Egerváry Research Group
on Combinatorial Optimization

Technical reportS

TR-2001-10. Published by the Egrervary Research Group, Pázmány P. sétány 1/C, H–1117, Budapest, Hungary. Web site: www.cs.elte.hu/egres . ISSN 1587–4451.

Restricted t-matchings in bipartite graphs

András Frank

March 2001

Restricted t-matchings in bipartite graphs

András Frank*

Abstract

Given a simple bipartite graph G and an integer $t \geq 2$, we derive a formula for the maximum number of edges in a subgraph H of G so that H contains no node of degree larger than t and H contains no complete bipartite graph $K_{t,t}$ as a subgraph. In the special case $t = 2$ this fomula was proved earlier by Z. Király [6], sharpening a result of D. Hartvigsen [4]. For any integer $t \geq 2$, we also determine the maximum number of edges in a subgraph of G that contains no complete bipartite graph, as a subgraph, with more than t nodes. The proofs are based on a general min-max result of [2] concerning crossing bi-supermodular functions.

1 Introduction

Throughout the paper we work with a bipartite graph $G = (S, T; E)$ with node set $V := S \cup T$. We will always assume, without any further reference, that G is simple. Let $t \geq 2$ be an integer. By a t-**matching** (resp., t-**factor**) we mean a subgraph H of G in which the degree $d_H(v)$ of every node v is at most t (resp., is exactly t). Note that in the literature such a subgraph is sometimes called a simple t-matching and in a t-matching multiple copies of an edge are also allowed. Since we never use multiple edges and work exclusively with subgraphs of G, the adjective "simple" will not be used.

For a subset $Z \subseteq V$, the number of edges induced by Z is denoted by $i_G(Z) = i(Z)$, while the number of edges with at least one end-node in Z is denoted by $e_G(Z) = e(Z)$. It is known from (bipartite) matching theory that G has a t-factor if and only if $|S| = |T|$ and

$$t|S| \leq t|Y| + i(V - Y) \tag{1}$$

holds for every subset $Y \subseteq V$ (and actually, it is enough to assume (1) only for subsets Y for which $V - Y$ induces a graph of maximum degree at most $t-1$). More generally, the maximum number of edges in a t-matching of G is equal to

$$\min_{Z \subseteq V} (t|Z| + i(V - Z)). \tag{2}$$

*The work was completed while the author visited the Institute for Discrete Mathematics, University of Bonn, July, 2000. Supported by the Hungarian National Foundation for Scientific Research, OTKA T029772. Department of Operations Research, Eötvös University, Kecskeméti u. 10-12, Budapest, Hungary, H-1053 and Traffic Lab Ericsson Hungary, Laborc u.1, Budapest, Hungary H-1037. e-mail: `frank@cs.elte.hu`

W.H. Cunningham and J. Geelen proposed to investigate the problem of maximum C_4-free (or square-free) 2-matchings of bipartite graphs where C_4 or a square is a circuit of length four. In a recent paper, D. Hartvigsen [4] provided an answer to this problem in the following sense. He introduced a linear program (P'), as follows.

$$\max x(E)$$
$$0 \leq x \leq 1$$
$$d_x(v) \leq 2 \text{ for every } v \in V, \tag{3}$$
$$x(K) \leq 3 \text{ for every 4-circuit } K \text{ of } G \tag{4}$$

where $d_x(v) := \sum(x(e) : e \in E, e \text{ is incident with } v)$.

Clearly, a $0-1$ vector satisfies these constraints if and only if it is the characteristic vector of a square-free 2-matching of G. Hartvigsen [4] announced the following result.

Theorem 1.1. *[4] The linear program (P') has an otpimum solution which is $0-1$-valued. The dual linear program to (P') has an optimum solution which is half-integer-valued.*

To prove this, Hartvigsen constructed a (combinatorial) strongly polynomial algorithm that computes a $0-1$-valued primal solution along with a half-integer-valued dual solution so that this pair of solutions satisfies the complementary slackness condition. This immediately gives rise to a min-max formula on the maximum cardinality of a square-free 2-matching of G but Hartvigsen did not explicitly mention this, only gave he a characterization for the existence of a square-free 2-factor. (That paper is an extended abstract and does not include the detailed algorithm and its proof.)

