the clause c_ℓ.

For the convenience of the reader we show in Fig. 2 an example of a special instance G obtained from a $\text{MAX3SAT}_{\underline{3}}$ instance. Next we state (without proof) two claims that can be easily checked.

Claim 1. *$V(S_j)$ can be partitioned into 7 cliques, and cannot be partitioned in less than 7 cliques.* □

Claim 2. *Let $X \subseteq \{x_{r,j}, x_{s,j}, x_{t,j}\}$. Then $V(S_j) \setminus X$ can be partitioned into 6 cliques, and cannot be partitioned in less than 6 cliques.* □

We say that the graph K_3 is a *triangle* and that the graph $K_4 - e$ is a *diamond*, where e is an edge of K_4.

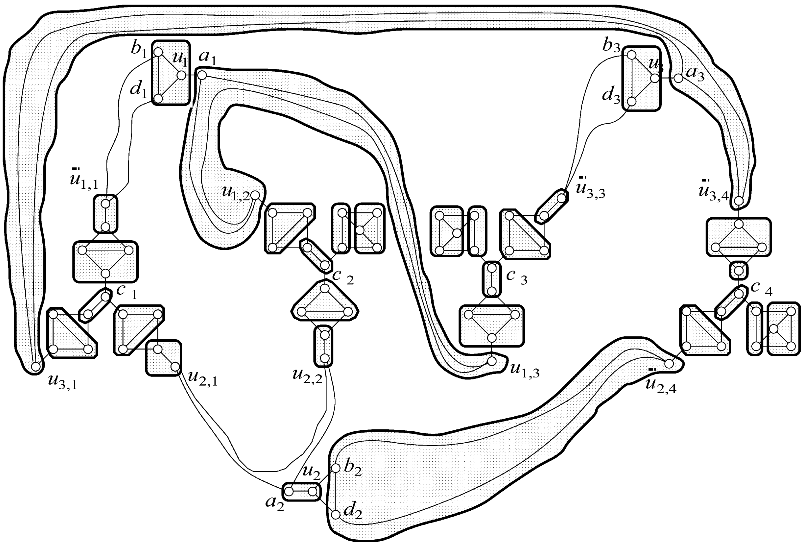

Fig. 3. Partition into cliques Z with $2 \times 3 + 6 \times 4 = 30$ cliques for graph G obtained from the satisfiable truth assignment $u_1 = \bar{u}_2 = u_3 = T$ for $I = (U, C)$.

Theorem 3 (*Fundamental Property of the Construction of $G = (V, E)$ from I*). *Let $I = (U, C)$ be an instance of* MAX3SAT$_{\underline{3}}(I)$*, with $|U| = n$ and $|C| = m$. A truth assignment for U with c satisfied clauses defines a partition into cliques Z of V of size $|Z| = 2n+7m-c$. Conversely, given a clique partition Z' of V, there exists a clique partition Z of V satisfying $|Z'| \geq |Z| = 2n+7m-c$ such that Z defines a truth assignment for U with c satisfied clauses, where each subgraph T_i requires two cliques in Z, each subgraph S_j corresponding to a satisfied clause requires* 6 *cliques in Z, and each subgraph S_j corresponding to a non-satisfied clause requires* 7 *cliques in Z.*

Proof. Given a truth assignment for U with c satisfied clauses, we shall define a suitable partition Z of cliques proving the first part of the theorem. First for each $i \in \{1, 2, \ldots n\}$ we define a partition for T_i. If u_i is true, then we add to Z the triangle of T_i induced by vertices u_i, b_i, d_i, and the triangle induced by vertices $a_i, u_{i,j}, u_{i,k}$. If u_i is false, then we add to Z the clique of T_i induced by vertices u_i, a_i, and the triangle induced by vertices $b_i, d_i, \bar{u}_{i,\ell}$. Note that for each clause c_j satisfied by some literal x_i, the vertex $x_{i,j}$ of S_j belongs to the partition defined so far. Hence, by Claim 1, if c_j is non-satisfied, then we partition $V(S_j)$ into additional 7 cliques; and by Claim 2, if c_j is satisfied, then we partition $V(S_j)$ into additional 6 cliques. Hence, Z is a partition of the set V, and its size is $|Z| = 2n + 6c + 7(m - c) = 2n + 7m - c$. See an example in Fig. 3.

