Approximation algorithms and hardness results for the clique packing problem

F. Chataigner [1] *   G. Manić [2] †   Y.Wakabayashi [1] ‡   R. Yuster [3] §

[1] Instituto de Matemática e Estatística – Universidade de São Paulo, SP, Brazil
[2] Instituto de Computação – Universidade Estadual de Campinas, SP, Brazil
[3] Department of Mathematics, University of Haifa, Israel

October, 2007

Abstract

For a fixed family $\mathcal{F}$ of graphs, an $\mathcal{F}$-packing in a graph G is a set of pairwise vertex-disjoint subgraphs of G, each isomorphic to an element of $\mathcal{F}$. Finding an $\mathcal{F}$-packing that maximizes the number of covered edges is a natural generalization of the maximum matching problem, which is just $\mathcal{F} = \{K_2\}$. In this paper we provide new approximation algorithms and hardness results for the $\mathcal{K}_r$-packing problem where $\mathcal{K}_r = \{K_2, K_3, \ldots, K_r\}$.

We show that already for $r = 3$ the $\mathcal{K}_r$-packing problem is APX-complete, and, in fact, we show that it remains so even for graphs with maximum degree 4. On the positive side, we give an approximation algorithm with approximation ratio at most 2 for *every* fixed r. For $r = 3, 4, 5$ we obtain better approximations. For $r = 3$ we obtain a *simple* $3/2$-approximation, achieving a known ratio that follows from a more involved algorithm of Halldórsson. For $r = 4$, we obtain a $(3/2 + \epsilon)$-approximation, and for $r = 5$ we obtain a $(25/14 + \epsilon)$-approximation.

Keywords: approximation algorithms, APX-hardness, clique, packing, triangle

1 Introduction

Let $\mathcal{F}$ be a fixed family of graphs. An *$\mathcal{F}$-packing* in a graph G is a set of pairwise vertex-disjoint subgraphs of G, each isomorphic to an element of $\mathcal{F}$. We say that an $\mathcal{F}$-packing *covers* an edge (resp. vertex) of G if one of the subgraphs of the packing contains that edge (resp. vertex). In this paper the *$\mathcal{F}$-packing problem* is the problem of finding the maximum number of *edges* that can be covered by an $\mathcal{F}$-packing. When $\mathcal{F} = \{K_2\}$, this simply corresponds to the maximum matching problem. Apart from its theoretical interest, this problem is also important from a practical point of view, as it arises naturally in applications such as scheduling.

*frederic@ime.usp.br; supported by FAPESP, Proc. 05/53840-0.
†gocam@ime.usp.br; supported by FAPESP, Proc. 2006/01817-7.
‡yw@ime.usp.br; partially supported by CNPq (Proc. 490333/04-4, 308138/04-0) and PRONEX–FAPESP/CNPq (Proc. 2003/09925-5).
§raphy@math.haifa.ac.il

Another related problem is that of finding an $\mathcal{F}$-packing in a graph G that covers the maximum number of *vertices*. To avoid confusion, we refer to this problem as the *$\mathcal{F}_V$-packing problem.* This problem is NP-hard, even when $\mathcal{F}$ consists of a single graph that has a component with at least three vertices [5]; and also when $\mathcal{F}$ contains only complete graphs with at least three vertices [6]. On the other hand, the $\mathcal{F}_V$-packing problem is polynomially solvable for some non-trivial classes of families $\mathcal{F}$, and many important results in matching theory can be generalized to those cases. For example, when $\mathcal{F} = \{K_2, \ldots, K_r\}$, $r > 2$, Hell and Kirkpatrick [6] showed that this problem is in P.

Let $\mathcal{K}_r = \{K_2, \ldots, K_r\}$. In contrast with the above result of Hell and Kirkpatrick, we show, in Section 3, that the $\mathcal{K}_r$-packing problem is APX-complete already for $r = 3$, and, in fact, already for graphs with maximum degree 4.

On the positive side, we show in Section 2 that a simple greedy algorithm yields a 2-approximation for the $\mathcal{K}_r$-packing problem. A modified greedy algorithm, which is based on application of the local search method of Hurkens and Schrijver [7] yields better approximation ratios for $r = 4, 5$. The analysis of these ratios is also somewhat more complicated than the analysis of the simple greedy algorithm. In particular, for $r = 4$ we obtain a $(3/2 + \epsilon)$-approximation and for $r = 5$ we obtain a $(25/14 + \epsilon)$-approximation.

In Section 4 we specifically address the $\mathcal{K}_3$-packing problem. We show, in fact, that a tighter analysis of the simple greedy algorithm yields a 3/2-approximation for it. More generally, we show that there is a $(1 + \frac{1}{3}\rho)$-approximation algorithm for the $\mathcal{K}_3$-packing problem, whenever there is a ρ-approximation algorithm for the triangle packing problem. In particular, for the class of graphs with maximum degree 4, using a result of [8] for the triangle packing problem, we derive a 1.4-approximation algorithm for the $\mathcal{K}_3$-packing problem for this class of graphs.

An extended abstract mentioning the results of this paper has appeared in the Proceedings of Eurocomb 2007 [3].

1.1 Basic definitions and notation

All graphs considered here are simple. If G is a graph, then V_G (resp. E_G) denotes its vertex (resp. edge) set. The number of vertices of G is denoted by n_G, and the maximum degree by $\Delta(G)$.

Let $\mathcal{F}$ be a fixed family of graphs. We recall that in the $\mathcal{F}$-packing problems to be investigated in this paper we are interested in maximizing the number of *edges* that are covered. The set $\{K_2, \ldots, K_r\}$ is abbreviated as $\mathcal{K}_r$. A complete graph of order k is called a *k-clique.* A 3-clique is called a *triangle.* If $\mathcal{A}$ is an $\mathcal{F}$-packing of a graph G, then the *value* of $\mathcal{A}$, denoted $\text{val}_{\mathcal{F}}(G, \mathcal{A})$ (or simply $\text{val}(\mathcal{A})$), is the number of edges of G that it covers. We denote by $\mathcal{P}_{\mathcal{A}}$ (resp. $\mathcal{Q}_{\mathcal{A}}$, $\mathcal{T}_{\mathcal{A}}$, $\mathcal{E}_{\mathcal{A}}$) the collection of all 5-cliques (resp. 4-cliques, triangles, edges) in $\mathcal{A}$. Furthermore, we denote by $G[\mathcal{A}]$ the subgraph of G induced by the set of edges in $\mathcal{A}$. If G is an instance of the $\mathcal{F}$-packing problem, then $\text{opt}_{\mathcal{F}}(G)$ denotes the value of an optimal solution for G. We call a graph *irredundant* if each of its edges belongs to some triangle.

In this paper we refer to a heuristic of Hurkens and Schrijver [7], denoted as $\text{HS}(k, t)$, that finds a maximum set of vertex-disjoint k-cliques in a graph G. It is a local search greedy heuristic that, starting with any collection of k-cliques, while possible, it replaces at most $p-1$ k-cliques in the current collection with a set of $p \leq t$ disjoint k-cliques that are not in the current collection, and updates the current collection. The parameter t is an integer. Its approximation ratio is $k/2 + \varepsilon$, where ε depends on t.

Given a parameter $\rho \geq 1$, a *ρ-approximation algorithm* for a maximization problem Π is a polynomial-time algorithm that, for any instance I of Π produces a solution S whose value, $\mathrm{val}_\Pi(I, S)$, is at least $\frac{1}{\rho}\mathrm{opt}_\Pi(I)$, where $\mathrm{opt}_\Pi(I)$ is the value of an optimal solution for I (we also say that ρ is the *approximation ratio*). If such an algorithm exists, we say that Π belongs to APX. Let Π_1 and Π_2 be optimization problems. An L-*reduction* from Π_1 to Π_2 consists of a pair of polynomial-time computable functions (f, g) such that, for two fixed positive constants α and β the following holds:

(C1) For every input instance I_1 of Π_1, $f(I_1)$ is an instance of Π_2, and $|\mathrm{opt}_{\Pi_2}(f(I_1))| \leq \alpha|\mathrm{opt}_{\Pi_1}(I_1)|$.

(C2) Given an instance I_1 of Π_1, and any feasible solution $\mathcal{A}$ for $f(I_1)$, we have that $g(I_1, \mathcal{A})$ is a feasible solution for the instance I_1 of Π_1, and $|\mathrm{opt}_{\Pi_1}(I_1) - \mathrm{val}_{\Pi_1}(I_1, g(I_1, \mathcal{A}))| \leq \beta|\mathrm{opt}_{\Pi_2}(f(I_1)) - \mathrm{val}_{\Pi_2}(f(I_1), \mathcal{A})|$.

We denote by $\Pi_1 \leq_L \Pi_2$ the existence of an L-reduction from Π_1 to Π_2. If $\Pi_1 \leq_L \Pi_2$ and $\Pi_2 \in$ APX, then $\Pi_1 \in$ APX. A problem Π is APX-*hard* if, for every $\Pi' \in$ APX, we have $\Pi' \leq_L \Pi$. If an APX-*hard* problem belongs to the class APX, then it is APX-complete. It is known that an APX-hard problem does not admit a PTAS, unless P = NP [1].

One of the reductions we show in Section 3 consider the following restricted version of the MAX SAT problem, denoted here simply as SAT, known to be APX-complete [1]. Given a collection of disjunctive clauses $C = \{c_1, c_2, \ldots, c_l\}$ over a set $X = \{x_1, x_2, \ldots, x_p\}$ of variables, such that each clause has at most 2 literals, and each variable appears in at most 3 of the clauses (counting both positive and negated occurrences), find a truth assignment for the variables of X that satisfies as many clauses as possible.