That formulation however was not completely correct as was pointed out by Z. Király in an unpublished manuscript around September, 1999. Király not only corrected Hartvigsen's characterization but proved a stronger result asserting that the linear programming dual to problem (P') has always an integer-valued optimum (and not only half-integer-valued). What actually Király proved was the following.

Theorem 1.2. *[6] The maximum cardinality of a square-free 2-matching in a bipartite graph $G = (S, T; E)$ is equal to*

$$\min_{Z \subseteq V} (2|Z| + i(V - Z) - c_2(Z)) \tag{5}$$

where $c_2(Z)$ is the number of those components of $G - Z$ which are a square. The optimal Z may be chosen in such a way that each component of $G - Z$ consists of a single node, or two adjacent nodes, or a square.

Király's proof is relatively simple though not algorithmic. Having Király's characterization at hand, Hartvigsen was able to revise his algorithm to provide an integer-valued dual optimum, as well. This improved version will appear in his detailed paper [5] (a first draft is available in September, 2000). In December 1999, I noticed a relation of the problem to a result in [2] and this led not only to a third approach to the square-free 2-matching problem but to its extensions, as well. (The present paper contains the details: see theorems 2.1 and 3.3). Having heard of theorem 2.1, Király was able to apply his proof technique and proved that result too. This will appear in [7].

As far as generalizations of theorem 1.2 are concerned, several possibilities show up naturally. For example, one may be interested in finding maximum cardinality 2-matchings not containing circuits of length six. J. Geelen [3] proved the NP-completeness of this problem. As the ordinary min-cost 2-factor problem is tractable through network flows, one may want to find a minimum cost square-free 2-factor of G. However, Z. Király [8] noticed that Geelen's proof can be modified to prove the NP-completeness of this problem. Finally, for higher integers t, the problem of maximum t-matchings not containing certain forbidden subgraphs is worth for investigation, too.

But what kind of forbidden subgraphs are hopeful for good characterizations? Geelen's NP-completeness observation above indicates that forbidding circuits longer than four is not promising. A circuit C_4 of length four, however, may also be considered as a complete bipartite graph $K_{2,2}$, and this fact will prove to be a suitable ground for generalizations. In what follows, $K_{k,l}$ denotes a complete bipartite graph, that is, a graph whose node set is partitioned into a k-element set K and an l-element set L, and its edge set is $\{uv : u \in K, v \in L\}$. When $k \geq 1, l \geq 1$, we speak of a **bi-clique**. The **size** of a bi-clique is the number ($= k + l$) of its nodes. If $k = 1$ or $l = 1$, the bi-clique is called **trivial.** A trivial bi-clique may be called a **star**.

The goal of this note is to exhibit two extensions of the result of Hartvigsen and Király. In the first one we consider t-matchings of G containing no $K_{t,t}$. That is, we are interested in subgraphs containing neither trivial bi-cliques of size $t + 1$ nor bi-cliques $K_{t,t}$. In the other case, subgraphs of G are considered containing no bi-clique of size larger than t for a given integer $t \geq 2$. In both cases, we derive a formula for the maximum number of edges in such subgraphs. The proofs are based on a general min-max theorem [2] concerning positively crossing bi-supermodular functions (which have already found several applications such as node- and edge-connectivity augmentation of directed graphs, directed splitting-off results, extensions of Győri's theorem on intervals). Before presenting the main results, we recall this formula.

Let S and T be two disjoint sets and let $A^* := \{st : s \in S, t \in T\}$ denote the set of all directed edges with tail in S and head in T. Let $\mathcal{A}^* := \{(A, B) : \emptyset \subset A \subseteq S, \emptyset \subset B \subseteq T\}$. The first member A of a pair (A, B) is called its **tail** while the second member B is its **head**. A pair (A, B) is called **trivial** if $|A| = 1$ or $|B| = 1$. We say that a directed edge st **covers** a pair $(A, B) \in \mathcal{A}^*$ if $s \in A, t \in B$. A subset $\mathcal{F}$ of $\mathcal{A}^*$ is called **independent** if no two members of $\mathcal{F}$ can be covered by an element of A^*, which is equivalent to saying that for any two members of $\mathcal{F}$ their heads or their tails are disjoint. Let $p : \mathcal{A}^* \to \mathbf{Z}_+$ be a nonnegative integer-valued function. We say that p is **positively crossing bi-supermodular** if

