Induced Matchings in Regular Graphs and Trees

Michele Zito*

Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, UK

Abstract. This paper studies the complexity of the Maximum Induced Matching problem (MIM) in regular graphs and trees. We show that the largest induced matchings in a regular graph of degree d can be approximated with a performance ratio less than d. However MIM is NP-hard to approximate within some constant $c > 1$ even if the input is restricted to various classes of bounded degree and regular graphs. Finally we describe a simple algorithm providing a linear time optimal solution to MIM if the input graph is a tree.

1 Introduction

If $G = (V, E)$ is a graph, a set $M \subseteq E$ is a *matching* in G if for all $e_1, e_2 \in M$ it is $e_1 \cap e_2 = \emptyset$. Let $V(M)$ be the set of vertices belonging to edges in the matching. A matching M is *maximal* if for every $e \in E \setminus M$, $M \cup e$ is not a matching; M is *induced* if for every edge $e = \{u, v\}$, $e \in M$ if and only if $u, v \in V(M)$ and $e \in E$. Let $\nu_I(G)$ denote the maximum cardinality of an induced matching in G. The maximum induced matching problem (MIM) is that of finding an induced matching in G with $\nu_I(G)$ edges.

The problem was introduced in [14] as a variation of the maximum matching problem and motivated as the "risk-free" marriage problem: find the maximum number of pairs such that each married person is compatible with no married person other than the one he (or she) is married to. Induced matchings have stimulated a lot of interest in discrete mathematics because finding large induced matchings is a subtask of finding a *strong edge-colouring* in a graph (see [5, 6] and [15, 11] for more recent results), a proper colouring of the edges such that no edge is adjacent to two edges of the same colour.

MIM is NP-complete even for bipartite graphs of maximum degree four [14]. One way of coping with the NP-completeness of an optimization problem is to relax the optimality requirement and look for the existence of polynomial time algorithms which guarantee solutions whose size is close to the size of the optimum. In what follows we say that a maximization problem P is approximable with *(performance) ratio* ρ if there is a polynomial time algorithm returning a solution whose size is at least ρ^{-1} times the size of an optimal solution. Not much is known about the approximability of MIM. In Section 2, fairly simple combinatorial arguments allow us to prove the existence of approximation algorithms giving a ratio smaller than d, for regular graphs of degree d. In Section 3, we

* Supported by EPSRC grant GR/L/77089.

establish a number of non-approximability results. We provide explicit bounds on the performance ratio such that MIM is NP-hard to approximate with ratio less than these bounds in several classes of bounded degree graphs including regular graphs of degree four.

A graph $G = (V, E)$ is a *tree* if it is connected and it has no cycle, it is *chordal* if any cycle of at least four vertices contains an edge connecting two non-consecutive vertices. MIM in chordal graphs can be reduced [3] to finding the largest independent set in a chordal graph and the latter problem admits an optimal polynomial time solution [7]. Since trees are chordal graphs, this argument (and an efficient implementation of Gavril's algorithm) implies the existence of a $O(|V|^2)$ algorithm for finding the maximum induced matching in a tree. In Section 4 we present an alternative algorithm which again solves MIM optimally if the input graph is a tree but it runs in $O(|V|)$ time.

2 Combinatorial Bounds

If $G = (V, E)$ is a graph let $\deg_G v$, the *degree of* v, be the number of vertices that are adjacent to v. Let $V_i(G) = \{v \in V : \deg_G v = i\}$ for all $i = 0, \ldots, |V| - 1$. Notations $V(G)$ and $E(G)$ will be used instead of V and E when necessary to prevent ambiguities. A (δ, Δ)-graph is a graph with minimum degree δ and maximum degree Δ. A (d, d)-graph is a regular graph of degree d. Let (δ, Δ)-MIM (resp. d-MIM) identify MIM when the input is restricted to (δ, Δ)-graphs (resp. regular graphs of degree d). In this section we look at positive approximation results for MIM in regular graphs. We describe two results that "come for free" in the sense that they do not require any involved algorithmic idea and their validity is implied by the combinatorial structure of the matching problem under consideration. The negative results in Section 3 show that there is not much scope for better results.

Definition 1. [9] *An* independence system *is a pair* $(E, \mathcal{F})$ *where* E *is a finite set and* $\mathcal{F}$ *a collection of subsets of* E *with the property that whenever* $F \subset H \in \mathcal{F}$ *then* $F \in \mathcal{F}$*. The elements of* $\mathcal{F}$ *are called* independent sets*. A* maximal independent set *is an element of* $\mathcal{F}$ *that is not subset of any other element of* $\mathcal{F}$*.*

Korte and Hausmann [9] analysed the independence system formed by all matchings in G and proved an upper bound of 2 on the ratio between the sizes of any two maximal matchings. In the next result a similar argument is applied to estimate the maximum ratio between two maximal induced matchings.