2 Approximation algorithm for the $\mathcal{K}_r$-packing problem

All algorithms we describe in this paper have a common structure. This common structure will be presented in a form of a generic algorithm, called here BASIC$_r$. To distinguish the different algorithms we can derive from this basic algorithm, we assume this algorithm calls a generic PROCEDURE $\mathcal{P}_q$ that outputs a $\{K_q\}$-packing of a given input graph G. The different algorithms are then obtained by substituting $\mathcal{P}_q$ by specific algorithms.

Algorithm BASIC$_r$

Input: A graph G.

Subroutine: PROCEDURE $\mathcal{P}_q$ that outputs a $\{K_q\}$-packing of a given input graph

Output: A $\{K_2, \ldots, K_r\}$-packing of G.

for $q = r$ **downto** 3 **do**
  $F_q \leftarrow$ a $\{K_q\}$-packing of G output by the PROCEDURE $\mathcal{P}_q$
  $G \leftarrow G - F_q$
$F_2 \leftarrow$ a maximum matching in G
return $F_2 \cup \ldots \cup F_q$

2.1 A simple greedy algorithm

Denote by $\mathcal{G}_r$ the simple Greedy Algorithm that consists of the algorithm $\textsc{Basic}_r$ in which the Procedure $\mathcal{P}_q$ (called at line 2) is an algorithm that simply selects a *maximal* set of vertex-disjoint q-cliques in a graph. We say $\mathcal{G}_r$ is a greedy algorithm because it selects first the larger cliques.

The next lemma is the key to obtain the approximation ratio of $\mathcal{G}_r$.

Lemma 2.1 *Let G be a graph, $r \geq 2$ an integer, and $\mathcal{A}$ a solution returned by the algorithm $\mathcal{G}_r$ applied to G. If C is a q-clique in G, where $2 \leq q \leq r$, then $\sum_{v \in V_C} d_{G[\mathcal{A}]}(v) \geq \frac{1}{2}q(q-1)$.*

Proof. The proof is by induction on r. For $r = 2$, it suffices to note that if C is a 2-clique in G then at least one of its vertices intersects an edge of $\mathcal{A}$ (a maximal matching returned by the algorithm). Thus, $\sum_{v \in V_C} d_{G[\mathcal{A}]}(v) \geq 1$ and the lemma holds.

Suppose now that $r > 2$. Let C be a q-clique of G, $2 \leq q \leq r$, and let l be the number of vertices in the intersection of C and F_r. Set $F' := \cup_{i=2}^{r-1} F_i$. Then, we have that

$$\sum_{v \in V_C} d_{G[\mathcal{A}]}(v) = \sum_{v \in V_{C \cap F_r}} d_{G[\mathcal{A}]}(v) + \sum_{v \in V_{C \cap F'}} d_{G[\mathcal{A}]}(v) = l(r-1) + \sum_{v \in V_{C - F_r}} d_{G[F']}(v).$$

Note that $C - F_r$ is isomorphic to K_{q-l}. When $q < r$ we have that $q - l < r$. When $q = r$, since $\mathcal{P}_r$ is the algorithm that simply selects a maximal set of vertex-disjoint r-cliques, we have that $l \geq 1$, and again $q - l < r$. Thus, we can apply the induction hypothesis on the last term of the equation, obtaining

$$\sum_{v \in V_C} d_{G[\mathcal{A}]}(v) \geq l(r-1) + \frac{1}{2}(q-l)(q-l-1) = \frac{1}{2}(l^2 + (2r - 2q - 1)l + q(q-1)). \tag{1}$$

If $q < r$, then $2r - 2q - 1 > 0$, so the minimum for the right-hand side of (1) is reached at $l = 0$. In that case we thus have $\sum_{v \in V_C} d_{G[\mathcal{A}]}(v) \geq \frac{1}{2}q(q-1)$. If, however, $q = r$, then the minimum for the right-hand side of (1) is reached at $l = \frac{1}{2}$. Since $\mathcal{P}_r$ selects a maximal set of vertex-disjoint r-cliques, we have that $l = 0$ is not possible. Thus, the minimum for the right-hand side of (1) is reached at $l = 1$ and is $\frac{1}{2}q(q-1)$. So, the proof is now complete. ■

Theorem 2.2 *For $r \geq 2$, the algorithm $\mathcal{G}_r$ is a 2-approximation algorithm for the $\mathcal{K}_r$-packing problem.*

Proof. Let $\mathcal{A}$ be a solution returned by the algorithm $\mathcal{G}_r$ applied to a graph G. Consider an optimal $\mathcal{K}_r$-packing $\mathcal{O}$ in G. Applying Lemma 2.1 to each clique C of $\mathcal{O}$, we get

$$2\,\mathrm{val}(\mathcal{A}) = \sum_{v \in V_G} d_{G[\mathcal{A}]}(v) \geq \sum_{C \in \mathcal{O}} \sum_{v \in V_C} d_{G[\mathcal{A}]}(v) \geq \sum_{C \in \mathcal{O}} \frac{1}{2}|V_C|(|V_C| - 1) = \sum_{C \in \mathcal{O}} |E_C| = \mathrm{val}(\mathcal{O}).$$

The first inequality follows from the fact that in the first sum we consider the degrees in $G[\mathcal{A}]$ of all vertices of G, and in the second sum we consider the degrees in $G[\mathcal{A}]$ of those vertices in

G that belong to the cliques of $\mathcal{O}$ (we may have vertices in $G[\mathcal{A}]$ that do not belong to $\mathcal{O}$). The second inequality follows from Lemma 2.1.

■

Remark 1. In the proof of Lemma 2.1 we did not use the fact that F_2 is a maximum matching (see step 4 of the algorithm BASIC_r). That is, we may substitute step 4 by "$F_2 \leftarrow$ a maximal matching in G", and obtain the same result. In other words, if we consider that $\mathcal{G}_r$ (the Greedy Algorithm) simply uses the PROCEDURE $\mathcal{P}_q$ for $q = r, \ldots, 2$, the statement of Theorem 2.2 holds.

Remark 2. We note that the upper bound 2 for the approximation ratio of algorithm $\mathcal{G}_r$ is not tight. In Section 4 we show that the algorithm $\mathcal{G}_3$ has, in fact, approximation ratio 3/2. The analysis is somewhat more delicate, however.

2.2 A modified greedy algorithm based on local search

Denote by $\mathcal{B}_r$ the algorithm BASIC_r in which the PROCEDURE $\mathcal{P}_q$ is the heuristic $\text{HS}(q,t)$, when $q = r$ (see Section 1.1), and for $2 < q < r$, $\mathcal{P}_q$ is the algorithm that simply selects a maximal set of vertex-disjoint q-cliques.

Theorem 2.3 *The algorithm $\mathcal{B}_4$ is a $(3/2+\varepsilon)$-approximation algorithm for the $\mathcal{K}_4$-packing problem.*

Proof. Let $\mathcal{O}$ be an optimal solution and $\mathcal{B}$ be the solution returned by the algorithm $\mathcal{B}_4$. Thus, $\text{val}(\mathcal{O}) = 6|\mathcal{Q}_\mathcal{O}| + 3|\mathcal{T}_\mathcal{O}| + |\mathcal{E}_\mathcal{O}|$ and $\text{val}(\mathcal{B}) = 6|\mathcal{Q}_\mathcal{B}| + 3|\mathcal{T}_\mathcal{B}| + |\mathcal{E}_\mathcal{B}|$.

Let q_i, $0 \le i \le 4$, be the number of 4-cliques of $\mathcal{Q}_\mathcal{O}$ that intersect precisely i vertices of $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$. Let t_i, $0 \le i \le 3$, be the number of triangles of $\mathcal{T}_\mathcal{O}$ that intersect precisely i vertices of $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$. Note that since $\mathcal{P}_3$ selects a maximal set of vertex-disjoint 3-cliques, we have that $t_0 = 0$. Furthermore, since $\text{HS}(4,t)$ returns a maximal collection of 4-cliques, we have that $q_0 = 0$. Suppose now that $q_1 > 0$. Then, there is a 4-clique, say D, in $\mathcal{Q}_\mathcal{O}$ that intersects precisely one vertex of $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$. The three other vertices of D would form a triangle that does not intersect $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$, contradicting the fact that $\mathcal{P}_3$ selects a maximal set of vertex-disjoint 3-cliques. Thus, $q_1 = 0$.

Observe that the number of vertices of $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$ covered by $\mathcal{Q}_\mathcal{O} \cup \mathcal{T}_\mathcal{O}$ is $2q_2 + 3q_3 + 4q_4 + t_1 + 2t_2 + 3t_3$. Thus, the number of vertices of $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$ not covered by $\mathcal{Q}_\mathcal{O} \cup \mathcal{T}_\mathcal{O}$ is $w := 4|\mathcal{Q}_\mathcal{B}| + 3|\mathcal{T}_\mathcal{B}| - (2q_2 + 3q_3 + 4q_4 + t_1 + 2t_2 + 3t_3)$. Hence, the number of edges of $\mathcal{E}_\mathcal{O}$ with at least one endpoint in a clique of $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$ is at most w.

