Egerváry Research Group
on Combinatorial Optimization

Technical reportS

TR-2002-07. Published by the Egrerváry Research Group, Pázmány P. sétány 1/C, H–1117, Budapest, Hungary. Web site: www.cs.elte.hu/egres . ISSN 1587–4451.

Generalized Induced Factor Problems

Zoltán Király and Jácint Szabó

October 2002

Generalized Induced Factor Problems

Zoltán Király* and Jácint Szabó**

Abstract

Given a graph $G = (V, E)$ and a family $\mathcal{H}$ of graphs, a subgraph G' of G is usually called an $\mathcal{H}$-factor, if it is a spanning subgraph of G and its every component is isomorphic to some member of $\mathcal{H}$. Here we focus on the case $K_2 \in \mathcal{H}$. Many nice results are known in the literature for this case. We show some very general theorems (Tutte type existence theorem, Tutte-Berge type minimax formula, Gallai-Edmonds type structure theorem) that can be considered as a common generalization of almost all such known results. In this paper we use a stricter and more general concept for $\mathcal{H}$-factors, namely where the components of G' must be induced subgraphs of G.

Keywords. matching, factor, structure theorem, matroid

1 Introduction

Generalized factor problems have a wide literature. These problems can be divided into three areas: (1) where the possible degrees are prescribed; (2) where the shape of the components of the factors are prescribed as given graphs and K_2 is allowed to be a component; and (3) where the shape is prescribed but K_2 is not allowed.

Most of the problems in (3) are NP-complete, the only known solvable case is the long-path factor problem [6, 13]. We now focus on the case (2). We give a common generalization of most of the known results in this area [14, 8, 4, 1, 3, 2, 5, 11, 12, 10, 7]. Our generalization is similar to the theorems of Loebl and Poljak [11, 12, 10] but further generalizes their results. We are dealing with induced factors, in this area the only known result was due to Kelmans [7]. This is a real generalization, because if we take all the supergraphs (definition see below) of the graphs of the prescribed family we arrive back to the non-induced factor problems. On the other hand, for

*Department of Computer Science, Eötvös University, Pázmány Péter sétány 1/C, Budapest, Hungary H-1117. The author is a member of the Egerváry Research Group (EGRES). e-mail: `kiraly@cs.elte.hu`

**Department of Operations Research, Eötvös University, Pázmány Péter sétány 1/C, Budapest, Hungary H-1117. The author is a member of the Egerváry Research Group (EGRES). e-mail: `jacint@cs.elte.hu`

Research is supported by OTKA grants T 030059, T 037547 and N 034040.

example, if we are looking for a factor having K_2 components and induced five-cycles, this problem was not settled yet.

Throughout the paper all graphs are finite and simple. K_n denotes the complete graph on n vertices. G will always denote a graph $G = (V, E)$. If $V' \subseteq V$ then $G[V']$ denotes the subgraph of G induced by V'. For any graph $H = (V_H, E_H)$ we use the notation $V(H)$ for V_H and $E(H)$ for E_H.

In this extended abstract we will assume the knowledge of basics of matching theory, for concepts and theorems used see e.g. [9].

Definition 1.1. If G is a graph, we call a graph $G^* = (V, E^*)$ a *supergraph* of G if $E \subseteq E^*$, i.e. if G is a spanning subgraph of G^*.

Definition 1.2. Let $\mathcal{H}$ be a family of graphs. A subgraph $G' = (V', E')$ of G is called an *$\mathcal{H}$-packing* if every (connected) component of G' is an *induced* subgraph of G and is isomorphic to some member of $\mathcal{H}$.

The size of an $\mathcal{H}$-packing is the number of vertices covered. An $\mathcal{H}$-packing is called maximum, if it has the largest possible size. $\text{minfree}_{\mathcal{H}}(G)$ denotes the minimum number of uncovered vertices, i.e. the number of uncovered vertices for any maximum $\mathcal{H}$-packing.

An $\mathcal{H}$-packing is called an *$\mathcal{H}$-factor*, if its size equals to the number of vertices of G, i.e. if it is a spanning subgraph.