$$p(X, Y) + p(X', Y') \leq p(X \cap X', Y \cup Y') + p(X \cup X', Y \cap Y') \tag{6}$$

holds whenever $p(X, Y) > 0, p(X', Y') > 0$, $X \cap X' \neq \emptyset, Y \cap Y' \neq \emptyset$. For a subset $\mathcal{F}$ of $\mathcal{A}^*$, let $p(\mathcal{F}) = \sum(p(X, Y) : (X, Y) \in \mathcal{F})$. For a vector $z : A^* \to \mathbf{R}$ and a pair $(X, Y) \in \mathcal{A}^*$ we use the notation $d_z(X, Y) := \sum(z(xy) : x \in X, y \in Y)$. We say that a non-negative vector z on A^* **covers** p or that z is a **covering** of p if $d_z(A,B) \geq p(A,B)$ holds for every member (A,B) of $\mathcal{A}^*$.

Theorem 1.3. *[2] For an integer-valued positively crossing bi-supermodular function p, the following min-max equality holds:* $\min(z(A^*)$: *z an integer-valued covering of p*) $= \max(p(\mathcal{F})$*: $\mathcal{F} \subseteq \mathcal{A}^*, \mathcal{F}$ independent).*

2 $K_{t,t}$-free t-matchings

We say that a t-matching H is $K_{t,t}$**-free** if it contains no $K_{t,t}$ as a subgraph, which is equivalent to saying that no component of H is a $K_{t,t}$. For a subset $Z \subseteq V$, let $c_t(Z)$ denote the number of those components of $G - Z$ which are a $K_{t,t}$.

Theorem 2.1. *The maximum number of edges in a $K_{t,t}$-free t-matching of a bipartite graph $G = (S,T;E)$ is equal to*

$$\gamma := \min_{Z \subseteq V}(t|Z| + i(V - Z) - c_t(Z)). \tag{7}$$

Moreover, it suffices to take the minimum only over those subsets Z of V for which all the non-$K_{t,t}$ components of $G - Z$ induce a $(t-1)$-matching.

Proof. First we show the second part. Let Z be a set minimizing (7) for which $|Z|$ is as large as possible. We show that each non-$K_{t,t}$ component K of $G - Z$ induces a $(t-1)$-matching. Suppose on the contrary that K has a node u which has at least t neighbours in K and let $Z' := Z + u$. Then $i(V - Z') \leq i(V - Z) - t$ and $c_t(Z') \geq c_t(Z)$. Hence $t|Z'| + i(V - Z') - c_t(Z') \leq (t|Z| + t) + (i(V - Z) - t) - c_t(Z') = t|Z| + i(V - Z) - c_t(Z)$, that is, Z' is another minimizer of (7) contradicting the maximum choice of Z.

Next we prove for any subset Z that the cardinality of a $K_{t,t}$-free t-matching H is at most $t|Z| + i(V - Z) - c_t(Z)$. Indeed, since H is a t-matching, the number of edges of H which are incident to a node in Z is at most $t|Z|$. Furthermore, since H is $K_{t,t}$-free, each $K_{t,t}$-component of $G - Z$ has an edge not in H, so the number of edges of H not incident to any element of Z, that is, the edges of H induced by $V - Z$, is at most $i(V - Z) - c_t(Z)$. Hence the total number of edges of H is indeed at most $t|Z| + i(V - Z) - c_t(Z)$.

Finally, we turn to the main content of the theorem and prove that $\max \geq \min$. For a number x, let $x^+ := \max(x, 0)$. Let us define $p : \mathcal{A}^* \to \mathbf{Z}_+$ as follows. $p(A,B) := (|A| + |B| - 2t + 1)^+$ if $A \cup B$ induces a non-trivial bi-clique of G, $p(A,B) := (|A| + |B| - t - 1)^+$ if $A \cup B$ induces a trivial bi-clique in G, and $p(A,B) := 0$ otherwise.