Let $\mathcal{M}_I(G)$ be the set of all induced matchings in a graph G. The pair $(E(G), \mathcal{M}_I(G))$ is an independence system. For every $S \subseteq E$ the *lower* (resp. *upper*) *rank* of S, $\underline{\rho}(S)$ (resp. $\bar{\rho}(S)$) is the size of the smallest (resp. largest) maximal induced matching included in S. By [9, Theorem 1.1], if M is a maximal induced matching, then

$$\frac{\nu_I(G)}{|M|} \leq \max_{S \subseteq E} \frac{\overline{\rho}(S)}{\underline{\rho}(S)}$$

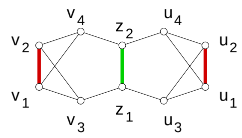

Fig. 1. A cubic graph with large induced matching.

Theorem 1. *Let G be a (δ, Δ)-graph and $(E(G), \mathcal{M}_I(G))$ be given. Then*

$$\max_{S \subseteq E(G)} \frac{\overline{\rho}(S)}{\underline{\rho}(S)} \leq 2(\Delta - 1).$$

Proof. Let M_1 and M_2 be two maximal induced matchings in G and let $e \in M_2 \setminus M_1$. Clearly $M_1 \cup \{e\} \subseteq E$ and, by the maximality condition, this set is not independent (i.e. it is not an induced matching anymore). Hence there exists $\phi(e) \in M_1$ at distance less than two from e and again since M_2 is maximal and independent $\phi(e) \in M_1 \setminus M_2$. Indeed ϕ defines a function from $M_2 \setminus M_1$ to $M_1 \setminus M_2$. Let f be one of the edges in the range of ϕ. A bound on the number of edges $e \in M_2 \setminus M_1$ that can be the pre-image of $f \in M_1 \setminus M_2$ is needed. There can be at most $2(\Delta - 1)$ such e. The result follows. $\square$

The last result gives a bound on the ratio of any algorithm that construct a maximal induced matching in a given graph.

Theorem 2. $(1, \Delta)$-MIM *can be approximated with ratio* $2(\Delta - 1)$.

This result can be slightly improved on regular graphs.

Theorem 3. *If G is a (δ, Δ)-graph then $\nu_I(G) \leq \frac{\Delta|V(G)|}{2(\Delta+\delta-1)}$. Moreover for every $d \geq 3$, with d odd, there exists a regular graph of degree d, with $2(2d-1)$ vertices and a maximum induced matching of size d.*

Proof. Let G be a (δ, Δ)-graph and M be a maximal induced matching in G. Let $R = V \setminus V(M)$. Each $v \in V(M)$ is adjacent to at least $\delta - 1$ vertices in R. Each $v \in R$ can be adjacent to at most Δ vertices in $V(M)$. Hence $(|V(G)| - 2|M|)\Delta \geq 2|M|(\delta - 1)$ and the result follows.

The second part can be proved by giving a recursive description of a family of graphs $\{G_i\}_{i \in \mathbb{N}^+}$. It is convenient to draw G_i so that all its vertices are on five different layers, called *far-left*, *mid-left*, *central*, *mid-right* and *far-right* layer. Figure 1 shows G_1. Vertices v_1 and v_2 (respectively u_1 and u_2) are in the far-left (respectively far-right) layer. Vertices v_3 and v_4 (respectively u_3 and u_4) are in the mid-left (respectively mid-right) layer. Vertices z_1 and z_2 are in the central layer. Moreover an horizontal axis separates odd-indexed vertices (which are below it) from even-indexed ones (which are above), with smaller indexes below higher ones.

Let G_{i-1}, for $i \geq 2$, be given. The graph G_i is obtained by adding four central vertices, two mid-left and two mid-right vertices. Since G_1 has two

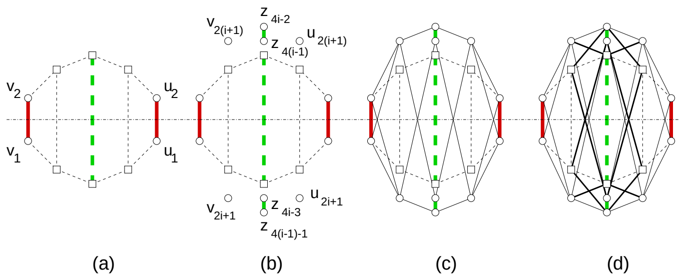

Fig. 2. A regular graph of degree d with large induced matching.

central and two pairs of mid vertices and easy inductions proves that G_{i-1} has $2[2(i-1)-1]$ central vertices and $2(i-1)$ mid-left and mid-right ones. Let $z_{4(i-1)-1}, z_{4(i-1)}, z_{4i-3}, z_{4i-2}$ be the four "new" central vertices, v_{2i+1} and $v_{2(i+1)}$, u_{2i+1} and $u_{2(i+1)}$ the mid-left and mid-right ones. G_i has all edges of G_{i-1} plus the following groups:

1. two edges connecting each of v_{2i+1}, $v_{2(i+1)}$, (respectively u_{2i+1} and $u_{2(i+1)}$) to v_1 and v_2 (respectively u_1 and u_2), plus edges

$$\{v_{2i+1}, z_{4(i-1)-1}\}, \{v_{2i+1}, z_{4(i-1)}\}, \qquad \{v_{2(i+1)}, z_{4i-3}\}, \{v_{2(i+1)}, z_{4i-2}\},$$

$$\{u_{2i+1}, z_{4(i-1)-1}\}, \{u_{2i+1}, z_{4(i-1)}\}, \qquad \{u_{2(i+1)}, z_{4i-3}\}, \{u_{2(i+1)}, z_{4i-2}\}$$

   All these edges are the continuous black lines in Figure 2.(c).
2. A final set of edges connects each of the even index mid vertices with the central vertices of G_{i-1} with indices $4j-2$ and $4j-3$ for $j = 0, 1, \ldots, i-1$. Each of the odd index mid vertices are connected with the central vertices of G_{i-1} with indices $4(j-1)$ and $4(j-1)-1$ for $j = 1, \ldots, i$. The squares in Figure 2 represent all mid vertices in G_{i-1}. The bold solid lines in Figure 2.(d) represent this kind of edges.

Graph G_1 has an induced matching of size three. For each $i \geq 2$ the matching in G_i is obtained by adding the two edges $\{z_{4i-2}, z_{4(i-1)}\}$ and $\{z_{4i-3}, z_{4(i-1)-1}\}$ to the matching in G_{i-1}. $\square$

Theorem 3 is complemented by the following result, giving a lower bound on the size of a particular family of induced matchings.

Theorem 4. *Let $f(\delta, \Delta) = (4\Delta^2 - 4\Delta + 2)/\delta$. If G is a (δ, Δ)-graph then $\nu_I(G) \geq |V(G)|/f(\delta, \Delta)$. Moreover for every $d \geq 2$ there exists a regular graph of degree d with $d \cdot f(d, d)$ vertices and a maximal induced matching of size d.*

Proof. Let G be a (δ, Δ)-graph and M a maximal induced matching in G. Each edge in G must be covered by at least an edge in M. Conversely every edge in M can cover at most $2(\Delta-1)^2 + 2\Delta - 1$ edges. Thus

$$|M| \geq \frac{|E(G)|}{2(\Delta-1)^2 + 2\Delta - 1} \geq \frac{\delta |V(G)|}{2(2\Delta^2 - 2\Delta + 1)}.$$

A *d-ary depth two tree* T_d is formed by connecting with an edge the roots of two identical copies of a complete d-ary tree on d^2-d+1 vertices. The graph obtained by taking d copies of T_d all sharing the same set of $(d-1)^2$ leaves is regular of degree d, it has $d \cdot f(d,d)$ vertices and a maximal induced matching of size d. □

Corollary 1. d-MIM *can be approximated with ratio* $d-(d-1)/(2d-1)$.

Proof. Let G be a regular graph of degree d. The proof follows from Theorem 3 and Theorem 4 and the use of any greedy algorithm that returns a maximal induced matching in G. □

3 Hardness of Approximation

In this section we investigate the non-approximability of MIM for various classes of bounded degree graphs. Although several notions of approximation preserving reductions have been proposed (see for example [4]) the L-reduction defined in [13] is perhaps the easiest one to use. Let P be an optimization problem. For every instance x of P, and every solution y of x, let $c_P(x,y)$ be the *cost* of the solution y. Let $opt_P(x)$ be the cost of an optimal solution.

Definition 2. *Let P and Q be two optimization problems. An* L-reduction *from P to Q is a four-tuple $(t_1, t_2, \alpha, \beta)$ where t_1 and t_2 are polynomial time computable functions and α and β are positive constants with the following properties:*

(1) t_1 maps instances of P to instances of Q and for every instance x of P, $opt_Q(t_1(x)) \leq \alpha \cdot opt_P(x)$.