Recall that a connected graph is called factor-critical if after deletion of any vertex the remaining graph has a perfect matching. K_1 is called trivial, all other factor-critical graphs are called non-trivial.

Definition 1.3. A connected graph F is called $\mathcal{H}$-critical, if it has no $\mathcal{H}$-factor, but after deleting any vertex the resulting graph has. The number of $\mathcal{H}$-critical components of graph G is denoted by $c_{\mathcal{H}}(G)$, while $c(G)$ denotes the number of connected components of G.

Claim 1.4. *If $K_2 \in \mathcal{H}$ then a factor-critical graph F is $\mathcal{H}$-critical if and only if F has no $\mathcal{H}$-factor.*

Definition 1.5. A connected graph is called a *propeller* if it has a vertex c (called the center) such that after deleting c, each component (called blade) is factor-critical. A propeller is called trivial if it has two vertices. A propeller with one blade has a perfect matching. A propeller with at least two blades has a unique center.

Definition 1.6. Let k be a positive integer and $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_k$ be arbitrary families of factor-critical graphs such that $K_1 \in \mathcal{F}_1$. Construct a propeller in the following way. Choose some factor-critical graphs $D_1, D_2, \ldots, D_{k'}$ from $1 \leq k' \leq k$ different families $\mathcal{F}_i$. Add a new vertex c and connect it to arbitrary nonempty subset of vertices of each D_i.

A family $\mathcal{P}$ of graphs is called a *propeller family* of order k (defined by $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_k$), if it consists of all the propellers constructed in the way described above.

Definition 1.7. A family $\mathcal{H}$ of graphs is called *admissible* if $\mathcal{H} = \mathcal{P} \cup \mathcal{F}$, where $\mathcal{P}$ is a propeller family and $\mathcal{F}$ is a family of non-trivial factor-critical graphs. The admissible family is of order k, if $\mathcal{P}$ is of order k.

Note that as $K_1 \in \mathcal{F}_1$, an admissible family always contains K_2.

2 Main Results

In this section $\mathcal{H}$ will always denote an admissible family.

Definition 2.1. A subgraph $G' = (V', E')$ of graph G is called nice, if $G - V'$ has a perfect matching.

Lemma 2.2. *Let K be a factor-critical graph, and K' be a spanning subgraph having only factor-critical and propeller components. Then either one of the factor-critical components is nice or a two-bladed subpropeller of a propeller component is nice.*

Proof. We call the factor-critical components of K' and the blades of non-trivial propeller components of K' altogether potatoes. The proof goes by induction on the number of potatoes. If this number is one, then K' has one factor-critical component and all other components are K_2s, that is the factor-critical one is nice. (We used that K has an odd, while a propeller with one blade has an even number of vertices).

In this proof the words 'blade' and 'center' will refer to only blades and centers of non-trivial propellers. If K' has a non-trivial propeller with one blade then we can replace this component by a perfect matching of it – and we are done by induction. So we may suppose that every non-trivial propeller has at least two blades. Let M_1 denote the matching consisting of the trivial propellers. Let x be any vertex in a potato and consider an M_2 perfect matching of $K - x$. Now $M_1 \triangle M_2$ consists of alternating path and cycles. Every center is an endpoint of one such path, and every potato has an odd number of outgoing path (except one, where x resides). As the number of blades is at least twice the number of centers, there must be an alternating path connecting two potatoes. If these two potatoes are not the blades of a two-bladed propeller then we can replace the union of the two potatoes and the alternating path connecting them by a perfect matching getting K'' with smaller number of potatoes; and we can use induction. If the alternating path connects the blades of a two-bladed propeller, we can replace the union of this propeller and the path by one factor-critical component. The number of potatoes decreased again, and observe, that if the nice subgraph given by the induction hypothesis is this new factor-critical component then our two-bladed propeller is also nice. □

Lemma 2.3. *If K is factor-critical but not $\mathcal{H}$-critical then K can be factorized by edges and one bigger graph from $\mathcal{H}$ which is either factor-critical or a two-bladed propeller.*