Now we prove the second part of the theorem. Let Z' be a clique partition of V. We consider a first case where for each subgraph T_i there are two cliques of Z' containing all the vertices of T_i. In this case we define a truth assignment as follows: set variable u_i as true if and only if a_i and u_i do not belong to the same clique of Z'. Let c be the number of clauses satisfied by this truth assignment. Let c_j be a clause satisfied by literal u_i. Since vertices a_i and u_i do not belong to the same clique, the clique containing a_i can be extended in order to contain vertex $u_{i,j}$ of S_j without increasing the size of Z', and in this case by Claim 2, subgraph S_j requires at least 6 additional cliques in Z'. Let c_j be a clause satisfied by literal $\bar{u}_i$. Since vertices a_i and u_i belong to the same clique, the second clique of T_i, induced by vertices b_i and d_i, can be extended in order to contain vertex $\bar{u}_{i,j}$ of S_j without increasing the size of Z', and in this case by Claim 2, S_j requires at least 6 additional cliques in Z'. Let c_j be a non-satisfied clause. We consider two subcases. The first subcase occurs when c_j contains a literal u_i. Since a_i and u_i belong to the same clique, and the only vertex of the Truth Setting components adjacent to $u_{i,j}$ is a_i, we have that vertex $u_{i,j}$ does not belong to any clique containing vertices of the Truth Setting components. The second subcase occurs when c_j contains a literal $\bar{u}_i$. Since a_i and u_i do not belong to the same clique, the second clique of T_i is formed by vertices b_i, d_i and u_i, and the only vertices of the Truth Setting components adjacent to $\bar{u}_{i,j}$ are b_i and d_i; thus, vertex $\bar{u}_{i,j}$ does not belong to any clique containing vertices of the Truth Setting components. Therefore, if c_j is a non-satisfied clause, then S_j requires, by Claim 2, 7 additional cliques in Z'. Hence, we have two cliques for each variable, 6 cliques for each satisfied clause and 7 cliques for each non-satisfied clause, that altogether imply $|Z'| \geq |Z| = 2n + 7m - c$ cliques.

The second case occurs when there are three or four cliques of Z' containing the vertices of T_i. In this case, we show how to update Z' in order to define the partition Z, such that $|Z'| \geq |Z|$ and there are two cliques of Z containing all the vertices of T_i. Suppose that there is an integer $i \in \{1, 2, \ldots, n\}$ for which there are three or four cliques of Z' containing the vertices of T_i. Then, the cliques of Z' containing the vertices of T_i may contain one, two or three vertices in the subgraphs S_j, S_k and S_ℓ corresponding to the clauses where variable u_i occurs. Let u_i occurs positively in c_j and c_k, and negatively in c_ℓ. There are three subcases to consider:

(1) The cliques of Z' containing vertices of T_i contain exactly one vertex $x \in V(S_j) \cup V(S_k) \cup V(S_\ell)$. In this case, we update Z' by partitioning the set of vertices $T_i \cup \{x\}$ into two cliques: the clique K induced by x and the neighborhood of x contained in T_i, and the clique $T_i - K$.
(2) The cliques of Z' containing vertices of T_i contain exactly two vertices $x, y \in S_j \cup S_k \cup S_\ell$. In this case, we have two additional subcases:
  (a) $x = u_{i,j}$ and $y = u_{i,k}$; then update Z' by partitioning $V(T_i) \cup \{u_{i,j}, u_{i,k}\}$ into the two cliques induced by $u_{i,j}, u_{i,k}, a_i$ and $b_i, d_f i, u_i$.
  (b) $x = u_{i,j}$ and $y = \bar{u}_{i,\ell}$; then update Z' by partitioning $V(T_i) \cup \{u_{i,j}, \bar{u}_{i,\ell}\}$ into the three cliques induced by $u_{i,j}, a_i$, by $b_i, d_i, \bar{u}_{i,\ell}$, and by $u_{i,j}$.
(3) The cliques in Z' containing vertices of T_i contain $u_{i,j}, u_{i,k}, \bar{u}_{i,\ell}$. In this case we partition $V(T_i) \cup \{u_{i,j}, u_{i,k}, \bar{u}_{i,\ell}\}$ into the three cliques induced by $u_{i,j}, u_{i,k}, a_i$, by b_i, d_i, u_i, and by $\bar{u}_{i,j}$. □

Next we state two obvious consequences from Theorem 3.