(2) for every instance x of P, t_2 maps pairs $(t_1(x), y')$ (where y' is a solution of $t_1(x)$) to a solution y of x so that

$$|opt_P(x) - c_P(x, t_2(t_1(x), y'))| \leq \beta |opt_Q(t_1(x)) - c_Q(t_1(x), y')|.$$

Theorem 5. *If P and Q are two maximization problems, there is an L-reduction from P to Q with parameters α and β, and it is NP-hard to approximate P with ratio c then it is NP-hard to approximate Q with ratio $\frac{\alpha\beta c}{(\alpha\beta-1)c+1}$.*

Proof. The result is derived from [12, Proposition 13.2]. Suppose by contradiction that there is an algorithm which approximates Q with ratio $\frac{\alpha\beta c}{(\alpha\beta-1)c+1}$. For every instance x of P let y' be the result of applying this algorithm to $t_1(x)$. Then, by definition of L-reduction,

$$\frac{opt_P(x) - c_P(x, t_2(t_1(x), y'))}{opt_P(x)} \leq \alpha\beta \frac{opt_Q(t_1(x)) - c_Q(t_1(x), y')}{opt_Q(t_1(x))}$$

By definition of performance ratio it is $\frac{opt_Q(t_1(x))}{c_Q(t_1(x), y')} \leq \frac{\alpha\beta c}{(\alpha\beta-1)c+1}$, therefore

$$\frac{opt_Q(t_1(x)) - c_Q(t_1(x), y')}{opt_Q(t_1(x))} \leq 1 - \frac{(\alpha\beta-1)c+1}{\alpha\beta c} = \frac{1}{\alpha\beta}\left(1 - \frac{1}{c}\right)$$

and the result follows. □

Let MIS denote the problem of finding a largest independent set in a graph (problem GT20 in [8]). Appellations (δ, Δ)-MIS and d-MIS are defined in the obvious way. There is [10] a very simple L-reduction from MIS to MIM with parameters $\alpha = \beta = 1$. Given a graph $G = (V, E)$, define $t_1(G) = (V', E')$ as follows:

$$V' = V \cup \{v' : v \in V\}, \quad E' = E \cup \{\{v, v'\} : v \in V\}.$$

If U is an independent set in G then $F = \{\{v, v'\} : v \in U\}$ is an induced matching in $t_1(G)$. Conversely if F is an induced matching in $t_1(G)$ the set $t_2(t_1(G), F)$ obtained by picking one endpoint from every edge in F is an independent set in G. Therefore the size of the largest independent set in G is $\nu_I(t_1(G))$.

The *s-padding of* a (δ, Δ)-graph G, G_s, is obtained by replacing every vertex v by a distinct set of vertices $v_1, \ldots, v_s$ with $\{v_i, u_j\} \in E(G_s)$ if and only if $\{u, v\} \in E(G)$. The following result is a consequence of the definition.

Lemma 1. *For any $s \geq 2$, if G is a (δ, Δ)-graph then G_s is a $(s \cdot \delta, s \cdot \Delta)$-graph.*

The key property of the s-padding of a graph is that it preserves the distance between two vertices. If $G = (V, E)$ is a graph then for every $u, v \in V(G)$, $dst_G(u, v)$ is the *distance* between u and v, defined as the number of edges in a shortest path between u and v.

Lemma 2. *For all graphs G and for every $s \geq 2$, $dst_G(u, v) = dst_{G_s}(u_i, v_j)$ for all $u, v \in V(G)$ with $u \neq v$ and all $i, j \in \{1, 2, \ldots, s\}$.*

Lemma 3. *For all graphs G and for every $s \geq 2$, $\nu_I(G) = \nu_I(G_s)$.*

Proof. Let M be an induced matching in G. Define $M_s = \{\{u_1, v_1\} \in E(G_s) : \{u, v\} \in M\}$. By Lemma 2 all edges in M_s are at distance at least two. Conversely if M_s is an induced matching in G_s define $M = \{\{u, v\} \in E(G) : \{u_i, v_j\} \in M_s$ for some $i, j \in \{1, \ldots, s\}\}$. M is an induced matching in G. $\square$

The following Lemmas show how to remove vertices of degree one, two and three from a $(1, \Delta)$-graph.

Lemma 4. *Any $(1, \Delta)$-graph G can be transformed in polynomial time into a $(2, \Delta)$-graph G' such that $|V_1(G)| = \nu_I(G') - \nu_I(G)$.*

Proof. Given a $(1, \Delta)$-graph G, the graph G' is obtained by replacing each vertex v of degree one in G by the gadget G_v shown in Figure 3. The edge $\{v, w\}$ incident to v is attached to v_0. The resulting graph has minimum degree two and maximum degree Δ. If M is an induced matching in G it is easy to build an induced matching in G' of size $|M| + |V_1(G)|$. Conversely every induced matching

Fig. 3. Gadgets replacing vertices of degree one and two.

M' in G' will contain exactly one edge from every gadget G_v. Replacing (if necessary) each of these edges by the edge $\{v_1, v_2\}$ could only result in a larger matching. The matching obtained by forgetting the gadget-edges is an induced matching in G and its size is (at least) $|M'| - |V_1(G)|$. □

Lemma 5. *Any* $(2, \Delta)$*-graph* G *can be transformed in polynomial time into a* $(3, \Delta)$*-graph* G' *such that* $|V_2(G)| = \nu_I(G') - \nu_I(G)$.