Proof. By Claim 1.4, K has an $\mathcal{H}$-factor. Applying Lemma 2.2 we get the statement. □

Definition 2.4. For a subset $I \subseteq V$ of vertices, let $D_1, D_2, \ldots, D_t$ denote the components of $G[I]$ that are both factor-critical and $\mathcal{H}$-critical, and let $q_{\mathcal{H}}(I) = t$ (clearly $q_{\mathcal{H}}(I) \leq c_{\mathcal{H}}(G[I])$). Let $\Gamma(I)$ denote the set of neighbors of I, i.e. $\{v \in V - I \mid \exists x \in I \; : \; vx \in E\}$. For $x \in \Gamma(I)$ we define a weight function in the following way (we assume k and $\mathcal{F}_1, \ldots, \mathcal{F}_k$ to be fixed).

$$w_I(x) = \Big|\{i \; ; \; 1 \leq i \leq k, \; \exists j \; \text{ and a nice induced subgraph } F \text{ of } D_j, \text{ such that}$$

$$F \in \mathcal{F}_i \text{ and } x \text{ has at least one neighbor in } F\}\Big|.$$

Let $\gamma_w(I)$ denote $w_I(\Gamma(I)) = \sum_{x \in \Gamma(I)} w_I(x)$.

Definition 2.5. For a subset $I \subseteq V$ of vertices, $\mathrm{def}_{\mathcal{H}}(I) = q_{\mathcal{H}}(I) - \gamma_w(I)$ denotes the $\mathcal{H}$-deficiency of subset I. The $\mathcal{H}$-deficiency of G is defined as $\mathrm{def}_{\mathcal{H}}(G) = \max_{I \subseteq V} \mathrm{def}_{\mathcal{H}}(I)$.

Theorem 2.6. *G has an $\mathcal{H}$-factor if and only if*

$$c_{\mathcal{H}}(G[I]) \leq \gamma_w(I) \text{ for } \forall I \subseteq V.$$

Proof. This theorem will be an easy consequence of the following theorem. □

Theorem 2.7.

$$\mathrm{minfree}_{\mathcal{H}}(G) = \mathrm{def}_{\mathcal{H}}(G).$$

Before the proof we should point out, that if $k = 1$, $\mathcal{F}_1 = \{K_1\}$ and $\mathcal{F} = \emptyset$ then $\mathcal{H} = \{K_2\}$ and in this case Theorem 2.6 gives (a well-known form of) Tutte's theorem and Theorem 2.7 gives the Tutte-Berge formula.

Proof. It is enough to prove the statement for a connected graph G. First we prove that $\mathrm{minfree}_{\mathcal{H}}(G) \geq \mathrm{def}_{\mathcal{H}}(G)$. Let I be a subset where $\mathrm{def}_{\mathcal{H}}(I)$ takes the maximum value. We are going to show that any $\mathcal{H}$-packing G' leaves uncovered some vertex in at least $\mathrm{def}_{\mathcal{H}}(I)$ components of $G[I]$. Using the notations of Definition 2.4, let $Y = \{D_1, \ldots, D_t\}$ and $A = \Gamma(\bigcup Y) \subseteq \Gamma(I)$. Construct the following bipartite graph $B = (A^*, Y, E_B)$. Let $A^* = A^1 \cup A^2 \cup \ldots \cup A^k$ where $A^i = \{x^i \mid x \in A\}$. Let $D_j x^i$ be an edge in E_B if there is a nice induced subgraph F of D_j that is connected to x by at least one edge and $F \in \mathcal{F}_i$. (We say that F ensures the edge $D_j x^i$.) Let $\Gamma_B(.)$ denote the neighborhood in B. Observe, that $|\Gamma_B(Y)| = \gamma_w(I)$. As $K_1 \in \mathcal{F}_1$ and any one-vertex subgraph of a factor-critical graph is nice, hence we have $A^1 \subseteq \Gamma_B(Y)$.

First repeat the following procedure until it applies. Suppose there exists F such that either F is a factor-critical component of G' or it is a blade of a non-trivial propeller component P of G'; and F has a vertex x in A. If F is a blade of P replace P in G' by $P - F$; by the property of a propeller family $(P - F) \in \mathcal{H}$. Replace F by a perfect matching of $F - x$.