Corollary 4. *If $I = (U, C)$ is an instance for* MAX3SAT$_{\underline{3}}$ *with $|U| = n$, $|C| = m$, and $f(I) = G$, then* $\mathrm{Opt}_{\mathrm{MINPIC}}(G) = 2n + 7m - \mathrm{Opt}_{\mathrm{MAX3SAT}_{\underline{3}}}(I)$.

Corollary 5. PIC *for planar cubic graphs is NP-complete.*

Let A and B be two optimization problems. We say that *AL-reduces* [15] to B if there are two polynomial-time algorithms f and g and positive constants α and β, such that for each instance I of A, algorithm f produces an instance $I' = f(I)$ of B such that the optima of I and I' satisfy $\mathrm{Opt}_B(I') \leq \alpha.\mathrm{Opt}_A(I)$. Given any feasible solution of I' with cost c', algorithm g produces a solution of I with cost c such that $|c - \mathrm{Opt}_A(I)| \leq \beta.|c' - \mathrm{Opt}_B(I')|$.

Papadimitriou and Yannakakis [14,15] defined the optimization Max SNP class and proved several complete problems under the L-reduction, in particular that MAX3SAT$_{\underline{3}}$ is Max SNP-complete. In Corollary 6 we establish bounds for the optimum value of an instance of MAX3SAT$_{\underline{3}}$ and use it to prove that MINPIC is Max SNP-hard for cubic graphs.

Corollary 6. *If $I = (U, C)$ is an instance of* MAX3SAT$_{\underline{3}}$ *with $|U| = n$ and $|C| = m$, then* $\lceil \frac{2n}{3} \rceil \leq \mathrm{Opt}_{\mathrm{MAX3SAT}_{\underline{3}}}(I) \leq m \leq \lceil \frac{3n}{2} \rceil$.

Proof. We notice first that $\mathrm{Opt}_{\mathrm{MAX3SAT}_{\underline{3}}}(I) \leq m$. In order to prove the first inequality, we exhibit a truth assignment for I with at least $\lceil \frac{2n}{3} \rceil$ satisfied clauses. For each variable $u_i \in U$, set $u_i = T$. Observe that in this truth assignment there are $2n$ true literals, and the minimum number of clauses fitting these literals is $\lceil \frac{2n}{3} \rceil$.

We remark that the instance used to prove in [15] that MAX3SAT$_{\underline{3}}$ is Max SNP-hard considers only clauses with sizes 2 and 3. Hence, the last inequality $m \leq \lceil \frac{3n}{2} \rceil$ follows from the fact that the greatest number of clauses in a MAX3SAT$_{\underline{3}}$ instance is reached when every clause of C has two literals. □

Theorem 7. MINPIC *for cubic graphs is Max SNP-hard.*

Proof. Given a truth assignment for U with c satisfied clauses, from Corollaries 4 and 6, $\mathrm{Opt}_{\mathrm{MINPIC}}(G) \leq 2n + 7m \leq 2n + 7\lceil \frac{3n}{2} \rceil \leq 2n + 14n = 16n = \frac{48}{2}\frac{2n}{3} \leq 24\lceil \frac{2n}{3} \rceil \leq 24\mathrm{Opt}_{\mathrm{MAX3SAT}_{\underline{3}}}(I)$. Then $\alpha = 24$ suffices. Let Z' be a feasible solution for G such that $|Z'| = c'$. By Theorem 3, we can define from Z', a truth assignment for U with c satisfied clauses such that $|Z'| \geq 2n + 7m - c$. Thus, $|\mathrm{Opt}_{\mathrm{MAX3SAT}_{\underline{3}}}(I) - c| = |2n + 7m - \mathrm{Opt}_{\mathrm{MAX3SAT}_{\underline{3}}}(I) - 2n - 7m + c| = |\mathrm{Opt}_{\mathrm{MINPIC}}(G) - (2n + 7m - c)|$. Therefore, $\beta = 1$ suffices. □