Proof. Let G be a $(2, \Delta)$-graph. Every vertex w of degree two is replaced by the graph G_w in Figure 3. The two edges $\{u, w\}$ and $\{v, w\}$ adjacent to w are replaced by edges $\{u, w_1\}$ and $\{v, w_2\}$. Let G' be the resulting $(3, \Delta)$-graph. If M is a maximal induced matching in G, a matching M' in G' is obtained by taking all edges in M and adding one edge from each of the graphs G_w. Figure 4 shows all the relevant cases. If $w \in V(M)$ then without loss of generality we can assume that $w_1 \in V(M')$ and one of the two edges adjacent to w_2 can be added to M'. If $w \notin V(M)$ then any of the four central edges in G_w can be added to M'. After these replacements no vertex in the original graph gets any closer to an edge in the matching. Inequality $\nu_I(G') \geq \nu_I(G) + |V_2(G)|$ follows from the argument above applied to a maximum induced matching in G.

Conversely for any induced matching M' in G' at most one edge from each copy of G_w belongs to M'. The copies of G_w with $M' \cap E(G_w) = \emptyset$ are called *empty*, all others are called *full*. Inequality $\nu_I(G) \geq \nu_I(G') - |V_2(G)|$ is proved by the following claims applied to a maximum induced matching in G'.

Claim 1 *Any maximal induced matching* M' *in* G' *can be transformed into another induced matching* M'' *in* G' *with* $|M'| \leq |M''|$ *and such that all gadgets in* M'' *are full.*

Claim 2 $M =_{df} M'' \cap E(G)$ *is an induced matching in* G.

To prove the first claim, an algorithm is described which, given an induced matching $M' \subseteq E(G')$, fills all empty gadgets in M'. The algorithm visits in turn all gadgets in G' that have been created by the reduction and performs the following steps:

(1) If the gadget G_w under consideration is empty some local replacements are performed that fill G_w.
(2) The gadget G_w is then marked as "checked".

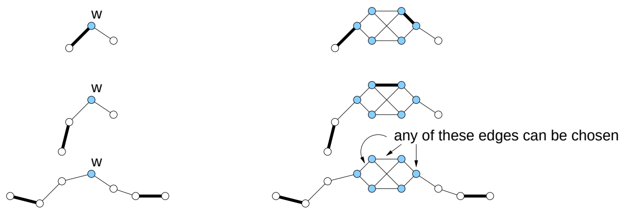

Fig. 4. Possible ways to define the matching in G' given the one in G.

Fig. 5. Filling an empty gadget, normal cases.

(3) A *maximality restoration* phase is performed in which, as a consequence of the local replacements in Step (1), some edges might be added to the induced matching.

Initially all gadgets are "unchecked". Let G_w be an unchecked gadget. If G_w is full the algorithm simply marks it as checked and carries on to the next gadget. Otherwise, since M' is maximal, at least one of the two edges adjacent to vertices w_1 and w_2 must be in M' for otherwise it would be possible to extend M' by picking any of the four central edges in G_w. Without loss of generality let $\{u, w_1\} \in M'$. Figure 5 shows all possible cases. If vertex v does not belong to another gadget then either of the configurations on the left of Figure 5 is replaced by the one shown on the right. If v is part of another gadget few subcases need to be considered. Figure 6 shows all possible cases and the replacement rule. In all cases after the replacement the neighbouring gadget is marked as checked. Notice that all replacement rules do not decrease the size of the induced matching. Also as the process goes by, new edges in $E(G)$ can only be added to the current matching during the maximality restoration phase. To prove the second claim, assume by contradiction that two edges $e = \{u, v\}$ and $f = \{w, y\}$ in M are at distance one. Notice that $dst_{G'}(e, f) = dst_G(e, f)$ unless all the shortest paths beween them contain a vertex of degree two. The existence of e and f is contradicted by the fact that M' and M'' are induced matchings in G' and all gadgets in G' are filled by M''. $\square$

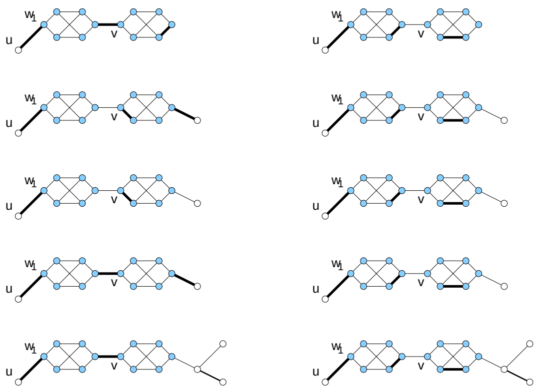

Fig. 6. Filling an empty gadget, special cases.