Now, let $z := |\mathcal{E}_\mathcal{O}| - w$. Note that at least $\max\{0, z\}$ edges of $\mathcal{E}_\mathcal{O}$ are disjoint from $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$. Furthermore, every triangle of $\mathcal{T}_\mathcal{O}$ (resp. 4-clique of $\mathcal{Q}_\mathcal{O}$) that intersects precisely 1 vertex (resp. 2 vertices) of $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$ contributes an edge that is disjoint from $\mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}$. Since $\mathcal{E}_\mathcal{B}$ is a maximum matching of $G - \{v\colon v \text{ is a vertex in } \mathcal{Q}_\mathcal{B} \cup \mathcal{T}_\mathcal{B}\}$, we have

$$|\mathcal{E}_\mathcal{B}| \ge q_2 + t_1 + \max\{0, z\}. \tag{2}$$

Using the facts that $|\mathcal{Q}_\mathcal{O}| = q_2 + q_3 + q_4$ and $|\mathcal{T}_\mathcal{O}| = t_1 + t_2 + t_3$, we can rewrite z obtaining

$$z = |\mathcal{E}_\mathcal{O}| - 4|\mathcal{Q}_\mathcal{B}| - 3|\mathcal{T}_\mathcal{B}| + 3|\mathcal{Q}_\mathcal{O}| + 2|\mathcal{T}_\mathcal{O}| - q_2 + q_4 - t_1 + t_3. \tag{3}$$

Since $\mathrm{val}(\mathcal{B}) = 6|\mathcal{Q}_{\mathcal{B}}| + 3|\mathcal{T}_{\mathcal{B}}| + |\mathcal{E}_{\mathcal{B}}|$, using (2) we get

$$\mathrm{val}(\mathcal{B}) \geq 6|\mathcal{Q}_{\mathcal{B}}| + 3|\mathcal{T}_{\mathcal{B}}| + q_2 + t_1 + \max\{0, z\} \geq 6|\mathcal{Q}_{\mathcal{B}}| + 3|\mathcal{T}_{\mathcal{B}}| + q_2 + t_1 + z.$$

Now substituting the value of z given in (3), we obtain

$$\mathrm{val}(\mathcal{B}) \geq 2|\mathcal{Q}_{\mathcal{B}}| + 3|\mathcal{Q}_{\mathcal{O}}| + 2|\mathcal{T}_{\mathcal{O}}| + |\mathcal{E}_{\mathcal{O}}|. \tag{4}$$

Combining the fact that $\mathcal{Q}_{\mathcal{B}}$ is the solution output by HS$(4, t)$, which has an approximation ratio $2 + \varepsilon$, and the fact that $\mathrm{opt}_{\mathcal{K}_4}(G) \geq |\mathcal{Q}_{\mathcal{O}}|$, we have

$$|\mathcal{Q}_{\mathcal{B}}| \geq (\frac{1}{2} - \varepsilon')\mathrm{opt}_{\mathcal{K}_4}(G) \geq (\frac{1}{2} - \varepsilon')|\mathcal{Q}_{\mathcal{O}}|.$$

The above inequality together with (4) imply that

$$\mathrm{val}(\mathcal{B}) \geq (4 - 2\varepsilon')|\mathcal{Q}_{\mathcal{O}}| + 2|\mathcal{T}_{\mathcal{O}}| + |\mathcal{E}_{\mathcal{O}}| \geq (\frac{2}{3} - \varepsilon')\Big(6|\mathcal{Q}_{\mathcal{O}}| + 3|\mathcal{T}_{\mathcal{O}}| + |\mathcal{E}_{\mathcal{O}}|\Big) = (\frac{2}{3} - \varepsilon')\mathrm{val}(\mathcal{O}).$$

■

Theorem 2.4 *The algorithm $\mathcal{B}_5$ is a $(25/14 + \varepsilon)$-approximation algorithm for the $\mathcal{K}_5$-packing problem.*

Proof. The proof is similar to the one presented for Theorem 2.3. Let $\mathcal{O}$ be an optimal solution and $\mathcal{B}$ be the solution returned by the algorithm $\mathcal{B}_5$. Thus, $\mathrm{val}(\mathcal{O}) = 10|\mathcal{P}_{\mathcal{O}}| + 6|\mathcal{Q}_{\mathcal{O}}| + 3|\mathcal{T}_{\mathcal{O}}| + |\mathcal{E}_{\mathcal{O}}|$ and $\mathrm{val}(\mathcal{B}) = 10|\mathcal{P}_{\mathcal{B}}| + 6|\mathcal{Q}_{\mathcal{B}}| + 3|\mathcal{T}_{\mathcal{B}}| + |\mathcal{E}_{\mathcal{B}}|$.

Let p_i, $0 \leq i \leq 5$, be the number of 5-cliques of $\mathcal{P}_{\mathcal{O}}$ that intersect precisely i vertices of $\mathcal{P}_{\mathcal{B}} \cup \mathcal{Q}_{\mathcal{B}} \cup \mathcal{T}_{\mathcal{B}}$. Let q_i, $0 \leq i \leq 4$, be the number of 4-cliques of $\mathcal{Q}_{\mathcal{O}}$ that intersect precisely i vertices of $\mathcal{P}_{\mathcal{B}} \cup \mathcal{Q}_{\mathcal{B}} \cup \mathcal{T}_{\mathcal{B}}$. Let t_i, $0 \leq i \leq 3$, be the number of triangles of $\mathcal{T}_{\mathcal{O}}$ that intersect precisely i vertices of $\mathcal{P}_{\mathcal{B}} \cup \mathcal{Q}_{\mathcal{B}} \cup \mathcal{T}_{\mathcal{B}}$. Similarly as in the proof of Theorem 2.3, we get $p_0 = q_0 = t_0 = q_1 = p_1 = p_2 = 0$.

Observe that the number of vertices of $\mathcal{P}_{\mathcal{B}} \cup \mathcal{Q}_{\mathcal{B}} \cup \mathcal{T}_{\mathcal{B}}$ not covered by $\mathcal{P}_{\mathcal{O}} \cup \mathcal{Q}_{\mathcal{O}} \cup \mathcal{T}_{\mathcal{O}}$ is $5|\mathcal{P}_{\mathcal{B}}| + 4|\mathcal{Q}_{\mathcal{B}}| + 3|\mathcal{T}_{\mathcal{B}}| - (3p_3 + 4p_4 + 5p_5 + 2q_2 + 3q_3 + 4q_4 + t_1 + 2t_2 + 3t_3)$. We now define $z := |\mathcal{E}_{\mathcal{O}}| - 5|\mathcal{P}_{\mathcal{B}}| - 4|\mathcal{Q}_{\mathcal{B}}| - 3|\mathcal{T}_{\mathcal{B}}| + 3p_3 + 4p_4 + 5p_5 + 2q_2 + 3q_3 + 4q_4 + t_1 + 2t_2 + 3t_3$.

Observe that at least $\max\{0, z\}$ edges of $\mathcal{E}_{\mathcal{O}}$ are disjoint from $\mathcal{P}_{\mathcal{B}} \cup \mathcal{Q}_{\mathcal{B}} \cup \mathcal{T}_{\mathcal{B}}$. Furthermore, every triangle of $\mathcal{T}_{\mathcal{O}}$ (resp. 4-clique of $\mathcal{Q}_{\mathcal{O}}$, 5-clique of $\mathcal{P}_{\mathcal{O}}$) that intersects precisely 1 vertex (resp. 2 vertices, 3 vertices) of $\mathcal{P}_{\mathcal{B}} \cup \mathcal{Q}_{\mathcal{B}} \cup \mathcal{T}_{\mathcal{B}}$ contributes an edge that is disjoint from $\mathcal{P}_{\mathcal{B}} \cup \mathcal{Q}_{\mathcal{B}} \cup \mathcal{T}_{\mathcal{B}}$. Since $\mathcal{E}_{\mathcal{B}}$ is a maximum matching of $G - \{v\colon v \text{ is a vertex in } \mathcal{P}_{\mathcal{B}} \cup \mathcal{Q}_{\mathcal{B}} \cup \mathcal{T}_{\mathcal{B}}\}$, we have

$$|\mathcal{E}_{\mathcal{B}}| \geq t_1 + q_2 + p_3 + \max\{0, z\}. \tag{5}$$

Using the facts that $|\mathcal{P}_{\mathcal{O}}| = p_3 + p_4 + p_5$, $|\mathcal{Q}_{\mathcal{O}}| = q_2 + q_3 + q_4$ and $|\mathcal{T}_{\mathcal{O}}| = t_1 + t_2 + t_3$, we can rewrite z obtaining

$$z = |\mathcal{E}_{\mathcal{O}}| - 5|\mathcal{P}_{\mathcal{B}}| - 4|\mathcal{Q}_{\mathcal{B}}| - 3|\mathcal{T}_{\mathcal{B}}| + 4|\mathcal{P}_{\mathcal{O}}| + 3|\mathcal{Q}_{\mathcal{O}}| + 2|\mathcal{T}_{\mathcal{O}}| - p_3 + p_5 - q_2 + q_4 - t_1 + t_3. \tag{6}$$

Now, using (5) we get

$$\mathrm{val}(\mathcal{B}) = 10|\mathcal{P}_{\mathcal{B}}| + 6|\mathcal{Q}_{\mathcal{B}}| + 3|\mathcal{T}_{\mathcal{B}}| + |\mathcal{E}_{\mathcal{B}}| \geq 10|\mathcal{P}_{\mathcal{B}}| + 6|\mathcal{Q}_{\mathcal{B}}| + 3|\mathcal{T}_{\mathcal{B}}| + p_3 + q_2 + t_1 + \max\{0, z\}.$$

Thus,

$$\text{val}(\mathcal{B}) \geq 10|\mathcal{P}_\mathcal{B}| + 6|\mathcal{Q}_\mathcal{B}| + 3|\mathcal{T}_\mathcal{B}| + p_3 + q_2 + t_1 + z.$$