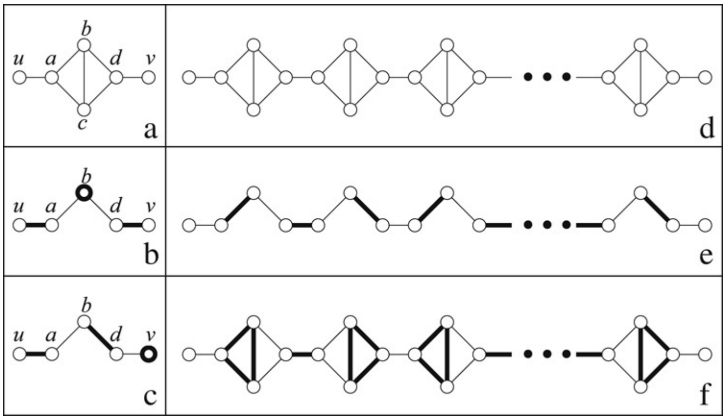

Fig. 4. (a) Diamond H and two vertices u, v in $G - H$; (b) an optimum partition into cliques of $G - c$; (c) another optimum partition of $G - c$; (d) chain of diamonds in G; (e) opening all diamonds in (d) and a corresponding optimum partition; (f) a partition with the same size as the one depicted in (e).

Given a pair of graphs S and G, we say that G is *S-free* if G does not contain S as a subgraph. The approximation algorithm presented in this paper takes as input a K_4-free maximum degree three connected graph $G = (V, E)$ and outputs a partition into cliques for V with size at most $\frac{5}{4}\text{Opt}_{\text{MINPIC}}(G)$. If G is a triangle-free graph, then the problem is easily solved by a polynomial maximum matching algorithm [5]. Moreover, we shall show that we need to consider only diamond-free graphs, since we will be able to transform a graph G into a diamond-free graph H, get a feasible $\frac{5}{4}$-approximation solution to H and obtain, in polynomial time a $\frac{5}{4}$-approximation solution to G. The way we get the $\frac{5}{4}$-approximation consists of taking the best of the two solutions for H: one using all triangles, and one using a maximum matching of H. We finally show that the value of the optimum solution is within $\frac{5}{4}$ of the best of these two solutions. For this purpose, we prove the lower bound $|T| + \frac{2}{3}|Z_T| - \frac{1}{3}|M_{Z_T}|$ for an optimum solution, where T is the collection of all triangles of $G = (V, E)$, $Z_T = V \setminus V(T)$, and M_{Z_T} is a maximum matching of the subgraph of G induced by Z_T.

Let $G = (V, E)$ be a graph with maximum degree three with a diamond H as a subgraph, where $V(H) = \{a, b, c, d\}$ and $E(H) = \{ab, ac, cb, bd, cd\}$. Let $u, v \in V$ be the vertices of $G - H$ adjacent to a and d, respectively, as shown in Fig. 4(a). Note that u and v may coincide. The operation *opening diamond* is the one which defines the graph $\text{open}(G) = G - c$ by the removal of the vertex c from G, as in Fig. 4(b).

Lemma 8. $\text{Opt}_{\text{PIC}}(\text{open}(G)) = \text{Opt}_{\text{PIC}}(G)$.

Proof. $\text{Opt}_{\text{PIC}}(G) \geq \text{Opt}_{\text{PIC}}(\text{open}(G))$, because $\text{open}(G)$ is a subgraph of G. We see that $\text{Opt}_{\text{PIC}}(G) \leq \text{Opt}_{\text{PIC}}(\text{open}(G))$, since we can add vertex c to the clique containing b in a partition into cliques for $\text{open}(G)$ in order to obtain a partition into cliques with same size for G. □

Lemma 8 allows us to design the following strategy for the algorithm: first, open all the diamonds of G obtaining the graph G' (see Fig. 4(d) and (e)); next, find a feasible partition into cliques β for G' (see Fig. 4(e)); finally, for each diamond H of G we update β by adding c to the same clique containing d in order to define a feasible partition into cliques α for G such that $|\beta| = |\alpha|$ (see Fig. 4(f)). By Lemma 8, $\text{Opt}_{\text{PIC}}(G) = \text{Opt}_{\text{PIC}}(G')$. Thus, the existence of an approximation ratio for β implies the same ratio for α. From now on, we assume diamond-free input instances for the algorithm.

Let $G = (V, E)$ be a diamond-free graph with maximum degree three. Thus, no two triangles of G intersect.

Let $t = (a, b, c)$ be a triangle of G. The *neighborhood* $N(t)$ of t is defined as $N(t) = \{v \in V - t \mid v \text{ is adjacent to some vertex of } t\}$. Given a set T' of triangles of G, we denote $N(T') = \bigcup_{t \in T'} N(t)$. Given an optimum solution α for PIC in G, we define three transformations.