When the procedure ends up resulting the $\mathcal{H}$-packing G'', clearly G'' will cover the same vertices inside I and has the property that all factor-critical components and blades of non-trivial propeller components are disjoint from A.

G'' can cover D_j only if it has a component H that intersects both A and D_j. By the previous paragraph $V(H) \cap A = \{c\}$, the center of H. (If H is a K_2, then we call this endvertex the center.) If F denotes $H[V(H) \cap V(D_j)]$ then each component of F is a blade of H, therefore it is factor-critical. Now D_j is factorized by some components of G'' and by some blades. By Lemma 2.2 there is a nice factor-critical or subpropeller component of this factor. But, as D_j is not $\mathcal{H}$-factorizable, this must be a blade of propeller P centered in A. If a blade of P is a nice subgraph of D_j we say that P settles D_j. A propeller with center $c \in A$ can settle at most $w_I(c)$ of the D_js therefore at least $\mathrm{def}_{\mathcal{H}}(I) = \mathrm{def}_{\mathcal{H}}(G)$ of the D_js remain uncovered.

Now we show that $\mathrm{minfree}_{\mathcal{H}}(G) \leq \mathrm{def}_{\mathcal{H}}(G)$. First take the Gallai-Edmonds decomposition (see 3.2.1 in [9]) of G into sets A, D, C. Let $D_1, D_2, \ldots, D_t$ denote the components of $G[D]$ that are $\mathcal{H}$-critical and $D'_1, D'_2, \ldots, D'_l$ denote the components of $G[D]$ that are $\mathcal{H}$-factorizable. Let $Z = \{D_1, D_2, \ldots, D_t\}$, $Z' = \{D'_1, D'_2, \ldots, D_l\}$ and $Y = Z \cup Z'$. As components of $G[D]$ are factor-critical, by Claim 1.4, every component of $G[D]$ belongs to Y. Define the bipartite graph B as before. By a simple property of the decomposition there is a matching M_1 in B covering A^1. By Ore's theorem (see 1.3.1 in [9]) there is a matching M_2 in B leaving at most $\mathrm{def}_{\mathcal{H}}(G)$ uncovered vertices in Z, since for every subset $J \subseteq D_1 \cup D_2 \cup \ldots \cup D_t$ we have $\mathrm{def}_{\mathcal{H}}(J) \leq \mathrm{def}_{\mathcal{H}}(G)$. By the theorem of Mendelsohn and Dulmage (see 1.4.3 in [9]) there is a matching M covering simultaneously A^1 and $|Z| - \mathrm{def}_{\mathcal{H}}(G)$ vertices of Z. For a vertex $c \in A$ the edges of M incident to $c^1, c^2, \ldots c^k$ defines a propeller $P \in \mathcal{H}$ with center c having blades as nice subgraphs of different components of $G[D]$. This can be easily completed to an $\mathcal{H}$-packing that leaves at most $\mathrm{def}_{\mathcal{H}}(G)$ vertices uncovered. Take a perfect matching of $G[C]$ and $D_i - F$ or $D'_j - F$ for all F that is a blade of propeller defined above. For the remaining components in Z' take an $\mathcal{H}$-factor. □

Note that we proved something stronger. On one hand we proved that $\mathrm{minfree}_{\mathcal{H}}(G) = \max\ \mathrm{def}_{\mathcal{H}}(I)$ where the maximum is taken only for Is that are the union of some components of D. On the other hand we proved, that every maximum $\mathcal{H}$-packing has a propeller center in every vertex of A; and, by Lemma 2.3, there are maximum $\mathcal{H}$-packings having only K_2 components in C and in each component of $D[G]$ having only K_2 components and at most one blade or factor-critical component or two-bladed propeller.

We can also extend the decomposition theorem of Gallai and Edmonds.