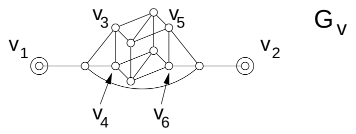

Fig. 7. Gadget connecting pairs of vertices of degree three.

Lemma 6. *Any* $(3, \Delta)$*-graph* G *can be transformed in polynomial time into a* $(4, \Delta)$*-graph* G' *such that* $|V_3(G)| = \frac{1}{2}\nu_I(G') - \nu_I(G)$.

Proof. Let G be a $(3, \Delta)$-graph. The graph G' is defined by taking two copies of G and connecting pairs of corresponding vertices of degree three with the gadget shown in Figure 7. The given gadget has the following important properties:

1. All maximal induced matchings in G_v contain exactly two edges.
2. There exists an induced matching M_v^* in G_v such that neither v_1 nor v_2 are adjacent to a vertex in $V(M)$.

If M is an induced matching in G then the union of two copies of M (one in each copy of G) and a copy of M^*v for each $v \in V_3(G)$ is an induced matching in G'. Conversely given a matching M' in G' the replacement of $M' \cap E(G_v)$ by $\{\{v_3, v_4\}, \{v_5, v_6\}\}$ for every $v \in V_3(G)$ can only lead to a possibly larger induced matching. □

The non-approximability of $(ks, (\Delta+1)s)$-MaxIndMatch (for $k = 1, 2, 3, 4$) and $4s$-MaxIndMatch follows from Theorem 5 applied to known results on independent set [1, 2].

Theorem 6. *Let* $h(\delta, \Delta, c) = \frac{(1+\lfloor \delta/3 \rfloor)[f(\delta,\Delta)+1]c}{f(\delta,\Delta)c+1}$. *Define*

$$g(0, \Delta, c) = c$$
$$g(i, \Delta, c) = h(i, \Delta, g(i-1, \Delta, c)) \qquad i \geq 1$$

For every $\Delta \geq 3$*, let* c_Δ *be a constant such that it is NP-hard to approximate* $(1, \Delta)$-MIS *with ratio* c_Δ*. Then for* $k = 1, 2, 3, 4$ *and every integer* $s > 0$ *it is NP-hard to approximate* $(ks, (\Delta+1)s)$-MaxIndMatch *with ratio* $g(k-1, \Delta+1, c_\Delta)$.

Proof. The result for $k = 1$ follows from the L-reduction at the beginning of the section for $s = 1$ and a further L-reduction based on s-paddings for $s \geq 2$. For $k \in \{2, 3, 4\}$ If G has minimum degree $k-1$, Theorem 4 implies $\nu_I(G) \geq |V_{k-1}(G)|/f(k-1, \Delta)$. The result follows using these bounds along with the reductions in Lemma 4, 5, and 6. □

Theorem 7. *Let* c_0 *be a constant such that* 3-MIS *is NP-hard to approximate with ratio* c_0*. Then for every integer* $s > 0$ *it is NP-hard to approximate* $4s$-MaxIndMatch *with ratio* $\frac{7740c_0}{7685c_0+55}$.

Proof. The reduction at the beginning of Section 3 and Theorem 5 imply that it is NP-hard to approximate $(1,4)$-MAXINDMATCH with ratio c_0. If the original cubic graph G has n vertices, then $t_1(G)$ has $|V_1(t_1(G))| = |V_4(t_1(G))| = n$, no vertex of degree two or three, $5n/2$ edges and the maximum number of edges at distance at most one from a given edge is 19. We call one such graph a *special* $(1,4)$-graph.

Claim 3 There is an L-reduction from $(1,4)$-MAXINDMATCH restricted to special $(1,4)$-graphs to $(3,4)$-MAXINDMATCH with parameters $\alpha = \frac{43}{5}$ and $\beta = 1$.

If G is a $(1,4)$-graph with $|V_2(G)| = |V_3(G)| = 0$, then replacing each vertex v of degree one with the gadget in Figure 3, gives a $(3,4)$-graph G'. The properties of special $(1,4)$-graphs and the same argument used to prove Theorem 4 imply $\nu_I(G) \geq \frac{5}{38}|V_1(G)|$. Therefore

$$\nu_I(G') = \nu_I(G) + |V_1(G)| \leq \nu_I(G) + \frac{38}{5}\nu_I(G) = \frac{43}{5}\nu_I(G)$$

Also, for every matching M' in G', define $t_2(G', M')$ as described in Lemma 4. It follows that $\nu_I(G) - |t_2(G', M')| \leq \nu_I(G') - |M'|$ and the claim is proved.