(1) *Transformation* 1 — Let u, v, x be the vertices of $N(t)$, respectively adjacent to a, b, c, as shown on the left in Fig. 5. These vertices are distinct, since G is diamond-free. If t has a vertex, say a (on the left in Fig. 5(a)) as a

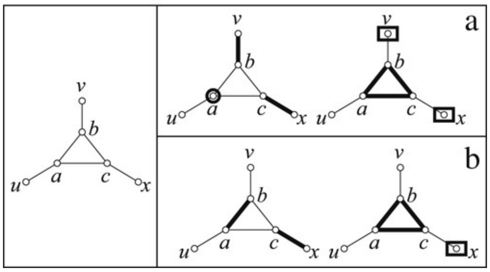

Fig. 5. (On the left) triangle t on vertices a, b and c plus the neighborhood $N(t) = \{u, v, x\}$. (On the right) (a) *Transformation* 1 — Triangle t with a vertex in an optimum solution α, and a new optimum solution with t as one of the cliques. (On the right) (b) *Transformation* 2 — Triangle t with an edge in an optimum solution α, and a new optimum solution with t as one of the cliques.

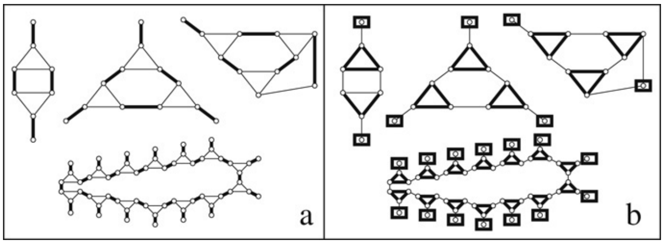

Fig. 6. *Transformation* 3 — (a) Four examples of subgraphs of G containing even cycles and corresponding partitions into cliques, and (b) other partitions obtained from the previous ones as explained in Transformation 3.

clique of α, then remove a from α, delete b and c from the cliques of α containing them and add t to α, as shown on the right in Fig. 5(a).

(2) *Transformation* 2 — Again, let u, v, x be the vertices of $N(t)$, respectively adjacent to a, b, c, as shown on the left in Fig. 5. If t has an edge, say ab, as a clique of α (on the left in Fig. 5(b)), then remove ab from α, delete c from the clique of α containing c, and add t to α, as shown on the right in Fig. 5(b).

(3) *Transformation* 3 — Suppose we have an even cycle C in G on vertices $v_1, v_2, \ldots, v_{2k}$ and on edges $v_1v_2, v_2v_3, \ldots, v_{2k-1}v_{2k}, v_{2k}v_1$, such that the edge $v_{2i+1}v_{2i+2}$ belongs to α for all $i \in \{0, 1, \ldots, k-1\}$, and for each remaining edge e of C there is a triangle t_e of G containing e. Add to α the k triangles containing the k edges of C out of α, and remove from α the k edges of C in α, as shown in Fig. 6.

Transformations 1–3 do not increase the number of cliques in α. We now consider an optimum solution α where none of the Transformations 1–3 can be applied. Let T be the collection of all the triangles of G, $T' = T \cap \alpha$, and $T'' = T - T'$. Since we cannot perform Transformations 1 and 2, every edge overlapping some t of T'' must be in α. Further, as we cannot perform Transformation 3, there is no cycle C of G with a sequence of alternating edges belonging to α such that every edge of C out of this sequence is contained in some t of T''. Hence, if we contract each triangle in T'' to a vertex, then the edges of α intersecting T'' form a forest. It follows that a component of this forest containing ℓ vertices of $V(G) \setminus V(T'')$ contains $\ell - 2$ vertices (contracted) of T'' and $2\ell - 3$ edges of α.

In order to prove the next results, consider $G = (V, E)$ a maximum degree three diamond-free graph where none of the Transformations 1–3 can be applied. Given a collection of triangles T' of G, we denote $Z_{T'} = V(G) \setminus V(T')$. If Z is a subset of $V(G)$, we denote by M_Z a maximum matching over Z.