Substituting the value of z given in (6) and discarding some terms we obtain

$$\text{val}(\mathcal{B}) \geq 5|\mathcal{P}_\mathcal{B}| + 2|\mathcal{Q}_\mathcal{B}| + 4|\mathcal{P}_\mathcal{O}| + 3|\mathcal{Q}_\mathcal{O}| + 2|\mathcal{T}_\mathcal{O}| + |\mathcal{E}_\mathcal{O}|. \tag{7}$$

Observe now that each element of $\mathcal{P}_\mathcal{B} \cup \mathcal{Q}_\mathcal{B}$ intersects $\mathcal{P}_\mathcal{O} \cup \mathcal{Q}_\mathcal{O}$ in at most 5 vertices. Thus,

$$|\mathcal{P}_\mathcal{B}| + |\mathcal{Q}_\mathcal{B}| \geq \frac{1}{5}(|\mathcal{P}_\mathcal{O}| + |\mathcal{Q}_\mathcal{O}|). \tag{8}$$

Indeed, if $|\mathcal{P}_\mathcal{B}| + |\mathcal{Q}_\mathcal{B}| < \frac{1}{5}(|\mathcal{P}_\mathcal{O}| + |\mathcal{Q}_\mathcal{O}|)$, there would be an element of $\mathcal{P}_\mathcal{O} \cup \mathcal{Q}_\mathcal{O}$ that does not intersect any of the elements from $\mathcal{P}_\mathcal{B} \cup \mathcal{Q}_\mathcal{B}$, contradicting the fact that the set $\mathcal{P}_\mathcal{B} \cup \mathcal{Q}_\mathcal{B}$ was found by algorithm $\mathcal{B}_5$.

Since HS$(5, t)$ has an approximation ratio $5/2 + \varepsilon$, we have

$$|\mathcal{P}_\mathcal{B}| \geq (\frac{2}{5} - \varepsilon')|\mathcal{P}_\mathcal{O}|.$$

Now multiplying inequality (8) by 2 and adding with the inequality above multiplied by 3, we get $5|\mathcal{P}_\mathcal{B}| + 2|\mathcal{Q}_\mathcal{B}| \geq (8/5 - 3\varepsilon')|\mathcal{P}_\mathcal{O}| + 2/5|\mathcal{Q}_\mathcal{O}|$. Combining this inequality with inequality (7) we obtain

$$\begin{aligned}
\text{val}(\mathcal{B}) &\geq (8/5 - 3\varepsilon')|\mathcal{P}_\mathcal{O}| + 2/5|\mathcal{Q}_\mathcal{O}| + 4|\mathcal{P}_\mathcal{O}| + 3|\mathcal{Q}_\mathcal{O}| + 2|\mathcal{T}_\mathcal{O}| + |\mathcal{E}_\mathcal{O}| \\
&= (28/5 - 3\varepsilon')|\mathcal{P}_\mathcal{O}| + 17/5|\mathcal{Q}_\mathcal{O}| + 2|\mathcal{T}_\mathcal{O}| + |\mathcal{E}_\mathcal{O}| \\
&= \frac{1}{10}(\frac{28}{5} - 3\varepsilon')10|\mathcal{P}_\mathcal{O}| + \frac{1}{6}(\frac{17}{5})\,6|\mathcal{Q}_\mathcal{O}| + \frac{1}{3}(2)\,3|\mathcal{T}_\mathcal{O}| + |\mathcal{E}_\mathcal{O}| \\
&\geq (\frac{28}{50} - \varepsilon')\Big(10|\mathcal{P}_\mathcal{O}|) + 6|\mathcal{Q}_\mathcal{O} + 3|\mathcal{T}_\mathcal{O}| + |\mathcal{E}_\mathcal{O}\Big) \\
&= (\frac{14}{25} - \varepsilon')\text{val}(\mathcal{O}).
\end{aligned}$$

■

We are not sure whether the ratio $(3/2 + \varepsilon)$ (resp. $(25/14 + \varepsilon)$) for the algorithm $\mathcal{B}_4$ (resp. $\mathcal{B}_5$) is tight. We note that for $r \geq 6$, using the same approach it is not possible to show that the ratio of the algorithm $\mathcal{B}_r$ is smaller than 2 (as we need a better ratio for HS(r, t)).

3 The APX-hardness of the $\mathcal{K}_3$-packing problem

In this section we prove that the $\mathcal{K}_3$-packing problem is APX-hard on graphs with maximum degree 5. We also show that this problem is APX-hard even on irredundant graphs with maximum degree 4. We recall that we defined a graph to be irredundant if each of its edges belongs to some triangle. As we know that the $\mathcal{K}_3$-packing problem has a constant approximation algorithm, we can conclude that it is an APX-complete problem.

We show first the result for graphs with maximum degree 5, and then for graphs with maximum degree 4. In both cases we consider the problem of finding the *maximum number of vertex-disjoint* *triangles* in a graph, denoted here as VTP, and known to be APX-complete [2]. (This problem is equivalent to the $\{K_3\}$-packing problem; it is just more convenient to simplify the counting arguments.)

The second proof is significantly more elaborate than the first: its structure is analogous to the reduction presented by Caprara and Rizzi [2] to show that the VTP problem is APX-complete on graphs with maximum degree 4.

Theorem 3.1 *The $\mathcal{K}_3$-packing problem is* APX*-hard on graphs with maximum degree* 5.

Proof. We show an L-reduction from the VTP problem to the $\mathcal{K}_3$-packing problem. For that, we shall exhibit a pair of functions (f, g), and constants α and β, in accordance with the definition of L-reduction given in Section 1.

Let G be an irredundant graph with $\Delta(G) = 4$. Define $G' := f(G)$ as the union of two copies, say G_1 and G_2, of G together with the set of edges

$$\{u_1u_2\colon u_1 \in V_{G_1}, u_2 \in V_{G_2}, \text{ and } u_1, u_2 \text{ correspond to the same vertex } u \in V_G\}.$$

We first show that

$$\mathrm{opt}_{\mathcal{K}_3}(G') = 3\,\mathrm{opt}_{\mathrm{VTP}}(G) + n_G. \tag{9}$$

Indeed, if $\mathcal{T}^*$ is an optimal solution of the VTP problem in G, then there is a $\{K_2, K_3\}$-packing of G' consisting of the triangles in G_1 and G_2 that are copies of triangles in $\mathcal{T}^*$, and set of edges $\{u_1u_2\colon u_1 \in V_{G_1}, u_2 \in V_{G_2}, \text{ and } u_1, u_2 \text{ correspond to the same vertex } u \text{ of } G \text{ not covered by } \mathcal{T}^*\}$. Since the number of vertices of G not covered by $\mathcal{T}^*$ is $n_G - 3\,\mathrm{opt}_{\mathrm{VTP}}(G)$, we have $\mathrm{opt}_{\mathcal{K}_3}(G') \geq 6\,\mathrm{opt}_{\mathrm{VTP}}(G) + n_G - 3\,\mathrm{opt}_{\mathrm{VTP}}(G) = 3\,\mathrm{opt}_{\mathrm{VTP}}(G) + n_G$. On the other hand, if an optimal solution of the $\mathcal{K}_3$-packing problem in G' has t' triangles and e' edges, since $e' \leq \frac{n_{G'} - 3t'}{2} = n_G - \frac{3}{2}t'$, we have $\mathrm{opt}_{\mathcal{K}_3}(G') = 3t' + e' \leq \frac{3}{2}t' + n_G$. Of course, $t' \leq 2\,\mathrm{opt}_{\mathrm{VTP}}(G)$, and thus $\mathrm{opt}_{\mathcal{K}_3}(G') \leq 3\,\mathrm{opt}_{\mathrm{VTP}}(G) + n_G$. Hence, statement (9) holds.

Let $\mathcal{T}^*$ be an optimal solution of the VTP problem in G. Suppose that there exists a triangle $T \in \mathcal{T}^*$, such that T has 5 neighbouring vertices in $V_G \setminus V_T$ that are not covered by $\mathcal{T}^*$. Since $\Delta(G) = 4$, one pair of them, say v_1, v_2 is adjacent to the same vertex, say x from V_T; another pair, say v_3, v_4 (disjoint from v_1, v_2), is adjacent to the same vertex, say y from V_T. Note that the third vertex of V_T, say z, has degree at least 3. Furthermore, since G is irredundant and $\Delta(G) = 4$, we have that $v_1v_2, v_3v_4 \in E_G$. Indeed, since G is irredundant, edge v_1x (resp. v_2x) has to be in some triangle. Since $d_G(x) = d_G(y) = \Delta(G) = 4$, the only possible triangle having edge xv_1, not using v_1v_2, is the triangle $[x, v_1, z]$ (see the Figure 1). But now, the only possible triangle having edge xv_2 is the triangle $[x, v_2, v_1]$, and hence, $v_1v_2 \in E_G$. Similarly, $v_3v_4 \in E_G$.

