Restricted t-matchings in Bipartite Graphs

Zoltán Király*

Abstract

The results contained in this paper were achieved in 1999, when the author participated the ,,Workshop on Matroids, Matching and Extensions" in Waterloo, Canada. This paper was written and distributed in 2001, when the author gave an EIDMA minicurse in Eindhoven, Netherlands. This is the uncorrected original version, except, that we added an Appendix containing the NP-completeness proof for the weighted $K_{2,2}$-free factor problem.

David Hartvigsen [H] recently gave an algorithm to find a square-free 2–factor in a bipartite graph and using this algorithm he proved several theorems. In a technical report [K] in 1999 we gave a simpler and stronger form of the Tutte-type theorem and gave a new Berge-type theorem on maximum square-free 2-matchings. Jim Geelen [G] showed that we cannot go further, deciding whether a bipartite graph has a C_6-free 2-factor is NP-complete. Using his ideas we can prove that if weights are given on the edges then finding the maximum weight square-free 2-factor in bipartite graphs is NP-hard, even with $0-1$ weights.

András Frank made the crucial observation that $C_4 = K_{2,2}$ and gave a min-max formula for $K_{t,t}$-free t-matchings. Now, following the lines of the technical report [K], we give a simple inductive proof for this generalization. The proof follows the idea of the inductive proof for the Hall's theorem given by Halmos and Vaughn [HV].

1 Introduction

Let $G = (V, E)$ be a simple bipartite graph with bipartition $V = X \cup Y$, $t \geq 2$ an integer, and let $f : V \to \{0, 1, 2, \ldots, t\}$ be a function on the vertices. For a vertex $v \in V$ we call the value $f(v)$ the weight of v. For a set $V' \subseteq V$ we write $f(V')$ for $\sum_{v \in V'} f(v)$. A subgraph $G' \subseteq G$ is called an *f-matching* if $\deg_{G'}(v) \leq f(v)$ $\forall v \in V$. We should note, that in the literature this is usually called a simple f-matching, but in

*Department of Computer Science and Communication Networks Laboratory, Eötvös University, Pázmány Péter sétány 1/C Budapest, Hungary H-1117. Research is supported by EGRES group (MTA-ELTE) and OTKA grants NK 67867, K 60802 and CNK 77780. E-mail: `kiraly@cs.elte.hu`

this paper the word "simple" will be omitted. The size of an f-matching is the number of edges contained. An f-matching is called an *f–factor* if $\deg_{G'}(v) = f(v)\ \forall v \in V$. If $f \equiv t$ we are speaking about t-matching and t-factor.

An f-matching (f-factor) is called $K_{t,t}$-free if it does not contain $K_{t,t}$ (the complete bipartite graph on $t + t$ vertices) as a subgraph. The maximum size of an $K_{t,t}$-free f-matching in G is denoted by $\nu_t^f(G)$.

A subgraph will be called a full weight $K_{t,t}$ if it is isomorphic to $K_{t,t}$ and the weights of all its vertices are exactly t. For $Z \subseteq V$ let $G - Z$ denote the subgraph of G induced by $V - Z$. Let $i(V - Z)$ denote the number of edges induced by $V - Z$, and $c_t(G - Z)$ denote the number of full weight $K_{t,t}$ components of $G - Z$. Moreover let $\mathbf{q(Z)} := f(Z) + i(V - Z) - c_t(G - Z)$. We call a subset Z *non-trivial* if it intersects both X and Y.

2 The theorem

The following theorem is a slight generalization of a theorem of Frank [F]. He proved the theorem for the case $f \equiv t$.

Theorem 1. *The maximum size of a $K_{t,t}$-free f-matching in bipartite graph G, $\nu_t^f(G)$, is equal to*

$$\tau_t^f(G) := \min\{q(Z);\ Z \subseteq V\}.$$

Proof. It is easy to see that $\nu_t^f \le \tau_t^f$, as for any subset Z the edges of a $K_{t,t}$-free f-matching are either incident to Z, there are at most $f(Z)$ such edges, or are induced by $V - Z$, but at least one edge must be missing from every $K_{t,t}$ component.

Now we prove $\nu_t^f \ge \tau_t^f$ by induction on $f(V) + |V|$. We start the induction with any edgeless graph, clearly $\nu_t^f = \tau_t^f = q(\emptyset) = 0$. If there exist a vertex v with zero weight use induction on $G - v$. From now on we suppose G has some edges and $f(v) \ge 1$ for all $v \in V$.