Theorem 9. *Let T be the collection of all triangles of G, α be an optimum solution of* MINPIC *for G, and T^* be the collection of triangles in α. Then, $|\alpha| = |T^*| + |Z_{T^*}| - |M_{Z_{T^*}}| \geq |T| + \frac{2}{3}|Z_T| - \frac{1}{3}|M_{Z_T}|$.*

Proof. First of all we prove that $|\alpha| = |T^*| + |Z_{T^*}| - |M_{Z_{T^*}}|$. Note that the size of α must be $|T^*|$ plus $|M_{Z_{T^*}}|$, by the definition of a maximum matching, plus the vertices of Z_{T^*} out of $M_{Z_{T^*}}$. Hence, $|\alpha| = |T^*| + |M_{Z_{T^*}}| + |Z_{T^*}| - 2|M_{Z_{T^*}}| = |T^*| + |Z_{T^*}| - |M_{Z_{T^*}}|$. Let $(Z^N, \overline{Z^N})$ be a partition of Z_T, where $Z^N = \{v \in Z_T \mid v \text{ is adjacent to some vertex of a triangle of } T - T^*\}$. By setting $\ell = |Z^N|$, we have by the previous remarks that $\ell - 2 = |T| - |T^*|$ and $2\ell - 3 = \ell - 2 + \ell - 1 = (|T| - |T^*|) + |Z^N| - 1$, which is the number of cliques of α in $G[(T - T^*) \cup N(T - T^*)]$.

Hence, the size of α is

$$|T^*| + |Z_{T^*}| - |M_{Z_{T^*}}| = |T^*| + \big(|T| - |T^*|\big) + |Z^N| - 1 + |\overline{Z^N}| - |M_{\overline{Z^N}}|, \tag{1}$$

since $|\overline{Z^N}| - |M_{\overline{Z^N}}|$ is the best partition into cliques of the subgraph induced by $\overline{Z^N}$. Note also that $|T| + \frac{2}{3}|Z_T| - \frac{1}{3}|M_{Z_T}| = |T| + \frac{2}{3}|Z^N| + \frac{2}{3}|\overline{Z^N}| - \frac{1}{3}|M_{Z_T}|$.

From (1) and the fact that $|M_{Z_T}| \geq |M_{\overline{Z^N}}|$, we have that $|T^*| + (|T| - |T^*|) + |Z^N| - 1 + |\overline{Z^N}| - |M_{Z_T}| \leq |\alpha|$. Hence, in order to prove that $|\alpha| \geq |T| + \frac{2}{3}|Z_T| - \frac{1}{3}|M_{Z_T}|$, it is enough to prove that: $|T| + \frac{2}{3}|Z^N| + \frac{2}{3}|\overline{Z^N}| - \frac{1}{3}|M_{Z_T}| \leq |T^*| + (|T| - |T^*|) + |Z^N| - 1 + |\overline{Z^N}| - |M_{Z_T}|$. Note that the last inequality occurs if and only if $1 \leq \frac{1}{3}\left(|Z^N| + |\overline{Z^N}| - 2|M_{Z_T}|\right) = \frac{1}{3}\left(|Z_T| - 2|M_{Z_T}|\right)$. Observe now that $|Z_T| - 2|M_{Z_T}|$ is the number of vertices out of a maximum matching in Z_T. Hence, the last inequality do not hold only when $i = |Z_T| - 2|M_{Z_T}| = 0$, 1, or 2. We show next that if $i = 0$, 1, or 2, then the solution with all triangles plus a maximum matching plus the remaining vertices is optimum.

Let β be such a solution with all triangles plus a maximum matching plus the remaining vertices, and $|\beta| = |T| + |Z_T| - |M_{Z_T}| = |T| + |M_{Z_T}| + i$ its size. If $i = 0$, then $|\alpha| = |\beta|$. Assume that γ is a partition into cliques for $G = (V, E)$ obtained from β by the removal of $r \geq 1$ triangles and p cliques of size two, such that $|\gamma| < |\beta|$. Consider now the case $i = 1$. Then all the vertices of G are partitioned into triangles and edges of γ, and so $\frac{3r+2p+1}{2} < r + p + 1$. Hence, $r < 1$, a contradiction. Finally, consider the case $i = 2$. We have two subcases: either there is one clique with one vertex in γ, or there is none. If there is one clique with one vertex in γ, then $\frac{3r+2p+1}{2} < r + p + 1$, and we have the same contradiction of the previous case. If there is no clique with one vertex in γ, then $\frac{3r+2p+2}{2} < r + p + 2$. Hence, $r < 2$. Then, $r = 1$, and therefore the removal of just one triangle defines a matching plus five vertices (the vertices of the removed triangle plus the two isolated vertices of β). These five vertices requires the additional number of three cliques in γ corresponding to the removed triangle and the two isolated vertices of β. Hence, $|\gamma| = |\beta|$. Thus, in the case that $i = 0$, 1 or 2, $|\alpha| \geq |T| + |M_{Z_T}| + |Z_T| - 2|M_{Z_T}| = |T| + |Z_T| - |M_{Z_T}| = |T| + \frac{2}{3}|Z_T| - \frac{|M_{Z_T}|}{3} + \frac{1}{3}\left(|Z_T| - 2|M_{Z_T}|\right) \geq |T| + \frac{2}{3}|Z_T| - \frac{|M_{Z_T}|}{3}$. □