Therefore, by Theorem 5, $(3,4)$-MAXINDMATCH is hard to approximate with ratio $c_1 = \frac{43c_0}{38c_0+5}$. The *special* $(3,4)$-graphs H generated by the last reduction have again a lot of structure. In particular, $|V_3(H)| = |V_4(H)|$, $|E(H)| = 7|V_3(H)|/2$ and again the maximum number of edges at distance at most one from a given edge is 23.

Claim 4 There is an L-reduction from $(3,4)$-MAXINDMATCH restricted to special $(3,4)$-graphs to 4-MAXINDMATCH with parameters $\alpha = \frac{91}{11}$ and $\beta = 1$.

The reduction was described in Lemma 6. Theorem 4 and the properties of special $(3,4)$-graphs imply $\nu_I(H) \geq \frac{11}{69}|V_3(H)|$. Therefore

$$\nu_I(H') \leq \frac{180}{11}\nu_I(H)$$

and thus, by Theorem 5, 4-MAXINDMATCH is hard to approximate with ratio $\frac{180c_1}{169c_1+11}$. Finally by Lemma 3 there is an L-reduction from 4-MAXINDMATCH to $4s$-MAXINDMATCH (for $s \geq 2$) with parameters $\alpha = \beta = 1$. □

4 Polynomial Time Solution on Trees

Although NP-complete for several classes of graphs including planar or bipartite graphs of maximum degree four and regular graphs of degree four, the problem of finding the largest induced matching admits a polynomial time solution on trees [3]. The algorithmic approach suggested by Cameron reduces the problem to that of finding the largest independent set in a graph H that can be defined starting from the given tree. If $G = (V,E)$ is a tree, the graph $H = (W,F)$ has $|V|-1$ vertices, one for each edge in G and there is an edge between two members of W if and only if the two original edges in G are either incident or connected by a single edge. Notice that $|F| = O(|V|^2)$. Moreover each induced matching in G is an independent set of the same size in H. Gavril's algorithm [7] finds the largest independent set in a chordal graph with n vertices and m edges in $O(n+m)$ time. Since the graph H is chordal, the largest induced matching in the tree can be found in $O(|V|^2)$ time. In this section we describe a simpler and more efficient way of finding a maximum induced matching in a tree. If $G = (V, E)$ is a tree we choose a particular vertex $r \in V$ to be the *root* of the tree (and we say that G is *rooted* at r). If $v \in V \setminus \{r\}$ then parent(v) is the unique neighbour of v in the path from v to r; if parent$(v) \neq r$ then grandparent$(v) =$ parent(parent(v)). In all other cases parent and grandparent are not defined. If $u =$ parent(v) then v is u's *child*. All children of the same node are *siblings* of each other. Let $c(v)$ be the number of children of node v. The *upper neighbourhood* of v (in symbols UN(v)) is empty if $v = r$, it includes r and all v's siblings if v is a child of r and it includes v's siblings, v's parent and v's grandparent otherwise. $E(\mathrm{UN}(v))$ is the set of edges in G connecting the vertices in UN(v).

Claim 5 *If $G = (V, E)$ is a tree and M is an induced matching in G then $|M \cap E(\mathrm{UN}(v))| \leq 1$, for every $v \in V$.*

Note that if M is an induced matching in G, any node v in the tree belongs to one of the following types with respect to the set of edges $E(\mathrm{UN}(v))$:

Type 1. the node $\{v, \mathrm{parent}(v)\}$ is part of the matching,
Type 2. either $\{\mathrm{parent}(v), \mathrm{grandparent}(v)\}$ or $\{\mathrm{parent}(v), w\}$ (where w is some siblings of v) belongs to the matching,
Type 3. Neither **Type 1.** nor **Type 2.** applies.

The algorithm for finding the largest induced matching in a tree G with n vertices handles an $n \times 3$ matrix Value such that Value$[i, t]$ is the size of the matching in the subtree rooted at i if vertex i is of type t.

Lemma 7. *If G is a tree with n vertices,* Value$[i, t]$ *can be computed in $O(n)$ time for every $i \in \{1, \ldots, n\}$ and $t = 1, 2, 3$.*

Proof. Let G be a tree with n vertices. We assume G is in adjacency list representation. If h is the height of the tree, some linear preprocessing is needed to define an array level$[i]$ (for $i = 0, \ldots, h$) such that level$[i]$ contains all vertices at distance i from the root.

The matrix Value can be filled in a bottom-up fashion starting from the deepest vertices of G. If i is a leaf of G then Value$[i, t] = 0$ for $t = 1, 2, 3$. In filling the entry corresponding to node $i \in V$ of type t we only need to consider the entries for all children of i.

(1) Value$[i, 1] = \sum_{k=1}^{c(i)}$ Value$[j_k, 2]$. Since $\{i, \mathrm{parent}(i)\}$ will be part of the matching, we cannot pick any edge from i to one of its children. The matching for the tree rooted at i is just the union of the matchings of the subtrees rooted at each of i's children.