We will use two procedures.

Edge-reduction(x, y): if $xy \in E$ then delete the edge xy and let $f'(x) = f(x) - 1,\ f'(y) = f(y) - 1$.

$K_{t,t}$-reduction$(x_1, \dots, x_t; y_1, \dots, y_t)$: if $x_i \in X,\ y_i \in Y,\ x_iy_j \in E$ for all $i, j = 1, \dots, t$, and $f(x_i) = f(y_i) = t$ for $i = 1, \dots, t$ then contract the x_i into a new vertex x, and contract the y_i into a new vertex y, delete the parallel edges as well as the edge xy, and let $f'(x) = f'(y) = 1$.

We will use the inverse of these operations, too. While inverting the reductions we suppose we have a $K_{t,t}$-free f'-matching F' of size $\nu_t^{f'}(G')$ and a set Z' with $q'(Z') = \tau_t^{f'}(G')$ in the reduced graph G', moreover that Z' is a maximal set satisfying this property.

Inverse-edge-reduction(x, y): add xy to E', let $f(x) = f'(x) + 1,\ f(y) = f'(y) + 1$, $Z = Z'$ and $F = F' + xy$. Observe that $q(Z) \le q'(Z') + 2$ and equality holds only when $x, y \in Z'$. F is an f-matching of size $|F'|+1$ and if it is not $K_{t,t}$-free then it has exactly one $K_{t,t}$ subgraph and this subgraph contains edge xy.

Inverse-$K_{t,t}$-reduction(x, y): We delete x and y, reconstruct G and put back weights t on the x_i and on the y_j. Construct f-matching F in G as follows. If F' has an edge incident with x then this edge came from the original edge vx_k at the contraction, otherwise (if F has no edge incident with x) let $k = 1$. Similarly if F' has an edge incident with y then this edge came from the original edge uy_l at the contraction, otherwise let $l = 1$. Start F with F', replacing edge vx by vx_k and edge uy by uy_l if they exist. Then add all edges x_iy_j except x_ky_l. Clearly $|F| = |F'| + t^2 - 1$ and F is an f-matching in G. We claim that F is $K_{t,t}$-free. Suppose that there is a $K_{t,t}$ in F. As x_ky_l is not an edge in F, it has some vertices outside the $K_{t,t}$ contracted. As F' was $K_{t,t}$-free, this $K_{t,t}$ must contain either edge vx_k or edge uy_l. It cannot contain both as x_ky_l is not an edge of F. And it cannot either contain one of them because this edge would be a cut-edge of this $K_{t,t}$. Let Z be a "blown-up" of Z', i.e., if $x \in Z'$ then replace it by $x_1, \ldots, x_t$ and if $y \in Z'$ then replace it by $y_1, \ldots, y_t$. If exactly one of x and y is in Z' then $q(Z) = q'(Z') + t^2 - 1$. If $x, y \notin Z'$ then, by the maximality of Z', all neighbors of x and y are in Z' – otherwise e.g. $q'(Z' + x) = q'(Z')$ using $f'(x) = 1$. Consequently $q(Z) = q'(Z') + t^2 - 1$ (as $i(V - Z) = i(V' - Z') + t^2$ and $c_t(G - Z) = c_t(G' - Z') + 1$). If $x, y \in Z'$ then $q(Z) = q'(Z') + 2t^2 - 1$.

Case I for every subset Z of vertices with $q(Z) = \tau_t^f(G)$, the set Z is trivial.

Take any edge uv and use Edge-reduction. By induction $\tau_t^{f'}(G') = \nu_t^{f'}(G')$. After making the inverse, if $\tau_t^f(G) = \tau_t^{f'}(G') + 2$ then we have the set Z with $q(Z) = \tau_t^f(G)$ and $u, v \in Z$, i.e., Z is non-trivial, a contradiction. If F is $K_{t,t}$-free then we are done, so suppose F has a $K_{t,t}$ subgraph H. Let $F^* = F$ for later usage and remember that F^* is an f-matching of G with exactly one $K_{t,t}$ component, namely H, and F^* has size $\tau_t^{f'}(G') + 1 \geq \tau_t^f(G)$.