Algorithm A
Input: Graph $G = (V, E)$ with maximum degree three.
Output: Partition into Cliques $A(G)$ for the set V of vertices of G.

(1) Update G by opening all diamonds of G.
(2) Set A_1 as the partition defined by the collection T of triangles of G plus the edges of M_{Z_T} plus the remaining $|Z_T| - 2|M_{Z_T}|$ vertices. Note that $|A_1| = |T| + |M_{Z_T}| + |Z_T| - 2|M_{Z_T}| = |T| + |Z_T| - |M_{Z_T}|$.
(3) Set A_2 as the partition defined by the edges of M_V plus the remaining $|V(G)| - 2|M_V|$ vertices. Note that $|A_2| = |M_V| + |V(G)| - 2|M_V| = |V(G)| - |M_V|$.
(4) Update A_1 and A_2 with the suitable triangles of the diamonds of G, as in Fig. 4(e) and (f).
(5) Set $A(G) = A_i$, where $|A_i| = \min\{|A_1|, |A_2|\}$.

Let A and B be two collections of cliques of G, and k a positive real number. We say that A is *within* k of B if $||A| - |B|| \leq k$.

Theorem 10. *The performance ratio of the algorithm A is at most* $\frac{5}{4}$.

Proof. In order to avoid overloading notation we write for short $Z = Z_T$. We consider two cases: $|T| \geq \frac{2}{3}|Z| - \frac{1}{3}|M_Z|$, or $|T| < \frac{2}{3}|Z| - \frac{1}{3}|M_Z|$.

(1) If $|T| \geq \frac{2}{3}|Z| - \frac{1}{3}|M_Z|$, then the performance ratio R_A of A satisfies
$R_A = \frac{|A(G)|}{\mathrm{Opt}_{\mathrm{PIC}}(G)} \leq \frac{|T|+|Z|-|M_Z|}{|T|+\frac{2}{3}|Z|-\frac{1}{3}|M_Z|} \leq \frac{5}{4}$ if and only if $12|T| + 12|Z| - 12|M_Z| \leq 15|T| + 10|Z| - 5|M_Z|$, which holds if and only if $|T| \geq \frac{2}{3}|Z| - \frac{7}{3}|M_Z|$. But $|T| \geq \frac{2}{3}|Z| - \frac{1}{3}|M_Z| \geq \frac{2}{3}|Z| - \frac{1}{3}|M_Z| - \frac{6}{3}|M_Z| = \frac{2}{3}|Z| - \frac{7}{3}|M_Z|$.
(2) If $|T| < \frac{2}{3}|Z| - \frac{1}{3}|M_Z|$, then we consider two additional cases, $|T^*| \leq \frac{|T|}{2}$ and $|T^*| > \frac{|T|}{2}$. In both cases, we prove that α is within $\frac{|T|}{2}$ of $A(G)$.

(a) If $|T^*| \leq \frac{|T|}{2}$, we prove that α is within $\frac{|T|}{2}$ of $|A_2| = n - |M_V|$. For this purpose we define a new partition β by removing all $|T^*|$ triangles from α and by adding to α for each triangle t of T^*, one edge and one vertex. Since β does not contain any triangle we have that $|A_2| \leq |\beta| = |\alpha| + |T^*| = (|T^*| + |Z_{T^*}| - |M_{Z_{T^*}}|) + |T^*| \leq (|T^*| + |Z_{T^*}| - |M_{Z_{T^*}}|) + \frac{|T|}{2}$.