Thus, by replacing T with $[x, v_1, v_2]$ and $[y, v_3, v_4]$, we obtain a solution for the VTP problem that has more triangles than $\mathcal{T}^*$ does, a contradiction. Hence, each triangle from $\mathcal{T}^*$ has at most 4 neighbours not covered by $\mathcal{T}^*$. Note, furthermore, that since G is irredundant, each vertex not covered by $\mathcal{T}^*$ is adjacent to at least one vertex covered by $\mathcal{T}^*$. Indeed, suppose that there is a vertex v not covered by $\mathcal{T}^*$, and not adjacent to any vertex covered by $\mathcal{T}^*$. Since G is irredundant, v is a vertex of a triangle T. Observe that none of the vertex of T is covered by $\mathcal{T}^*$, and thus, $\mathcal{T}^*$ is not an optimal solution of the VTP problem in G, a contradiction. It thus follows that the number of vertices in G not covered by $\mathcal{T}^*$ is at most $4\,\mathrm{opt}_{\mathrm{VTP}}(G)$, that

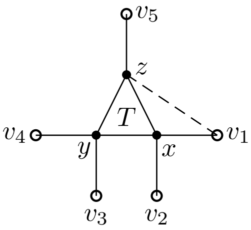


Figure 1: Vertices v_1, v_2, v_3, v_4 and v_5 are the neighbours of V_T that are not covered by $\mathcal{T}^*$.

is, $n_G - 3\,\mathrm{opt}_{\mathrm{VTP}}(G) \le 4\,\mathrm{opt}_{\mathrm{VTP}}(G)$. Using (9) we have $\mathrm{opt}_{\mathcal{K}_3}(G') \le 10\,\mathrm{opt}_{\mathrm{VTP}}(G)$. Thus, for $\alpha = 10$ condition (C1) of the definition of L-reduction is satisfied.

Given a $\{K_2, K_3\}$-packing $\mathcal{A}$ of $G' = f(G)$, we define $g(G, \mathcal{A})$ as a largest of the two sets $\mathcal{T}_{\mathcal{A}} \cap G_1$, $\mathcal{T}_{\mathcal{A}} \cap G_2$. Suppose, without loss of generality, that $g(G, \mathcal{A}) = \mathcal{T}_{\mathcal{A}} \cap G_1$. Let $t'_1 := |\mathcal{T}_{\mathcal{A}} \cap G_1|$, $t'_2 := |\mathcal{T}_{\mathcal{A}} \cap G_2|$, $e'_1 := |\mathcal{E}_{\mathcal{A}} \cap G_1|$, $e'_2 := |\mathcal{E}_{\mathcal{A}} \cap G_2|$, and e' be the number of edges in $\mathcal{E}_{\mathcal{A}}$ with one endpoint in G_1 and the other in G_2. Of course, $t'_1 \le \mathrm{opt}_{\mathrm{VTP}}(G)$. Thus, $\frac{1}{2}t'_1 + \frac{3}{2}t'_1 - 2\,\mathrm{opt}_{\mathrm{VTP}}(G) \le 0$. Since $t'_2 \le t'_1$, we have $\frac{1}{2}t'_1 + \frac{3}{2}t'_2 - 2\,\mathrm{opt}_{\mathrm{VTP}}(G) \le 0$, or equivalently,

$$\mathrm{opt}_{\mathrm{VTP}}(G) - t'_1 \le 3\,\mathrm{opt}_{\mathrm{VTP}}(G) + (\frac{3}{2}t'_1 + \frac{3}{2}t'_2 + e'_1 + e'_2 + e') - (3t'_1 + 3t'_2 + e'_1 + e'_2 + e'). \quad (10)$$

Now, $3t'_1 + 3t'_2 + 2e'_1 + 2e'_2 + 2e' \le n_{G'} = 2n_G$, and hence, $\frac{3}{2}t'_1 + \frac{3}{2}t'_2 + e'_1 + e'_2 + e' \le n_G$. Thus, from (10) we have $\mathrm{opt}_{\mathrm{VTP}}(G) - t'_1 \le 3\,\mathrm{opt}_{\mathrm{VTP}}(G) + n_G - (3t'_1 + 3t'_2 + e'_1 + e'_2 + e')$. Using (9), we get $\mathrm{opt}_{\mathrm{VTP}}(G) - t'_1 \le \mathrm{opt}_{\mathcal{K}_3}(G') - \mathrm{val}_{\mathcal{K}_3}(G', \mathcal{A})$. Thus, condition (C2) holds with $\beta = 1$. ■

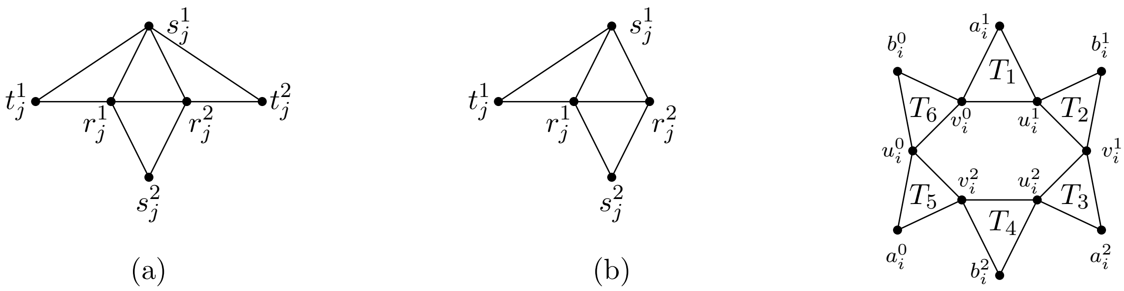
Theorem 3.2 *The $\mathcal{K}_3$-packing problem is* APX*-hard on the class of irredundant graphs with maximum degree* 4.

Proof. We show an L-reduction from the SAT problem we have defined in Section 1. For that, as in the previous proof, we shall exhibit a pair of functions (f, g), and constants α and β, according to the definition of L-reduction given in Section 1. Let $\varphi = (C, X)$ with $C = \{c_1, c_2, \ldots, c_l\}$ and $X = \{x_1, x_2, \ldots, x_p\}$ be an instance of SAT. Let m_i denote the number of occurrences of x_i. We may assume, without loss of generality, that $m_i \ge 2$ (for if x_i appears only in one clause we can set x_i to the value which satisfies that clause). We define $G' := f(\varphi)$ in the following way.

To each clause c_j we associate a *test component* $\mathcal{C}_j$. The test component of a clause with two literals consists of 4 triangles $[t^1_j, s^1_j, r^1_j]$, $[s^1_j, r^1_j, r^2_j]$, $[s^2_j, r^1_j, r^2_j]$, $[s^1_j, r^2_j, t^2_j]$ (see Figure 2(a)), whereas the test component associated with a clause with one literal consists of 3 triangles $[t^1_j, s^1_j, r^1_j]$, $[s^1_j, r^1_j, r^2_j]$, $[s^2_j, r^1_j, r^2_j]$ (see Figure 2(b)).

To each variable x_i we associate a *truth component* $\mathcal{X}_i$, (see Figure 2(c)). This component consists of $2m_i$ triangles $T_1, \ldots, T_{2m_i}$, where $T_{2k-1} = [a^k_i, v^{k-1}_i, u^k_i]$ and $T_{2k} = [b^k_i, u^k_i, v^k_i]$, $k = 1, \ldots, m_i$ (all upper indices being modulo m_i). The *parity* of T_k is the parity of k.

The graph G' is obtained by connecting the test and truth components as follows. Let c_j be a clause with two literals and let x_1, x_2 be the variables which occur in c_j. If x_i occurs positive (resp. negated) in c_j, then identify the vertex t^i_j of the test component $\mathcal{C}_j$, with a vertex a^k_i (resp.



(a) (b) (c)

Figure 2: (a) The test component of a clause c_j that has two literals. (b) The test component of a clause c_j that has one literal. (c)The truth component of a variable x_i with $m_i = 3$.

b_i^k) of the truth component $\mathcal{X}_i$ which has not yet been involved in any identification. Similarly, let c_j be a clause with one literal, say, x_1. If x_1 occurs positive (resp. negated) in c_j, then identify the vertex t_j^1 of $\mathcal{C}_j$, with a vertex a_1^k (resp. b_1^k) of $\mathcal{X}_1$ which has not yet been involved in any identification. Note that G' is irredundant and $\Delta(G') = 4$.

A maximal $\{K_2, K_3\}$-packing $\mathcal{A}$ of G' is called *canonical* if, for each truth component, it contains either all even or all odd triangles, and for each test component $\mathcal{C}_j$ it contains the triangle $[r_j^1, r_j^2, s_j^2]$, and possibly one of the edges $t_j^1 s_j^1$ or $t_j^2 s_j^1$. First, we show that the following statement holds.

$$\text{Given a non-canonical } \{K_2, K_3\}\text{-packing } \mathcal{A} \text{ of } G'\text{, one can find in polynomial time a canonical packing of } G' \text{ whose value is at least the value of } \mathcal{A}. \tag{11}$$

We will construct the desired packing $\mathcal{A}'$ from $\mathcal{A}$ (we start with $\mathcal{A}' = \mathcal{A}$). Initially, for each test component $\mathcal{C}_j$, $1 \le j \le l$, we remove from $\mathcal{A}'$ the triangles and edges that are in $\mathcal{C}_j$ and add $[r_j^1, r_j^2, s_j^2]$ to it. Furthermore, if one of the edges $t_j^1 s_j^1$, $t_j^2 s_j^1$ is covered by $\mathcal{A}$, then we add to $\mathcal{A}'$ the one that is covered by $\mathcal{A}$. Observe that for each $\mathcal{C}_j$, the value of $\mathcal{A}$ restricted to $\mathcal{C}_j$ is at most 4. Moreover, if the value of $\mathcal{A}$ restricted to $\mathcal{C}_j$ is exactly 4, then one of the edges $t_j^1 s_j^1$, $t_j^2 s_j^1$ is covered by $\mathcal{A}$. Thus, so far the resulting packing $\mathcal{A}'$ has a value that is at least the value of $\mathcal{A}$.

Moreover, for each i, $1 \le i \le p$, if the triangles of the truth component $\mathcal{X}_i$ that are in $\mathcal{T}_\mathcal{A}$ are not all of the same parity, we do the following (depending on the number of occurrences of x_i).