Theorem 2.8. *Let $D_{\mathcal{H}}$ consist of the vertices of G that can be left uncovered by a maximum $\mathcal{H}$-packing. Let $A_{\mathcal{H}} = \Gamma(D_{\mathcal{H}})$ and $C_{\mathcal{H}} = V - (D_{\mathcal{H}} \cup A_{\mathcal{H}})$. Then*

1. *every component of $G[D_{\mathcal{H}}]$ is $\mathcal{H}$-critical*
2. *the set Z^* of this components is the minimal subset of Z (defined above) with the property $|Z^*| - \Gamma_B(Z^*) = \mathrm{def}_{\mathcal{H}}(G) = \mathrm{def}_{\mathcal{H}}(D_{\mathcal{H}})$*
3. *$A_{\mathcal{H}} \subseteq A$*
4. *every maximum $\mathcal{H}$-packing leaves one vertex uncovered in exactly $\mathrm{def}_{\mathcal{H}}(G)$ different components of $G[D_{\mathcal{H}}]$*

The proof is straightforward, we leave it to the full version of the paper.

We should point out that we are able to prove even more general theorems. The proofs are not much more difficult, but stating these theorems requires much space, so they will only contained in the full version. At this point we just indicate the four directions we can extend these results to.

1. In the definition of propeller families, instead of putting factor-critical graphs into $\mathcal{F}_i$, we may take (F, T) pairs where $\emptyset \neq T \subseteq V(F)$, and consider only propellers where the neighborhood of the center intersects F in exactly T.
2. We need not take all propellers with blades in different $\mathcal{F}_i$s, we may have a matroid on $\{1, 2, \ldots, k\}$, and consider a propeller belonging to the family only if the index set of the $\mathcal{F}_i$s containing a blade is independent.
3. For some members of $\mathcal{H}$ we may drop the condition that in an $\mathcal{H}$-packing they must appear as *induced* subgraph of G.
4. We may have a local form, instead of giving a universal family of allowed graphs, we are given some factor-critical and propeller *subgraphs* of graph G.

The special case of the following theorem, where $\mathcal{H}$ contains only K_2 and some factor-critical graphs and components of a packing need not be induced, was proved in [1]. (This proof is much more simple than the original one.)

Theorem 2.9. *Every $\mathcal{H}$-critical graph is factor-critical.*

Proof. Let G be an $\mathcal{H}$-critical graph. Since every graph in $\mathcal{H}$ is connected, G itself is also connected. By Theorem 2.8, $D_{\mathcal{H}} = V(G)$ and every component of $D_{\mathcal{H}}$ is a component of D, i.e. factor-critical by the original Gallai-Edmonds theorem. □

Note that by this theorem $q_{\mathcal{H}}(I) = c_{\mathcal{H}}(G[I])$ for all $I \subseteq V$.

The corollaries of these theorems for special allowed families will be listed in the next section. Here we review some algorithmic consequences of these results. Suppose allowed family $\mathcal{H}$ is fixed and our problem is, given a graph G, to find a maximum $\mathcal{H}$-packing of G.

Theorem 2.10. *If $\mathcal{H}$ is a* finite *allowed family of graphs, that is its order, k is finite, for all i family $\mathcal{F}_i$ is finite and $\mathcal{F}$ is finite, then there is a polynomial time algorithm for the problem of finding a maximum $\mathcal{H}$-packing of input graph G.*

Proof. As $\mathcal{H}$ is fixed and finite, its size $|\mathcal{H}|$ can be considered as a constant for this problem. The following constants will play an important role. Let k denote the order of $\mathcal{H}$, c' denote the maximum number of vertices in any factor-critical graph in $\mathcal{F} \cup \mathcal{F}_1 \cup \mathcal{F}_2 \cup \ldots \cup \mathcal{F}_k$ and finally $c = 2c' + 1$. First we prove the statement for a factor-critical input graph K. If K is $\mathcal{H}$-critical then any maximum matching is a maximum $\mathcal{H}$-packing, so Edmonds' matching algorithm does the job. If not then, by Lemma 2.3, it has a nice induced subgraph which is either a factor-critical graph in $\mathcal{F}$ or a two-bladed propeller from $\mathcal{P}$, therefore this nice subgraph can have at most c

vertices. We can examine all induced subgraphs of K with at most c vertices, check whether it is in $\mathcal{F}$ or it is a two-bladed propeller in $\mathcal{P}$, and, using Edmonds' algorithm, whether it is nice.