(b) If $|T^*| > \frac{|T|}{2}$, then we prove that α is within $\frac{|T|}{2}$ of $|A_1| = |T| + |Z| - |M_Z|$. For this purpose we define a clique covering (not necessarily a partition into cliques) β by adding to α all triangles of $T - T^*$. Note that the number of elements in β is $|\beta| = (|T^*| + |Z_{T^*}| - |M_{Z_{T^*}}|) + |T - T^*|$. Observe that β has as many triangles as A_1, and since Z is partitioned into a maximum matching plus the remaining vertices, we have that $|A_1| = |T| + |Z| - |M_Z| \leq (|T^*| + |Z_{T^*}| - |M_{Z_{T^*}}|) + |T - T^*| < (|T^*| + |Z_{T^*}| - |M_{Z_{T^*}}|) + \frac{|T|}{2}$.

Hence, in either case, α is within $\frac{|T|}{2}$ of $A(G)$. Therefore, $R_A = \frac{|A(G)|}{\mathrm{Opt}_{\mathrm{PIC}}(G)} \leq \frac{\frac{|T|}{2} + \mathrm{Opt}_{\mathrm{PIC}}(G)}{\mathrm{Opt}_{\mathrm{PIC}}(G)} < \frac{5}{4}$ if and only if $|T| < \frac{\mathrm{Opt}_{\mathrm{PIC}}(G)}{2}$. But, $\mathrm{Opt}_{\mathrm{PIC}}(G) \geq |T| + \frac{2|Z|}{3} - \frac{|M_Z|}{3} > |T| + |T| = 2|T|$. □

3. Conclusions

We conjecture that MINPIC has a polynomial-time approximation algorithm with a fixed ratio for an input graph having a fixed maximum degree. Another question left open is whether MINPIC admits a Polynomial Time Approximation Scheme (PTAS) for planar cubic graphs. A positive evidence for this claim is that there is a PTAS [12] for MAXP3SAT, which is the decision version that we use to prove the NP-completeness of PIC for planar cubic graphs.

Acknowledgements

We thank the referees for their careful reading and valuable suggestions which helped to improve the results and the presentation of the paper.

References

[1] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Sccamela, M. Protasi, Complexity and Approximation (Combinatorial Optimization Problems and their Approximability Properties), Springer, Berlin, 1999.
[2] M. Bellare, O. Goldreich, M. Sudan, Free bits, PCPs and non-approximability — towards tight results, SIAM J. Comput. 27 (1998) 804–915.
[3] P. Berman, M. Karpinski, A.D. Scott, Approximation hardness and satisfiability of bounded occurrence instances of SAT, in: Electronic Colloquium on Computational Complexity, ECCC, 2003.
[4] R.L. Brooks, On colouring nodes of a network, Proc. Cambridge Philos. Soc. 37 (1941) 194–197.
[5] J. Edmonds, Paths, trees and flowers, Canad. J. Math. 17 (1965) 449–467.
[6] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W. H. Freeman & Company, 1979.
[7] M.R. Garey, D.S. Johnson, L. Stockmeyer, Some simplified NP-complete graph problems, Theoret. Comput. Sci. 1 (1976) 237–267.
[8] M.M. Halldórsson, A still better performance guarantee for approximate graph coloring, Inform. Process. Lett. 45 (1993) 19–23.
[9] M.M. Halldórsson, Approximating k-set cover and complementary graph coloring. in: Proc. 5th Conference on Integer Programming and Combinatorial Optimization, IPCO'1996, 1996, pp. 118–131.
[10] H.B. Hunt III, M.V. Marathe, V. Radhakrishnan, R.E. Stearns, The complexity of planar counting problems, SIAM J. Comput. 27 (4) (1998) 1142–1167.
[11] R.M. Karp, Reducibility among combinatorial problems, in: R.E. Miller, J.W. Thatcher (Eds.), Complexity of Computer Computations, Plenum Press, 1972, pp. 85–103.
[12] S. Khanna, Motwani, Towards a syntactic characterization of PTAS, in: 28th ACM Symp. on Theory of Computing, STOC'96, 1996, pp. 329–337.
[13] D. Lichtenstein, Planar formulae and their uses, SIAM J. Comput. 43 (1982) 329–393.
[14] C.H. Papadimitriou, M. Yannakakis, Optimization, approximation, and complexity classes, J. Comput. System Sci. 43 (1991) 425–440.
[15] C.H. Papadimitriou, Computational Complexity, Addison-Wesley, 1994.