Now we apply $K_{t,t}$-reduction to H in the original graph G. Using induction and the inverse procedure we are done unless $\tau_t^f(G) \geq \tau_t^{f'}(G') + t^2$ and consequently $x, y \in Z'$. In this case all edges of H are induced by Z and $q(Z) = \tau_t^{f'}(G') + 2t^2 - 1 < \tau_t^f(G) + t^2$. A simple calculation shows that this cannot be the case: $\tau_t^f(G) \leq |F^*| \leq (f(Z) - t^2) + i(V - Z) - c_t(G - Z) = q(Z) - t^2 < \tau_t^f(G)$ as F^* uses at least t^2 edges induced by Z.

Case II there exist subset Z of vertices with $q(Z) = \tau_t^f(G)$ so that Z is non-trivial. Choose Z^* to be a maximal set with these properties. Recall that the maximality implies that each vertex $v \in V - Z^*$ has degree strictly less than $f(v)$ in graph $G - Z^*$ unless v is in a full weight $K_{t,t}$ component of $G - Z^*$, when this degree is $f(v) = t$. Now apply $K_{t,t}$-reduction for all $K_{t,t}$ components of $G - Z^*$, one after the other, then apply Edge-reduction to edges induced by $V - Z^*$, again, one after the other in any order, finally simply delete all edges induced by Z^* resulting in a graph $G^* = (V^* = X^* \cup Y^*, E^*)$ with weight function f^*. The way we arrived at G^* ensures edges neither go inside Z^* nor outside Z^*. Let G_1 denote the subgraph of G^* induced by $(Z^* \cap Y^*) \cup (X^* - Z^*)$ and G_2 the subgraph induced by $(Z^* \cap X^*) \cup (Y^* - Z^*)$. Note that – using the fact that Z^* is non-trivial – both G_1 and G_2 have less vertices than G.

We claim that $\tau_t^{f^*}(G_1) = f^*(Z^* \cap Y^*)$ and $\tau_t^{f^*}(G_2) = f^*(Z^* \cap X^*)$. By symmetry it is enough to prove the first statement. Suppose we have a subset Z_1 of the vertices of G_1 with $q_1(Z_1) < f^*(Z^* \cap Y^*)$, choose Z_1 to be maximal with this property. Defining $Z' = Z_1 \cup (Z^* \cap X^*)$ we have $q^*(Z') = q_1(Z_1) + f^*(Z^* \cap X^*) < f^*(Z^*)$. We make the inverse steps in reversed order. As $Z' \cap Y^* \subseteq Z^* \cap Y^*$ at each inverse step we have $y \notin Z'$. So in every inverse Edge-reduction step $q(Z')$ is increased by one, and in every inverse $K_{t,t}$-reduction step it is increased by $t^2 - 1$ because x cannot have neighbors outside Z_1 by the maximality of Z_1 and y cannot have neighbors outside $Z^* \cap X^*$ by the maximality of Z^*. Finally we get a set Z with $q(Z) < q(Z^*)$, a contradiction.

By induction there is a $K_{t,t}$-free f^*-matching F_1 in G_1 of size $f^*(Z^* \cap Y^*)$ and a $K_{t,t}$-free f^*-matching F_2 in G_2 of size $f^*(Z^* \cap X^*)$. Clearly $F' = F_1 \cup F_2$ is a $K_{t,t}$-free f^*-matching in G^*. Now make the inverse steps in reversed order resulting in an f-matching F with size $q(Z^*) = \tau_t^f(G)$ (using the maximality of Z^* and that in every step $x, y \notin Z^*$). Finally we claim that F is $K_{t,t}$-free. As F' was $K_{t,t}$-free and in inverse $K_{t,t}$-free -reduction steps we do not create a $K_{t,t}$, so a $K_{t,t}$ subgraph H of F can appear only in an inverse Edge-reduction step when putting back edge xy. As Z^* induces no edges we can assume for example that H has no vertices in $Z^* \cap Y$. This means that in the original graph x had at least t neighbors in $Y - Z^*$ but x is not a vertex of a $K_{t,t}$ component of $G - Z^*$. This contradicts to the maximality of Z^*. □

We list some immediate corollaries.