Claim 2.2. *p is positively crossing bi-supermodular.*

Proof. Let (A,B) and (X,Y) be two pairs for which $p(A,B) > 0$ and $p(X,Y) > 0$, $X \cap A \neq \emptyset, Y \cap B \neq \emptyset$. Suppose first that they are non-trivial. Note that if $(A \cap X, B \cup Y)$ is trivial, then, by $t \geq 2$, $(|A \cap X| + |B \cup Y| - 2t + 1)^+ \leq p(A \cap X, B \cup Y)$ and similarly if $(A \cup X, B \cap Y)$ is trivial, then $(|A \cup X| + |B \cap Y| - 2t + 1)^+ \leq p(A \cup X, B \cap Y)$.

Hence $p(A,B)+p(X,Y) = |A|+|B|-2t+1+|X|+|Y|-2t+1 = |A\cap X|+|B\cup Y|-2t+1+|A\cup X|+|B\cap Y|-2t+1 \le (|A\cap X|+|B\cup Y|-2t+1)^+ + (|A\cup X|+|B\cap Y|-2t+1)^+ \le p(A\cap X, B\cup Y)+p(A\cup X, B\cap Y)$. (The last inequality is satisfied with equality if none of $(A\cap X, B\cup Y)$ and $(A\cup X, B\cap Y)$ is trivial.)

When both pairs are trivial, then both $(A\cap X, B\cup Y)$ and $(A\cup X, B\cap Y)$ are trivial, as well, from which (6) follows.

Finally, suppose that one of the two pairs, say (A,B), is trivial while (X,Y) is non-trivial. Then at least one of $(A\cap X, B\cup Y)$ and $(A\cup X, B\cap Y)$ is also trivial and (6) follows again. □

In what follows we will not distinguish in notation between an (undirected) edge of G connecting u and v and a directed edge in A^* with tail u and head v. Both will be denoted by uv. Also, when no ambiguity may arise, we do not distinguish between a one-element set $\{a\}$ and its only element a.

Lemma 2.3. *If $z : A^* \to \mathbf{Z}_+$ is a minimal covering of p, then $z(uv)$ may be positive on an edge $uv \in A^*$ only if $uv \in E$. Morover, z is $0-1$-valued and the edge set $E_z := \{uv \in E : z(uv) = 0\}$ is a $K_{t,t}$-free t-matching.*

Proof. If uv is not an edge of G, then uv does not belong to any bi-clique of G, that is, uv does not cover any pair (A,B) with positive $p(A,B)$ and hence the minimality of z implies $z(uv) = 0$.

Suppose now indirectly that $z(uv) \ge 2$ for some $uv \in A^*$. By the minimality of z, there is a pair $(A,B) \in \mathcal{A}^*$ for which $2 \le d_z(A,B) = p(A,B)$ and $u \in A, v \in B$. We may assume that $|A| \le |B|$. Then $|B| \ge 2$ for otherwise $|A| = |B| = 1$ and then $p(A,B) = (2-t-1)^+ = 0$. Hence $B' := B-v$ is nonempty and $p(A,B') \ge p(A,B)-1$. We have $d_z(A,B') \le d_z(A,B)-z(uv) = p(A,B)-z(uv) \le p(A,B)-2 \le p(A,B')-1$ contradicting the assumption that z covers p.

To show that E_z is a t-matching, assume indirectly that for some node u of G there are $t+1$ edges $uv_1, uv_2, \ldots, uv_{t+1}$ in E_z incident to u. Let $A := \{u\}, B := \{v_1, \ldots, v_{t+1}\}$. Then $p(A,B) = 1$ and $d_z(A,B) = 0$ contradicting that z is a covering of p. Therefore E_z is indeed a t-matching.

Finally, let $A \subseteq S$ and $B \subseteq T$ be subset of nodes so that $|A| = |B| = t$ and $A\cup B$ induces a bi-clique K of G. Then $p(A,B) = (|A|+|B|-2t+1)^+ = 1$, and since z covers p there must be an edge uv for which $u \in A, v \in B$ and $z(uv) = 1$. Therefore K cannot belong to E_z, that is, E_z is $K_{t,t}$-free. □

Let $\mathcal{F}$ be an independent subset of $\mathcal{A}^*$ for which $p(\mathcal{F})$ is maximum, and subject to this, $\mathcal{F}$ has a maximum number of trivial pairs. We collect some properties of trivial and non-trivial members of $\mathcal{F}$. Clearly, $p(A,B) \ge 1$ for $(A,B) \in \mathcal{F}$.