1. $m_i = 3$.
   We may assume, without loss of generality, that x_i appears negated in one clause, say c_j, and positive in two clauses (for if x_i appears only negated or only positive, we can set it to the value that satisfies all the clauses in which it appears in). Let t_j^k, $k \in \{1, 2\}$ be the vertex of $\mathcal{C}_j$ incident with $\mathcal{X}_i$. Then, we remove from $\mathcal{A}'$ the triangles and edges that are in $\mathcal{X}_i$, and add all even triangles of $\mathcal{X}_i$ to $\mathcal{A}'$. Furthermore, if $t_j^k s_j^1$ is in $\mathcal{A}'$, we remove it. We next show that after those changes the value of $\mathcal{A}'$ is at least the value of $\mathcal{A}$.

   (a) If there is no triangle of $\mathcal{X}_i$ that is in $\mathcal{T}_\mathcal{A}$, then there are at most 6 edges of $\mathcal{X}_i$ that are in $\mathcal{E}_A$, one from each triangle. Hence, the value of packing decreases by at most 7. Since the value of the packing increases by 9, we have that the value of $\mathcal{A}'$ increases.

(b) If there is exactly one triangle of $\mathcal{X}_i$ that is in $\mathcal{T}_\mathcal{A}$, then there are at most 5 edges of $\mathcal{X}_i$ that are in $\mathcal{E}_A$, one from each other triangle. Thus, the value of $\mathcal{A}'$ decreases by at most 9, and increases by 9.

(c) If there are exactly two triangles of $\mathcal{X}_i$ that are in $\mathcal{T}_\mathcal{A}$, then, there are at most 2 edges of $\mathcal{X}_i$ that are in $\mathcal{E}_A$ (see examples on Figure 3(a) and (b)). Hence, we have that the value of packing $\mathcal{A}'$ decreases by at most 9, and increases by 9.

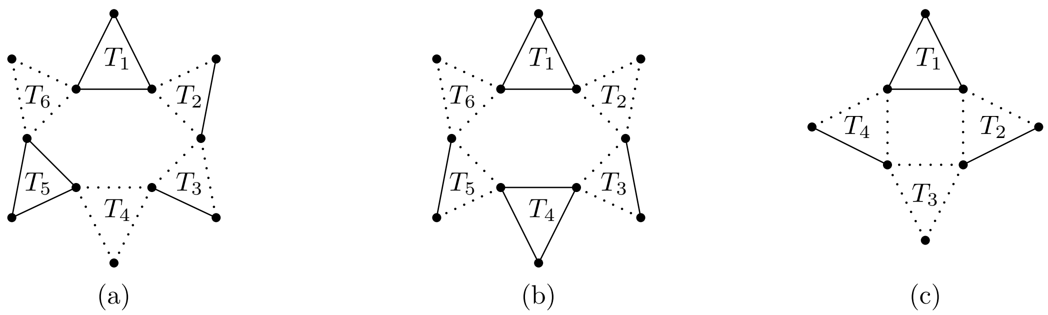


Figure 3: Triangles and edges with full lines are in $\mathcal{A}$. (a) $m_i = 3$, there are exactly two triangles of $\mathcal{X}_i$ that are in $\mathcal{T}_\mathcal{A}$, and they are of the same parity. (b) $m_i = 3$ and there are exactly two triangles of $\mathcal{X}_i$ that are in $\mathcal{T}_\mathcal{A}$, not of the same parity. (c) $m_i = 2$ and there is a triangle of $\mathcal{X}_i$ that is in $\mathcal{T}_\mathcal{A}$. In all cases, there are at most 2 edges of $\mathcal{X}_i$ that are in $\mathcal{E}_A$.

2. $m_i = 2$.
   We may assume, without loss of generality, that x_i appears negated in one clause, say c_j, and positive in another. Then, we remove from $\mathcal{A}'$ the triangles and edges that are in $\mathcal{X}_i$, and add two even triangles of $\mathcal{X}_i$ to $\mathcal{A}'$. Furthermore, if $t_j^1 s_j^1$ is in $\mathcal{A}'$, we remove it. We next show that those changes yield a packing $\mathcal{A}'$ whose value is at least the value of $\mathcal{A}$.

   (a) If there is no triangle of $\mathcal{X}_i$ that is in $\mathcal{T}_\mathcal{A}$, then there are at most 4 edges of $\mathcal{X}_i$ that are in $\mathcal{E}_A$, one from each triangle. Hence, the value of packing $\mathcal{A}'$ decreases by at most 5. Since the value of the packing increases by 6, we have that the value of $\mathcal{A}'$ increases.

   (b) If there is a triangle of $\mathcal{X}_i$ that is in $\mathcal{T}_\mathcal{A}$, then there is only one such triangle, say T_k. Furthermore, there are at most 2 edges of $\mathcal{X}_i$ that are in $\mathcal{E}_A$, since the number of vertices in $\mathcal{X}_i - V_{T_k}$ is 5 (see an example on Figure 3(c)). Hence, the value of $\mathcal{A}'$ decreases by at most 6, and increases by 6.

Finally, for each test component $\mathcal{C}_j$, if s_j^1 is not already an endpoint of an edge in $\mathcal{E}_{A'}$, then whenever possible, we add one of the edges $t_j^1 s_j^1$ or $t_j^2 s_j^1$ to $\mathcal{A}'$. That is, if the corresponding clause c_j has two literals, then, if t_j^1 is not covered by $\mathcal{A}'$, we add $t_j^1 s_j^1$ to $\mathcal{A}'$; otherwise, if t_j^2 is not covered by $\mathcal{A}'$, we add $t_j^2 s_j^1$ to $\mathcal{A}'$. If, however, the clause c_j has one literal, then if t_j^1 is not covered by $\mathcal{A}'$, we add $t_j^1 s_j^1$ to $\mathcal{A}'$.

An example of the construction of $\mathcal{A}'$ is shown in Figure 4.

Note that the resulting packing $\mathcal{A}'$ is a canonical packing of G' whose value is at least the value of $\mathcal{A}$. We have thus proved (11).

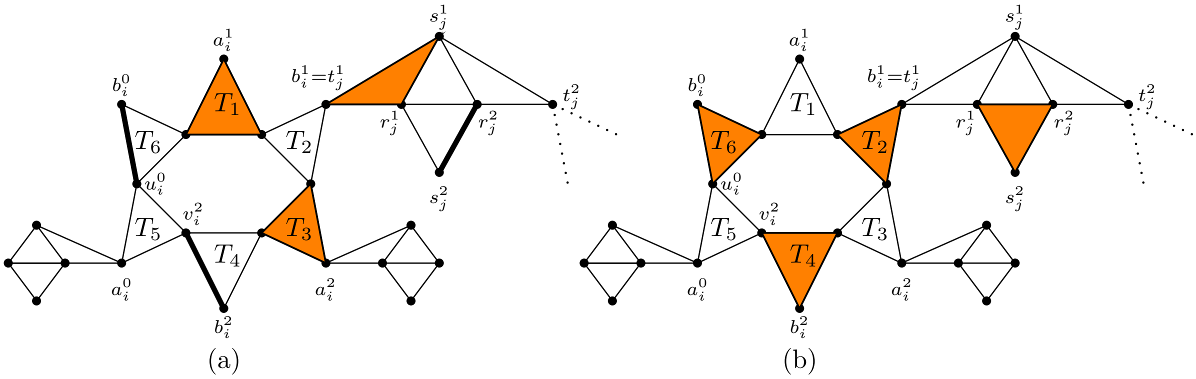


Figure 4: An example of the construction of $\mathcal{A}'$ (case $m_i = 3$, and x_i appears negated in only one clause, say c_j, and positive in two other clauses). Dotted lines indicate edges of another truth component. (a) Shows a non-canonical $\{K_2, K_3\}$-packing $\mathcal{A}$ restricted to $\mathcal{X}_i$ and $\mathcal{C}_j$ (highlighted edges and triangles are in $\mathcal{A}$). In the first step, $[t_j^1, s_j^1, r_j^1]$ and $s_j^2 r_j^2$ are removed from $\mathcal{A}'$, and $[r_j^1, r_j^2, s_j^2]$, $t_j^1 s_j^1$ are added to $\mathcal{A}'$. In the second step, $T_1, T_3, u_i^0 b_i^0, v_i^2 b_i^2, t_j^1 s_j^1$ are removed from $\mathcal{A}'$ and triangles T_2, T_4, T_6 are added to $\mathcal{A}'$. (b) The resulting packing $\mathcal{A}'$.

We observe that a given canonical packing $\mathcal{A}'$ of G' corresponds to a truth assignment for the variables in X in the following way. If $\mathcal{A}'$ contains all even (resp. odd) triangles of the truth component $\mathcal{X}_i$, then x_i is set to true (resp. false). On the other hand, given a truth assignment for the variables in X, we can construct a canonical packing $\mathcal{A}'$ of G' in the following way. If x_i is true (resp. false), we add all even (resp. odd) triangles of $\mathcal{X}_i$ to $\mathcal{A}'$. For each test component $\mathcal{C}_j$ we add the triangle $[r_j^1, r_j^2, s_j^2]$ to $\mathcal{A}'$. Moreover, if the corresponding clause c_j has two literals, then if t_j^1 is not covered by $\mathcal{A}'$, we add $t_j^1 s_j^1$ to $\mathcal{A}'$; otherwise, if t_j^2 is not covered by $\mathcal{A}'$, we add $t_j^2 s_j^1$ to the packing. If, however, the clause c_j has one literal, then if t_j^1 is not covered by $\mathcal{A}'$, we add $t_j^1 s_j^1$ to $\mathcal{A}'$.