Corollary 2. *Let $t \geq 2$ be an integer and $f : V \to \{0, 1, 2, \ldots, t\}$ be a function on the vertices. A bipartite graph G has a $K_{t,t}$-free f–factor iff*

$$\text{for all } Z \subseteq V: \quad q(Z) = f(Z) + i(V - Z) - c_t(G - Z) \geq \frac{f(V)}{2}. \qquad \square$$

Corollary 3. *Let $t \geq 2$ be an integer. The maximum size of a $K_{t,t}$-free t-matching in bipartite graph G is equal to*

$$\min\{t|Z| + i(V - Z) - c_t(G - Z);\ Z \subseteq V\}. \qquad \square$$

Corollary 4. *Let $t \geq 2$ be an integer. A bipartite graph G has a $K_{t,t}$-free t–factor iff*

$$\text{for all } Z \subseteq V: \quad t|Z| + i(V - Z) - c_t(G - Z) \geq \frac{t|V|}{2}. \qquad \square$$

For the case of $t = 2$ we state this corollary in the old-fashioned form of [K] as well. Let $I(G - Z)$ denote the set of isolated vertices of $G - Z$, $K(G - Z)$ denote the set of two-vertex components of $G - Z$ and $C(G - Z)$ denote the set of C_4 components of $G - Z$.

Corollary 5. *A bipartite graph G has a C_4-free 2–factor iff*

$$\text{for all } Z \subseteq V: \quad |Z| \geq |I(G - Z)| + |K(G - Z)| + |C(G - Z)|.$$

Proof. Seeing the necessity is easy, in a C_4-free 2–factor at least two edges must leave every counted component of $G - Z$ but at most $2|Z|$ edges can enter Z. To see the sufficiency suppose that G does not admit a C_4-free 2–factor. By Corollary 4 there is a subset Z with $2|Z| + i(V - Z) - c_2(G - Z) < |V|$, choose a maximal set Z with this property. Using the maximality of Z it is easy to see that every component of $G - Z$ is either an isolated vertex or a K_2 or a C_4: if $G - Z$ has a vertex v with at least two neighbors in $G - Z$ then $2|Z + v| + i(V - Z - v) - c_2(G - Z - v) \leq 2|Z| + i(V - Z) - c_2(G - Z)$ unless v is in a C_4 component of $G - Z$. Thus using this fact and $|Z| < |V - Z| - i(V - Z) + c_2(G - Z)$, we get $|Z| < |V - Z| - |K(G - Z)| - 4|C(G - Z)| + |C(G - Z)| = (I(G - Z) + 2|K(G - Z)| + 4|C(G - Z)|) - |K(G - Z)| - 3|C(G - Z)| = |I(G - Z)| + |K(G - Z)| + |C(G - Z)|$. □

3 Appendix

Problem 6. Given a bipartite graph and $0 - 1$ weights on the edges, and a limit K. Decide whether there exist a C_4-free (that is, a $K_{2,2}$-free) 2-factor of weight at least K.

Theorem 7. *Problem 6 is NP-complete.*

Proof. Clearly being in NP, it is enough to prove that some known NP-complete problem is reducible to Problem 6.

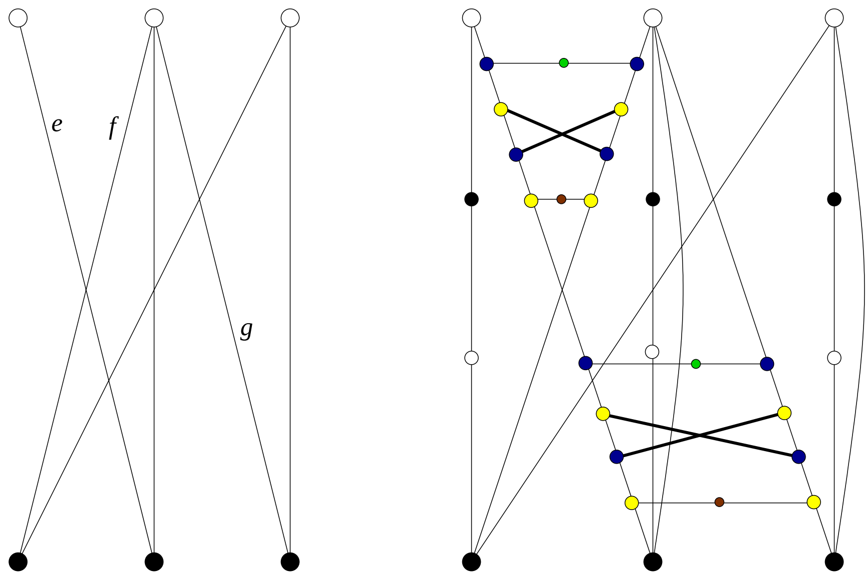

Figure 1: The construction for a small graph. Here $e_1^1 = e$, $e_1^2 = f$, $e_2^1 = e$, $e_2^2 = g$. Bold edges have weight 1, others have weight 0.