Claim 2.4. *For every node $a \in S$, $\mathcal{F}$ contains at most one (trivial) pair of form (a,B). For every node $b \in T$, $\mathcal{F}$ contains at most one (trivial) pair of form (A,b).*

Proof. By symmetry, it is enough to prove only the first part. If (a,B_1) and (a,B_2) belong to $\mathcal{F}$, then $B_1\cap B_2 = \emptyset$ by the independence of $\mathcal{F}$. Hence $\mathcal{F}' := \mathcal{F} - \{(a,B_1),\ (a,B_2)\} \cup \{(a, B_1\cup B_2)\}$ is also independent. Moreover $p(\mathcal{F}') = p(\mathcal{F}) -$

$(|B_1| - t) - (|B_2| - t) + (|B_1 \cup B_2| - t) = p(\mathcal{F}) + t$, contradicting the maximality of $p(\mathcal{F})$. □

Let $a_1, \ldots, a_k$ be those elements of S for which there are trivial members $(a_1, B_1), \ldots, (a_k, B_k)$ in $\mathcal{F}$ (there may be none). Let $b_1, \ldots, b_l$ be those elements of T for which there are trivial members $(A_1, b_1), \ldots, (A_l, b_l)$ in $\mathcal{F}$ (there may be none). Let $Z_1 := \{a_1, \ldots, a_k, b_1, \ldots, b_l\}$.

Let (A, B) be a non-trivial member of $\mathcal{F}$.

Claim 2.5. $|A| = |B| = t$.

Proof. We may assume that $|A| \leq |B|$. From $0 < p(A, B) = |A| + |B| - 2t + 1$ we have $|A| + |B| \geq 2t$. So if $|A| = |B| = t$ does not hold, then $|B| > t$. As (A, B) is non-trivial, $A' := A - a$ is non-empty for an element a of A. Let $\mathcal{F}' := \mathcal{F} - \{(A, B)\} \cup \{(A', B), (\{a\}, B)\}$. Now $p(a, B) = |B| - t \geq 1$ and $p(A', B) \geq p(A, B) - 1$ and hence $p(a, B) + p(A', B) \geq p(A, B)$, that is, $p(\mathcal{F}') \geq p(\mathcal{F})$ contradicting the extreme choice of $\mathcal{F}$. □

Claim 2.6. $(A \cup B) \cap Z_1 = \emptyset$.

Proof. Suppose indirectly that $(A \cup B) \cap Z_1 \neq \emptyset$. By symmetry we may assume that $A \cap Z_1 \neq \emptyset$ and let $a_i \in A \cap Z_1$. (Recall that (a_i, B_i) is a trivial member of $\mathcal{F}$, and then $|B_i| \geq t+1$). Let $\mathcal{F}' := \mathcal{F} - \{(A, B), (a_i, B_i)\} \cup \{(a_i, B \cup B_i)\}$. Then $\mathcal{F}'$ is independent. We have $p(A, B) = |A| + |B| - 2t + 1 = 1, p(a_i, B_i) = |B_i| - t, p(a_i, B \cup B_i) = |B \cup B_i| - t$ from which $p(A, B) + p(a_i, B_i) = 1 + |B_i| - t < |B_i \cup B| - t = p(a_i, B \cup B_i)$ from which $p(\mathcal{F}') > p(\mathcal{F})$, a contradiction. □

Claim 2.7. *If $(A, B), (A', B')$ are non-trivial members of $\mathcal{F}$, then $(A \cup B) \cap (A' \cup B') = \emptyset$.*

Proof. Suppose indirectly that the intersection is nonempty. By symmetry we may assume that $A \cap A' \neq \emptyset$, in which case $B \cap B' = \emptyset$. Let $a \in A \cap A'$, and $\mathcal{F}' := \mathcal{F} - \{(A, B), (A', B')\} \cup \{(a, B \cup B')\}$. Then $\mathcal{F}'$ is independent. Now $p(A, B) + p(A', B') = 1 + 1 = 2$ and $p(a, B \cup B') = |B \cup B'| - t = 2t - t = t \geq 2$. That is, $p(\mathcal{F}') \geq p(\mathcal{F})$, contradicting the extreme choice of $\mathcal{F}$. □