Consider now a canonical packing $\mathcal{A}'$ and the corresponding truth assignment for the variables in X. Let c_j be a clause with two literals, and let x_1, x_2 be the variables which occur in c_j. Note that t_j^i (for $i = 1, 2$) is not covered by a triangle of $\mathcal{A}'$ that belongs to the corresponding truth component $\mathcal{X}_i$, if and only if, x_i is set to the value that satisfies c_j. Thus, from the construction of the canonical packing we have that the following statements are equivalent: clause c_j is satisfiable; at least one of t_j^1, t_j^2 is not covered by a triangle of $\mathcal{A}'$ that belongs to the corresponding truth component; exactly one of $t_j^1 s_j^1$, $t_j^2 s_j^1$ is in $\mathcal{E}_{\mathcal{A}'}$; the value of $\mathcal{A}'$ restricted to $\mathcal{C}_j$ is 4. Similar statements hold for a clause with one literal. Thus, the value of $\mathcal{A}'$ restricted to $\mathcal{C}_j$ is 4 (resp. 3), if and only if, c_j is satisfiable (resp. not satisfiable). Moreover, exactly m_i triangles of each $\mathcal{X}_i$

are in $\mathcal{A}'$. Thus, the following claim holds.

A canonical packing $\mathcal{A}'$ of G' with value $\sum_{i=1}^{p} 3m_i + 4k + 3(l-k)$ corresponds to a truth assignment for the variables in X that satisfies exactly k clauses of φ, and vice versa. (12)

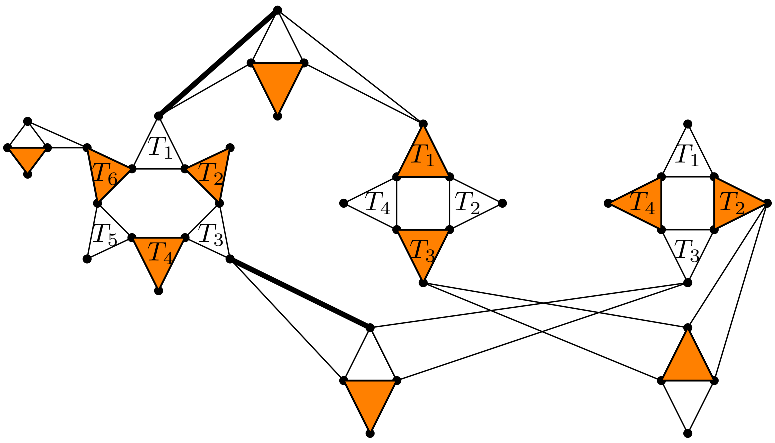


Figure 5: An example of a canonical packing $\mathcal{A}'$ of G' and a corresponding truth assignment for the variables of the SAT problem instance $\varphi = (x_1 \vee x_2) \wedge \neg x_1 \wedge (x_1 \vee x_3) \wedge (x_2 \vee \neg x_3)$: x_1 and x_3 are set to true, x_2 is set to false.

Now, given a $\{K_2, K_3\}$-packing $\mathcal{A}$ of $G' := f(\varphi)$, we define a truth assignment $g(\varphi, \mathcal{A})$ in the following way. First, find a canonical packing $\mathcal{A}'$ of G' with value at least the value of $\mathcal{A}$. Set a variable x_i to true (resp. false) if $\mathcal{A}'$ contains all even (resp. odd) triangles of the truth component $\mathcal{X}_i$.

We next show that

$$\mathrm{opt}_{\mathcal{K}_3}(G') = \sum_{i=1}^{p} 3m_i + \mathrm{opt}_{\mathrm{SAT}}(\varphi) + 3l. \tag{13}$$

Indeed, from (12) we have that an optimal solution of SAT(φ) corresponds to a canonical packing $\mathcal{A}'$ of G' with the value $\sum_{i=1}^{p} 3m_i + 4\,\mathrm{opt}_{\mathrm{SAT}}(\varphi) + 3(l - \mathrm{opt}_{\mathrm{SAT}}(\varphi))$. Thus, $\mathrm{opt}_{\mathcal{K}_3}(G') \geq \sum_{i=1}^{p} 3m_i + \mathrm{opt}_{\mathrm{SAT}}(\varphi) + 3l$. On the other hand, let $\mathcal{A}$ be a $\{K_2, K_3\}$-packing of G'. If the corresponding feasible solution $g(\varphi, \mathcal{A})$ of SAT(φ) satisfies k clauses, we have that $k \leq \mathrm{opt}_{\mathrm{SAT}}(\varphi)$. Furthermore, $\mathrm{val}_{\mathcal{K}_3}(G', \mathcal{A}) \leq \mathrm{val}_{\mathcal{K}_3}(G', \mathcal{A}')$, and by (12), $\mathrm{val}_{\mathcal{K}_3}(G', \mathcal{A}') = \sum_{i=1}^{p} 3m_i + k + 3l$. Hence, we have that $\mathrm{opt}_{\mathcal{K}_3}(G') \leq \sum_{i=1}^{p} 3m_i + \mathrm{opt}_{\mathrm{SAT}}(\varphi) + 3l$. We have thus proved (13).

Since each clause has at most 2 literals, we have $\sum_{i=1}^{p} m_i \leq 2l$. Furthermore, note that the optimal value of SAT problem on φ is at least $\frac{l}{2}$, since at least half of the clauses can be satisfied by a simple greedy approach. Thus, from (13) we have $\mathrm{opt}_{\mathcal{K}_3}(G') \leq 9l + \mathrm{opt}_{\mathrm{SAT}}(\varphi) \leq 19\,\mathrm{opt}_{\mathrm{SAT}}(\varphi)$. Hence, taking $\alpha = 19$ we can conclude that condition (C1) of the definition of L-reduction holds.

Finally, suppose that $\mathrm{val}_{\mathrm{SAT}}(\varphi, g(\varphi, \mathcal{A})) = k$, that is, the truth assignment $g(\varphi, \mathcal{A})$ satisfies exactly k clauses of φ. Hence, from (12) we have $\mathrm{val}_{\mathcal{K}_3}(G', \mathcal{A}') = \sum_{i=1}^{p} 3m_i + k + 3l$. From this, the equality (13), and the fact that $\mathrm{val}_{\mathcal{K}_3}(G', \mathcal{A}') \geq \mathrm{val}_{\mathcal{K}_3}(G', \mathcal{A})$, we have

$$\mathrm{opt}_{\mathrm{SAT}}(\varphi) - \mathrm{val}_{\mathrm{SAT}}(\varphi, g(\varphi, \mathcal{A})) \leq \mathrm{opt}_{\mathcal{K}_3}(G') - \mathrm{val}_{\mathcal{K}_3}(G', \mathcal{A}).$$

Thus, (C2) holds if we take $\beta = 1$. ■

4 Approximation algorithm for the $\mathcal{K}_3$-packing problem

Let us denote by $\mathcal{C}_3(\rho)$ an algorithm for the $\mathcal{K}_3$-packing problem that consists of the algorithm Basic$_3$ together with a Procedure $\mathcal{P}_3$ that is a ρ-approximation algorithm for the VTP problem. We are interested in the performance ratio of $\mathcal{C}_3(\rho)$.

Theorem 4.1 *Let $\mathcal{P}_3$ be a ρ-approximation algorithm for the* VTP *problem which produces for any input graph G a triangle packing that is maximal. Then the algorithm $\mathcal{C}_3(\rho)$ is a $(1 + \frac{1}{3}\rho)$-approximation algorithm for the $\mathcal{K}_3$-packing problem.*

Proof. Let G be a graph and $\mathcal{A}$ the solution returned by the algorithm $\mathcal{C}_3(\rho)$ applied to G. Let $\mathcal{O}$ be an optimal solution for the $\mathcal{K}_3$-packing problem on G with the largest possible number of triangles in common with $\mathcal{A}$. Let t_i (resp. o_i), $0 \leq i \leq 3$, be the number of triangles of $\mathcal{A}$ (resp. $\mathcal{O}$) that intersect exactly i vertices of $\mathcal{T}_{\mathcal{O}}$ (resp. $\mathcal{T}_{\mathcal{A}}$).

We show first that $t_0 = 0$. Suppose that $t_0 > 0$ and that T is a triangle of $\mathcal{A}$ that intersects no triangle of $\mathcal{O}$. If at most two edges of $\mathcal{E}_{\mathcal{O}}$ are adjacent to T, then we can replace these edges with T, obtaining a $\{K_2, K_3\}$-packing with value greater than the value of $\mathcal{O}$, a contradiction. Thus, there are 3 edges of $\mathcal{E}_{\mathcal{O}}$ adjacent to T. Removing these edges and adding T to $\mathcal{O}$, we get an optimal solution of the $\mathcal{K}_3$-packing problem that has more triangles in common with $\mathcal{A}$ than $\mathcal{O}$ does, which is again a contradiction. Thus, $t_0 = 0$. Since $\mathcal{P}_3$ returns a maximal triangle packing, o_0 must be zero.