For general input graph G we use the proof of Theorem 2.7. First, using Edmonds' matching algorithm we get the Gallai-Edmonds decomposition. By the method used above we can determine Z' and Z and construct bipartite graph B, and for every edge of it store one nice F subgraph that ensured this edge. Constructing M_1 as well as M_2 is simple using the alternating path method for bipartite graphs. Finding M in $M_1 \cup M_2$ is an easy task. As we stored the nice subgraphs that ensured any edge of B, constructing the propellers centered in A is trivial. For components of Z' having no blade in them use the previous paragraph to construct an $\mathcal{H}$-factor. For the remaining part construct a maximum matching by Edmonds' algorithm. □

Note that as alternating path method gives the minimum set with maximum deficiency as well, essentially the same algorithm gives us the decomposition of Theorem 2.8 as well. We also have algorithms for some weighted versions.

Theorem 2.11. *The family of node-sets*

$$\{X \subseteq V \mid \exists G' \ \mathcal{H}\text{-packing such that } G' \text{ covers } X\}$$

gives the independent sets of a matroid.

The proof is left for the full version. The node-weighted $\mathcal{H}$-packing problem is the following. Given graph G and a non-negative weight function on the vertices, we are looking for the maximum weight node-set that can be covered by an $\mathcal{H}$-packing. The matroid given by the previous theorem is simple, so we can use greedy algorithm in polynomial time, and thus we get:

Theorem 2.12. *If allowed family $\mathcal{H}$ is finite, then the node-weighted $\mathcal{H}$-packing problem can be solved in polynomial time.*

3 Special cases

When the order of $\mathcal{H}$ is one, as all one-bladed propeller can be factorized by edges, we may assume that $\mathcal{H} = \{K_2\} \cup \mathcal{F}$, where $\mathcal{F}$ is a family of factor-critical graphs.

Theorem 2.6 is specialized to the following theorem of Cornuéjols, Hartvigsen and Pulleyblank [1].

Theorem 3.1. *G has an $\mathcal{H}$-factor if and only if $c_{\mathcal{H}}(G - X) \leq |X|$ for all $X \subseteq V$.*

We note that for this case one has a polynomial algorithm for finding a maximum $\mathcal{H}$-packing if one can decide for a factor-critical graph whether it is $\mathcal{H}$-factorizable in polynomial time.

Definition 3.2. The graph K is a $\triangle$-*cactus* if $K = K_1$ or each two-connected block is a triangle.

The following theorem was proved by Cornuéjols, Hartvigsen and Pulleyblank [4]. Let $\mathcal{H}$ consist of K_2 and the supergraphs of all odd cycles of length at least 5. Using the ear-decomposition of factor-critical graphs, it is not hard to see, that a factor-critical graph is $\mathcal{H}$-critical if and only if it is a $\triangle$-cactus. Thus, using Theorem 2.6 we get:

Theorem 3.3. *G can be factorized by edges and non-induced odd cycles of length at least* 5 *if and only if for every subset X of vertices, the $\triangle$-cactus components of $G-X$ is at most $|X|$.*

Now we consider the case when $k \geq 2$ and $K_1 \in \mathcal{F}_i$ for $1 \leq i \leq k$. We note this case is equivalent to the case where $\mathcal{F}_i = \{K_1\}$ for all i. This means that $\mathcal{P}$ consists of the $\leq k$-stars, calling a connected graph an $\leq k$-star, if it has a specified vertex, the center c, such that $S-c$ is the union of i isolated vertices ($1 \leq i \leq k$). The $\mathcal{H}$-factor problem is to factorize a graph with induced $\leq k$-stars and induced factor-critical graphs of $\mathcal{F}$.

The special case of Theorem 2.7 is the following:

Theorem 3.4.

$$\text{minfree}_{\mathcal{H}}(G) = \max\{c_{\mathcal{H}}(G-X) - k|X| \; ; \; X \subseteq V(G)\}.$$

A basic case is when $\mathcal{F} = \emptyset$, this is the problem of factorization by induced $\leq k$-stars. This problem was first studied by Kelmans [7].