Claim 2.8. *If $xy \in E$ and $x \in A \cup B$, then $y \in Z_1$.*

Proof. Suppose that $y \notin Z_1$. By symmetry we may assume that $x \in A$. Let $B' := B + y$. Then $\mathcal{F}' := \mathcal{F} - \{(A, B)\} \cup \{(x, B')\}$ is independent. Now $p(A, B) = 1$ and $p(x, B') = |B'| - t = 1$. Hence $p(\mathcal{F}') \geq p(\mathcal{F})$, contradicting the extreme choice of $\mathcal{F}$. □

Let $\mathcal{F}_1 = \{(a_1, B_1), \ldots, (a_k, B_k), (A_1, b_1), \ldots, (A_l, b_l)\}$, that is, $\mathcal{F}_1$ consists of the trivial members of $\mathcal{F}$. Let $\mathcal{F}_2$ consist of the non-trivial members of $\mathcal{F}$. We have shown that members of $\mathcal{F}_2$ are pairwise disjoint $K_{t,t}$-subgraphs which may be connected only to elements of Z_1 $(= \{a_1, \ldots, a_k, b_1, \ldots, b_l\})$. In other words the members of $\mathcal{F}_2$ are components of $G - Z_1$ hence $p(\mathcal{F}_2) = |\mathcal{F}_2| \leq c_t(Z_1)$. Furthermore, since $\mathcal{F}_1$ is independent, we cannot have both $a_i \in A_j$ and $b_j \in B_i$ for $i = 1, \ldots, k,\ j = 1, \ldots, l$. Hence $p(\mathcal{F}_1) = \sum |A_i| + \sum |B_j| - t|Z_1| \leq e(Z_1) - t|Z_1| = |E| - i(V - Z_1) - t|Z_1|$ where $e(Z_1)$ denotes the number of edges of G with at least one end-node in Z_1. By the definition of γ in (7), $\gamma \leq t|Z_1| + i(V - Z_1) - c_t(Z_1)$. The combination of these inequalities gives rise to $p(\mathcal{F}) = p(\mathcal{F}_1) + p(\mathcal{F}_2) \leq |E| - i(V - Z_1) - t|Z_1| + c_t(Z_1) \leq |E| - \gamma$. By Theorem 1.3, $z(E) = p(\mathcal{F}) \leq |E| - \gamma$ for the minimum covering z of p and hence the $K_{t,t}$-free t-matching E_z has cardinality $|E_z| = |E| - z(E) \geq \gamma$, as required. □

3 Subgraphs with no large bi-cliques

In the preceding section we were interested in subgraphs of a bipartite graph $G = (S, T; E)$ not containing two types of bi-cliques: $K_{1,t+1}$ and $K_{t,t}$. Here we want to find for a given integer $t \geq 2$ the maximum number of edges of a subgraph of G that does not contain any bi-clique of size larger than t. Such a bi-clique will be called **large** (with respect to t). (The problem for $t = 1$ is void as it asks for the maximum number of edges in a subgraph with no edges.) For $t = 2$, the problem is to maximize the number of edges in a subgraph containing no to adjacent edges: this is exactly the (ordinary) matching problem solved by Kőnig. When $t = 3$, the problem requires finding a largest square-free 2-matching, Hartvigsen's problem. A subset F of edges of G will be called a **t-covering** if F covers every large bi-clique.

We investigate the following equivalent form of the general case $t \geq 2$: what is the minimum cardinality $\tau_t(G)$ of a t-covering in G? For any bi-clique H let $p_t(H) := (|V(H)| - t)^+$. Clearly, for large bi-cliques $p_t(H) := |V(H)| - t$. Later, for a family $\mathcal{H}$ of bi-cliques, we will use the notation $p_t(\mathcal{H}) := \sum(p_t(X) : X \in \mathcal{H})$. First we make an easy observation.

Claim 3.1. *If H is a bi-clique, then any t-covering F of H has at least $|V(H)| - t$ edges, that is,*

$$\tau_t(H) \geq p_t(H). \tag{8}$$

Proof. The claim is trivial if $|V(H)| \leq t$ so we may assume that $H = K_{k,l}$ is a large bi-clique with $k \leq l$. Let K and L denote the two stable sets of H (with $|K| = k, |L| = l$.)