Now, counting the vertices that are in the intersection of triangles from $\mathcal{A}$ and $\mathcal{O}$ we get

$$3t_3 + 2t_2 + t_1 = 3o_3 + 2o_2 + o_1. \tag{14}$$

We next define e_1 (resp. e_0) as the number of edges in $\mathcal{E}_{\mathcal{O}}$ with at least one (resp. none) of its endpoints in a triangle of $\mathcal{A}$. Clearly, e_1 is at most the number of vertices v of the triangles in $\mathcal{A}$ such that v is not covered by a triangle from $\mathcal{O}$, that is,

$$e_1 \leq 2t_1 + t_2. \tag{15}$$

Let $G' := G - \{v\colon v \text{ is a vertex of a triangle in } \mathcal{T}_{\mathcal{A}}\}$. Note that a matching of G' can be obtained by taking one edge of each triangle of $\mathcal{O}$ that has exactly one vertex in common with a triangle of $\mathcal{A}$, and taking the edges of $\mathcal{E}_{\mathcal{O}}$ that have no vertex in common with any triangle of $\mathcal{A}$. Hence, as $\mathcal{E}_{\mathcal{A}}$ is a maximum matching of G', we have $|\mathcal{E}_{\mathcal{A}}| \geq o_1 + e_0$. From this, and the inequality (15), we have

$$|\mathcal{E}_{\mathcal{O}}| = e_1 + e_0 \leq 2t_1 + t_2 + |\mathcal{E}_{\mathcal{A}}| - o_1. \tag{16}$$

We now consider the ratio r of the value of $\mathcal{O}$ to the value of $\mathcal{A}$, that is, $r := (3|\mathcal{T}_\mathcal{O}|+|\mathcal{E}_\mathcal{O}|)/(3|\mathcal{T}_\mathcal{A}|+|\mathcal{E}_\mathcal{A}|)$. Using (16) and the fact that $|\mathcal{T}_\mathcal{O}| = o_3 + o_2 + o_1$, we get

$$r \le \frac{3(o_3 + o_2 + o_1) + (2t_1 + t_2 + |\mathcal{E}_\mathcal{A}| - o_1)}{3|\mathcal{T}_\mathcal{A}| + |\mathcal{E}_\mathcal{A}|}.$$

Since $|\mathcal{E}_\mathcal{A}| \ge 0$, and $r \ge 1$, we can remove $|\mathcal{E}_\mathcal{A}|$ in the last inequality, obtaining

$$r \le \frac{3(o_3 + o_2 + o_1) + (2t_1 + t_2 - o_1)}{3|\mathcal{T}_\mathcal{A}|}.$$

Using (14), we have

$$r \le \frac{(3t_3 + 2t_2 + t_1) + (o_2 + 2o_1) + (2t_1 + t_2 - o_1)}{3|\mathcal{T}_\mathcal{A}|} = \frac{3(t_3 + t_2 + t_1) + (o_2 + o_1)}{3|\mathcal{T}_\mathcal{A}|} = \frac{3|\mathcal{T}_\mathcal{A}| + (o_2 + o_1)}{3|\mathcal{T}_\mathcal{A}|}.$$

Since $o_2 + o_1 \le |\mathcal{T}_\mathcal{O}|$, we have $r \le 1 + \frac{1}{3}\frac{|\mathcal{T}_\mathcal{O}|}{|\mathcal{T}_\mathcal{A}|}$. As $|\mathcal{T}_\mathcal{O}| \le \mathrm{opt}_{\mathrm{VTP}}(G)$, and $\mathcal{P}_3$ is a ρ-approximation algorithm for the VTP problem,

$$\frac{|\mathcal{T}_\mathcal{O}|}{|\mathcal{T}_\mathcal{A}|} \le \frac{\mathrm{opt}_{\mathrm{VTP}}(G)}{|\mathcal{T}_\mathcal{A}|} \le \rho, \text{ and hence, } r \le 1 + \frac{1}{3}\rho.$$

■

Corollary 4.2 *There is a* $(\frac{3}{2} + \varepsilon)$*-approximation algorithm for the* $\mathcal{K}_3$*-packing problem.*

Proof. Hurkens and Schrijver [7] showed that HS$(3, t)$ is a $(\frac{3}{2} + \varepsilon)$-approximation algorithm for the VTP problem (ε is inversely proportional to t). So it suffices to apply Theorem 4.1 with $\mathcal{P}_3 = \mathrm{HS}(3, t)$ and $\rho = \frac{3}{2} + \varepsilon$. ■

Corollary 4.3 *There is a* 1.4*-approximation algorithm for the* $\mathcal{K}_3$*-packing problem on graphs with maximum degree* 4.

Proof. It follows from Theorem 4.1 and the result of [8] showing that there is a ρ-approximation algorithm for the triangle packing problem on graphs with maximum degree 4, where ρ is slightly less than 1.2. ■

A more precise analysis of the greedy algorithm $\mathcal{G}_3$ gives the following result.

Theorem 4.4 *The algorithm* $\mathcal{G}_3$ *is a* 3/2*-approximation for the* $\mathcal{K}_3$*-packing problem. Furthermore, the ratio* 3/2 *is tight.*

Proof. It is similar to the proof of Theorem 2.3. Let $\mathcal{O}$ be an optimal solution and $\mathcal{B}$ be the solution returned by the algorithm $\mathcal{G}_3$. Let t_i, $0 \le i \le 3$, be the number of triangles of $\mathcal{T}_\mathcal{O}$ that intersect precisely i vertices of $\mathcal{T}_\mathcal{B}$. Note that $t_0 = 0$.

Now let $z := |\mathcal{E}_\mathcal{O}| - 3|\mathcal{T}_\mathcal{B}| + t_1 + 2t_2 + 3t_3$. Then at least $\max\{0, z\}$ edges of $\mathcal{E}_\mathcal{O}$ are disjoint from $\mathcal{T}_\mathcal{B}$. Furthermore, every triangle of $\mathcal{T}_\mathcal{O}$ that intersects precisely 1 vertex of $\mathcal{T}_\mathcal{B}$ contributes an edge that is disjoint from $\mathcal{T}_{\mathcal{B}}$. Since $\mathcal{E}_{\mathcal{B}}$ is a maximum matching of $G - \{v\colon v \text{ is a vertex in } \mathcal{T}_{\mathcal{B}}\}$, we have

$$|\mathcal{E}_{\mathcal{B}}| \geq t_1 + \max\{0, z\}. \tag{17}$$

Using the facts that $|\mathcal{T}_{\mathcal{O}}| = t_1 + t_2 + t_3$, we can rewrite z obtaining

$$z = |\mathcal{E}_{\mathcal{O}}| - 3|\mathcal{T}_{\mathcal{B}}| + 2|\mathcal{T}_{\mathcal{O}}| - t_1 + t_3. \tag{18}$$

Now substituting the value of z in the inequality $\mathrm{val}(\mathcal{B}) \geq 3|\mathcal{T}_{\mathcal{B}}| + t_1 + z$, we get

$$\mathrm{val}(\mathcal{B}) \geq 3|\mathcal{T}_{\mathcal{B}}| + t_1 + |\mathcal{E}_{\mathcal{O}}| - 3|\mathcal{T}_{\mathcal{B}}| + 2|\mathcal{T}_{\mathcal{O}}| - t_1 + t_3 \geq 2|\mathcal{T}_{\mathcal{O}}| + |\mathcal{E}_{\mathcal{O}}| \geq \frac{2}{3}\mathrm{val}(\mathcal{O}).$$

To see that the ratio 3/2 of algorithm $\mathcal{G}_3$ is tight, consider the following graph G: it consists of 4 triangles T_0, T_1, T_2, T_3, such that T_1, T_2 and T_3 are pairwise vertex-disjoint and each of them "hangs" in a different vertex of T_o (G has 3 vertices of degree 4 and 6 vertices of degree 2). ■

5 Concluding remarks

The approximation algorithm $\mathcal{C}_3(\rho)$ that we presented for the $\mathcal{K}_3$-packing problem makes use of a routine to find an approximate solution for the VTP problem. From our result, it follows that any improvement on the $(\frac{3}{2} + \varepsilon)$-approximation ratio for the VTP problem would yield an improvement on the approximation ratio for the $\mathcal{K}_3$-packing problem.

Halldórsson [4] presented an algorithm for a version of the minimum 3-set cover problem, with the constraint that the sets found are pairwise disjoint, in addition to forming a cover of the vertices of the input graph. His algorithm is also another approach for the $\mathcal{K}_3$-packing problem. Using the results presented in [4], one can deduce that its approximation ratio is 3/2. This algorithm is however not as simple as the greedy algorithm $\mathcal{G}_3$.

It would be interesting to study the $\mathcal{F}$-packing problem for other families $\mathcal{F}$.

References

[1] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi. *Complexity and approximation*, Springer-Verlag, 1999.

[2] A. Caprara and R. Rizzi. Packing triangles in bounded degree graphs. *Inform. Process. Lett.*, 84(4):175–180, 2002.

[3] F. Chataigner, G. Manic, Y. Wakabayashi and R. Yuster. Approximation algorithms and hardness results for the clique packing problem. In *Proceedings of Eurocomb 2007*, volume 29 of *Electron. Notes Discrete Math.*, pages 397–401 (electronic), Amsterdam, 2007. Elsevier.

[4] M. Halldórsson. Approximating k-set cover and complementary graph coloring. *Lecture Notes in Computer Science*, 1084:118–131, 1996.

[5] P. Hell and D. G. Kirkpatrick. On the complexity of general graph factor problems. *SIAM J. Comput.*, 12(3):601–609, 1983.

[6] P. Hell and D. G. Kirkpatrick. Packings by cliques and by finite families of graphs. *Discrete Math.*, 49(1):45–59, 1984.

[7] C. A. J. Hurkens and A. Schrijver. On the size of systems of sets every t of which have an SDR, with an application to the worst-case ratio of heuristics for packing problems. *SIAM J. Discrete Math.*, 2(1):68–72, 1989.

[8] G. Manić and Y. Wakabayashi. Packing triangles in low degree graphs and indifference graphs. *Discrete Math.*, 308(8):1455–1471, 2008.