Using the idea of Geelen [G], we reduce the bipartite version of Problem GT55 in [GJ], namely

Theorem 8 (Garey-Johnson)**.** *Given a bipartite graph, and some pairs of edges* $\{e_1^1, e_1^2\}$, $\{e_2^1, e_2^2\}, \dots \{e_k^1, e_k^2\}$. *It is NP-complete to decide whether there is a perfect matching that uses at most one of* $\{e_i^1, e_i^2\}$ *for all* $1 \le i \le k$.

The construction of the reduction is as follows. Suppose the lower vertex class is $\{u_1, \dots, u_n\}$, and the upper vertex class is $\{v_1, \dots, v_n\}$. First we put new length-three paths for all $1 \le j \le n \;:\; u_j, x_j, y_j, v_j$ (where x_j and y_j are new vertices). Observe that the 2-factors of this new graph correspond to the perfect matchings of the original graph.

Then for each prescribed pair of edges $\{e_i^1, e_i^2\}$ we put a gadget. Subdivide e_i^1 by vertices $a_i^1, a_i^2, a_i^3, a_i^4$, and e_i^2 by vertices $b_i^1, b_i^2, b_i^3, b_i^4$ (downwards). Connect a_i^1 and b_i^1 by a path of length 2, a_i^2 and b_i^3 by an edge of weight 1, a_i^3 and b_i^2 by an edge of weight 1, and a_i^4 and b_i^4 by a path of length 2. All edges, where it was not stated, have weight 0. (See Figure 1.) It is easy to check, that the resulting graph is bipartite.

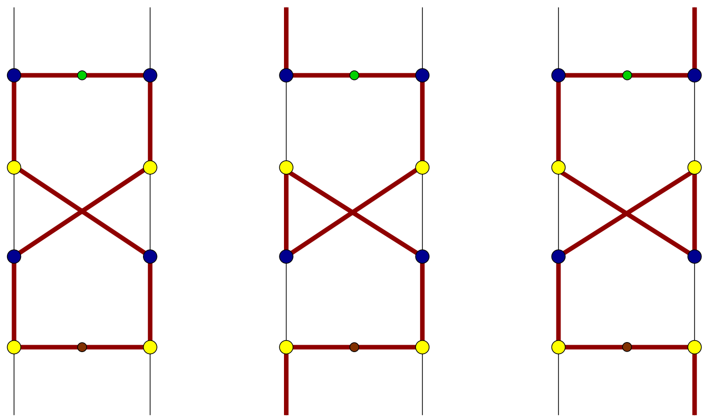

Figure 2: The three possibilities a weight $2k$ C_4-free 2-factor can cross a gadget. Here bold edges are the edges of the 2-factor.

It is a routine to prove, that the original graph has a perfect matching that uses at most one of $\{e_i^1, e_i^2\}$ for all $1 \le i \le k$, if and only if, there is a C_4-free 2-factor of weight $\ge K = 2k$ in the new graph: As k is the number of gadgets, $2k$ is the sum of all weights. So every 2-factor of weight $2k$ must use every ,,horizontal" edge (that is the weight 1 edges and edges of the added 2-paths) of each gadget. If the 2-factor is C_4-free, then there are three possibilities to do so, see Figure 2.

Remark. By changing the weights from $0-1$ to $1-2$, and appropriately adjusting the value of K, one can easily see, that the same proof gives the NP-completeness of the decision version of the maximum weight C_4-free 2-matching problem.

References

[GJ] M. Garey and D. Johnson; Computers and Intractability, Freeman, 1979.

[HV] P. R. Halmos, H. E. Vaughn, *The marriage problem*, Amer. J. Math. **72**, 1950, 214-215. MR11, 423.

[H] D. Hartvigsen, *The Square-Free 2-Factor Problem in Bipartite Graphs*, Integer Programming and Combinatorial Optimization, 7th IPCO Conference Graz, 1999, 234-241. LNCS 1610.

[F] A. Frank, *Restricted t-matchings in Bipartite Graphs*, Discrete Applied Mathematics **131**, 2003, 337-346.

[G] J. Geelen, personal communication, 1999, Waterloo

[K] Z. Király, C_4*–free 2-factors in Bipartite Graphs*, Egres Technical Report TR-2001-13, www.cs.elte.hu/egres/

[LP] L. Lovász, M. D. Plummer, *Matching Theory*, Akadémiai Kiadó, Budapest, 1986