If $l \geq t$, then, for every element v of K, $L + v$ induces a large bi-clique and therefore at least $l - t + 1$ edges incident to v must be in F. Hence $|F| \geq k(l - t + 1) = (l - t + 1) + (k - 1)(l - t + 1) \geq (l - t + 1) + (k - 1) = k + l - t = p_t(H)$.

In case $l < t$ let K' denote the subset of elements of K which are incident to F. If $K' = K$, then $|F| \geq |K'| = k > k + l - t$. If $K' \subset K$, then $(K - K') \cup L$ induces a bi-clique not covered by F. Hence $|(K-K')\cup L|\leq t$, from which $|F|\geq|K'|\geq k+l-t=p_t(H)$, as required. □

Let us call a bi-clique $K_{k,l}$ **essential** if $2\leq k<t, 2\leq l<t,\ k+l>t$. The following claim is not really required for our discussion but the min-max formula below uses essential bi-cliques and stars hence it may be useful to establish the τ_t-value of these bi-cliques.

Claim 3.2. *If $H=K_{k,l}$ is a trivial bi-clique (that is, a star: $k=1$ or $l=1$) or an essential bi-clique, then (8) holds with equality.*

Proof. As the claim is obvious for stars, we assume that $2\leq k\leq l<t$. By Claim 3.1, we only have to show that there is a t-covering F of H with $|F|=k+l-t$. Since $0<k+l-t<k$, H has a matching F of $k+l-t$ edges. We claim that F covers all large bi-cliques H' of H. Indeed, if H' is not covered by F, then H' may contain at most one end-node of each matching-edge. Hence $|H'|\leq|K|+|L|-|F|=k+l-(k+l-t)=t$, contradicting that H' is large. □

The main result of this section is as follows.

Theorem 3.3. *In a bipartite graph $G=(S,T;E)$ the minimum cardinality $\tau_t=\tau_t(G)$ of a t-covering of bi-cliques ($t\geq 2$) is equal to*

$$\max\{p_t(\mathcal{H}):\mathcal{H}\text{ is a set of pairwise edge-disjoint bi-cliques}\}. \tag{9}$$

The optimal $\mathcal{H}$ may be chosen to consist of stars and essential bi-cliques.

Proof. Let M denote the maximum in (9). By Claim 3.1 any t-covering F contains at least $p_t(H)$ edges from every bi-clique, hence F contains at least $p_t(\mathcal{H})$ edges from the members of edge-disjoint bi-cliques in $\mathcal{H}$. Hence $\tau_t\leq M$ follows.

To see the non-trivial direction, let us define $p:\mathcal{A}^*\to\mathbf{Z}_+$ as follows. $p(A,B):=p_t(A,B)$ if $(A\cup B)$ induces a bi-clique of G and $p(A,B):=0$ otherwise. It is straightforward to see that p is positively crossing bi-supermodular and we can apply Theorem 1.3. It follows immediately from the definition of p and p_t that the maximum in Theorem 1.3 is M. Hence there exists a covering $z:A^*\to\mathbf{Z}_+$ of p for which $z(A^*)=M$.

$z(uv)$ cannot be positive on any edge $uv\in A^*-E$ since such an uv does not belong to any bi-clique of G, that is, uv does not cover any pair (A,B) with positive $p(A,B)$ and hence the minimality of z implies $z(uv)=0$.

We claim that z is $0-1$-valued. Indeed, let indirectly $z(uv)\geq 2$ for some $uv\in A^*$. By the minimality of z, there is a pair $(A,B)\in\mathcal{A}^*$ for which $2\leq d_z(A,B)=p(A,B)$ and $u\in A, v\in B$. We may assume that $|A|\leq|B|$. Then $|B|\geq 2$ for otherwise $|A|=|B|=1$ and hence $p(A,B)=(2-t)^+=0$. Hence $B':=B-v$ is nonempty and $p(A,B')\geq p(A,B)-1$. We have $d_z(A,B')\leq d_z(A,B)-z(uv)=p(A,B)-z(uv)\leq p(A,B)-2\leq p(A,B')-1$ contradicting the assumption that z covers p.

Since $p(A,B)$ is positive whenever $A\cup B$ induces a large bi-clique, it follows that the edge set $F_z:=\{uv\in E:z(uv)=1\}$ is a t-covering of G for which $|F_z|=z(A^*)=M$, which proves the min-max formula.