(2) Value$[i, 2] = \sum_{k=1}^{c(i)}$ Value$[j_k, 3]$. We cannot pick any edge from i to one of its children here either.

(3) If node i has $c(i)$ children then Value$[i, 3]$ is the maximum between $\sum_{k=1}^{c(i)}$ Value$[j_k, 3]$ and a number of terms

$$s_{j_k} = 1 + \text{Value}[j_k, 1] + \sum_{l \neq k} \text{Value}[j_l, 2]$$

If the upper neighbourhood of i is unmatched we can either combine the matchings in the subtrees rooted at each of i's children (assuming these children are of type 3) or add to the matching an edge from i to one of its children j_k (the one that maximises s_{j_k}) and complete the matching for the subtree rooted at i with the matching for the subtree rooted at j_k (assuming j_k is of type 1) and that of the subtrees rooted at each of i's other children (assuming these children are of type 3).

Option (3) above is the most expensive involving the maximum over a number of sums equal to the degree of the vertex under consideration. Since the sum of the degrees in a tree is linear in the number of vertices the whole table can be computed in linear time. □

Theorem 8. MIM *can be solved optimally in polynomial time if G is a tree.*

Proof. The largest between Value$[r, 1]$, Value$[r, 2]$ and Value$[r, 3]$ is the size of the largest matching in G. By using appropriate date structures it is also possible to store the actual matching. The complexity of the whole process is $O(n)$. □

5 Conclusions

In this paper we investigated the complexity of finding a largest induced matching in a graph. We suggested a couple of simple heuristics for solving the problem with $O(\Delta)$-bounded ratio on graphs with maximum degree Δ. The definition of algorithms achieving better performance ratios is an open problem. In Section 3 we complemented these positive results with a number of non-approximability results. In particular there is a constant c such that it is NP-hard to find induced matchings whose size approximates $\nu_I(G)$ with ratio c even if G is regular of degree $4s$ for any integer $s > 0$. We believe that similar results hold for cubic graphs and, using the s-padding technique described in Section 3, for regular graphs of degree $3s$ for any integer $s > 0$. Proving hardness on regular graphs of degree d for every integer $d > 2$ maybe harder. Finally we presented an algorithm which solves MIM optimally if the input graph is a tree. Our algorithm is simple in that it does not reduce the original problem to another one, its complexity improves the one of the best algorithm known and it is clearly optimal in the sense that to define an induced matching in a tree $\Omega(n)$ operations are needed.

References

1. N. Alon, U. Feige, A. Wigderson, and D. Zuckerman. Derandomized Graph Products. *Computational Complexity*, 5:60–75, 1995.

2. P. Berman and M. Karpinski. On Some Tighter Inapproximability Results. In *Proc. 26th I.C.A.L.P.*, Springer - Verlag, 1999.
3. K. Cameron. Induced Matchings. *Discr. Applied Math.*, 24(1-3):97–102, 1989.
4. P. Crescenzi. A Short Guide to Approximation Preserving Reductions. In *Proc. 12th Conf. on Comput. Complexity*, pages 262–273, Ulm, 1997.
5. P. Erdős. Problems and Results in Combinatorial Analysis and Graph Theory. *Discr. Math.*, 72:81–92, 1988.
6. R. J. Faudree, A. Gyárfas, R. H. Schelp, and Z. Tuza. Induced Matchings in Bipartite Graphs. *Discr. Math.*, 78(1-2):83–87, 1989.
7. F. Gavril. Algorithms for Minimum Coloring, Maximum Clique, Minimum Covering by Cliques and Maximum Independent Set of a Chordal Graph. *SIAM J. Comp.*, 1(2):180–187, June 1972.
8. M. R. Garey and D. S. Johnson. *Computer and Intractability, a Guide to the Theory of NP-Completeness.* Freeman and Company, 1979.
9. B. Korte and D. Hausmann. An Analysis of the Greedy Heuristic for Independence Systems. *Ann. Discr. Math.*, 2:65–74, 1978.
10. S. Khanna and S. Muthukrishnan. Personal communication.
11. J. Liu and H. Zhou. Maximum Induced Matchings in Graphs. *Discr. Math.*, 170:277–281, 1997.
12. C. Papadimitriou. *Computational Complexity.* Addison-Wesley, 1994.
13. C. H. Papadimitriou and M. Yannakakis. Optimization, Approximation and Complexity Classes. *J. Comp. Sys. Sciences*, 43:425–440, 1991.
14. L. J. Stockmeyer and V. V. Vazirani. NP-Completeness of Some Generalizations of the Maximum Matching Problem. *I.P.L.*, 15(1):14–19, August 1982.
15. A. Steger and M. Yu. On Induced Matchings. *Discr. Math.*, 120:291–295, 1993.