To see the second half of the theorem, let us choose an optimal $\mathcal{H}$ in (9) so that $|\mathcal{H}|$ is maximum. We claim that $\mathcal{H}$ consists of stars and essential bi-cliques. Suppose indirectly that $\mathcal{H}$ contains a bi-clique $D := K_{k,l}$ with stable sets K, L where $2 \leq |K| = k \leq |L| = l, l \geq t$. Let $u \in K$, $D' := (u, L)$ and $D'' := (K - u, L)$ and let $\mathcal{H}' := \mathcal{F} - \{D\} \cup \{D', D''\}$. Since $p_t(D) = k+l-t, p_t(D') = 1+l-t, p_t(D'') = k-1+l-t$, we have $p_t(\mathcal{H}') = p_t(\mathcal{H}) - (k+l-t) + (1+l-t) + (k-1+l-t) \geq p_t(\mathcal{H}) + l - t \geq p_t(\mathcal{H})$. Hence $\mathcal{H}'$ is another optimal packing of bi-cliques contradicting the maximality of $|\mathcal{H}|$. $\square$

Let us formulate the theorem in an equivalent form, too.

Theorem 3.4. *In a bipartite graph $G = (S, T; E)$ the maximum number of edges of a subgraph not containing large bi-cliques (:bi-cliques with more than t nodes) is equal to* $\min\{|E| - \sum_i (|V(D_i)| - t) : \{D_1, \ldots, D_k\}$ *is a set of pairwise edge-disjoint stars and essential bi-cliques*$\}$. $\square$

The proof of Theorem 1.3 in [2] is not algorithmic and hence the present proof is not algorithmic either. On the other hand in [2] we showed, relying on the ellipsoid method and the theorem itself, that there is a polynomial time algorithm to compute the minimum in Theorem 1.3. Furthermore, T. Fleiner [1] showed how to compute the maximum. Therefore there are polynomial algorithms to compute the extrema in question but it remains a challenge to find alternative, combinatorial algorithms as well.

As far as weighted extensions are considered, we mentioned already Király's observation [8] that in bipartite graphs finding a minimum weight (or equivalently the maximum weight) square-free 2-factor is NP-complete. Therefore so is the more general problem of finding a maximum weight square-free 2-matching. However, this latter problem is tractable for a class of weight functions which includes the cardinality function. Let $w : V \to \mathbf{R}$ be a node-function and define a weight function $c : E \to \mathbf{R}$ on the edge-set E by $c(uv) := w(u) + w(v)$. Then c is called a **node-induced** weight-functions on E.

For node-induced weight-functions [2] contained a weighted extension of Theorem 1.3, as well, for the case when a minimum weight covering of a positively crossing bi-supermodular function is considered. Relying on this, the same approach we used above may be used to extend, theorems 2.1 and 3.3 for induced weight-functions, and a formula may be given for the minimum weight of t-covering of bi-cliques or for the maximum weight of a t-matching not containing $K_{t,t}$.

References

[1] T. Fleiner, Uncrossing a family of set-pairs, Combinatorica, in print (2000).

[2] A. Frank and T. Jordán, Minimal edge-coverings of pairs of sets, J. Combinatorial Theory, Ser. B, Vol. 65, No. 1 (1995, September) pp. 73-110.

[3] J. Geelen, The C_6-free 2-factor problem in bipartite graphs is NP-complete, oral communication (December, 1999).

[4] D. Hartvigsen, The square-free 2-factor problem in bipartite graphs, in: Integer Programming and Combinatorial Optimization (eds., G. Cornuejols, R. Burkard, G.J. Woeginger) Springer Verlag, Lecture Notes in Computer Science 1610, 1999, pp. 234-240. (Extended abstract).

[5] D. Hartvigsen, Finding maximum square-free 2-matchings in bipartite graphs, (September 2000), manuscript, (detailed and improved version of [4]).

[6] Z. Király, Square-free 2-matchings in bipartite graphs, unpublished manuscript (1999, September).

[7] Z. Király, $K_{t,t}$-free t-matchings in bipartite graphs, to be submitted (2000).

[8] Z. Király, The minimum cost square-free 2-factor problem in bipartite graphs is NP-complete, oral communication.

Special thanks are due to Zoltán Király and Tibor Jordán for several helpful discussions.