Definition 3.5. A graph is an *oct* (odd clique tree), if its each two-connected block is an odd clique. Let *oct*(G) denote the number of oct components of G.

Note that the oct-s are factor-critical graphs.

Lemma 3.6. *Let F be a connected graph, and suppose $O = F - x - y$ is an induced oct subgraph of F, such that $xy \in E(F)$. Now either F is an oct or it can be factorized by edges and exactly one induced two-star.*

Proof. Suppose that F is not an oct. If there exists a vertex $v \in V(O)$ such that $vx \in E(F)$ and $vy \notin E(F)$, then $\{x, y, v\}$ induces a two-star, and $O - v$ has a perfect matching, so we are done. Thus we may assume that the nodes of O are either adjacent to both x and y or none of them. Let S denote the nodes of O adjacent to x, y. As F is connected and is not an oct, $|S| \geq 2$. Let $u_1, u_2 \in S$.

Suppose there exists a block B of O such that $S \subseteq V(B)$. F is not an oct, so there is a node $v \in V(B) \setminus S$. Now $\{x, u_1, v\}$ induces a two-star, it is easy to see that $O - \{u_1, u_2, v\}$ has a perfect matching, and the rest of the G is the edge yu_2, so we are done, since we have constructed an appropriate factor of G.

If there is no block B of O such that $S \subseteq V(B)$, then there exist nodes v', v'' of S such that none of the blocks of F includes both. Let P be $v' = u_1, u_2, \ldots, u_m = v''$, the (unique) shortest path between v' and v'' in O. It is clear that $m \geq 3$, and that $u_2, \ldots, u_{m-1}$ are cut-vertices in O. For $i = 1$ and $i = 3$ let w_i be such a vertex that w_i is adjacent to u_2 and u_i, but w_i is not in P. These w_i-s exist, since each block of O has at least 3 vertices, and P contains at most 2 vertices from a block. As u_2 is a cut-vertex, $w_1 w_3 \notin E(O)$. Using the two-star induced by $\{w_1, u_2, w_3\}$, the edges xv' and yv'', and a perfect matching of $O - \{v', v'', w_1, u_2, w_3\}$ we are done. $\square$

Lemma 3.7. *If the factor-critical graph F is not an oct, then it can be factorized by edges and exactly one induced two-star.*

Proof. Let O be a maximal, nice, induced, oct subgraph of F. O exists, because an arbitrary vertex of F is a nice oct subgraph. Let M be a perfect matching of $F - V(O)$. F is not an oct, so $O \neq F$, hence M is not empty, let $xy \in M$. By the maximality of O, the graph induced by $V(O) \cup \{x, y\}$ is not an oct, so by Lemma 3.6 it can be factorized by edges and exactly one induced two-star. Using the remaining part of M we get the desired factor of F. □

It can be easily checked that an oct cannot be factorized by induced $\leq k$-stars. By Lemma 3.7, a factor-critical graph, which is not an oct, can be factorized by edges and an induced two-star. Hence by Theorem 2.7 we obtain the theorem of Kelmans [7].

Theorem 3.8. *If $k > 1$, then the critical graphs for the induced $\leq k$-star factorization problem are the oct-s. A maximum induced $\leq k$-star packing of G leaves $\max_{X \subseteq V(G)} oct(G - X) - k|X|$ vertices uncovered.*

Taking $k = |V|$ we get:

Theorem 3.9. *A connected graph can be factorized by induced stars iff it is not an oct.*

Now we consider the general case, that is when $\mathcal{F} \neq \emptyset$.

Definition 3.10. An oct O is an $\mathcal{F}$*-oct*, if O does not have an induced subgraph belonging to $\mathcal{F}$. Let $oct_{\mathcal{F}}(G)$ denote the number of $\mathcal{F}$-oct components of G.

Lemma 3.11. *An oct can be $\mathcal{H}$-factorized iff it has an induced subgraph belonging to $\mathcal{F}$.*

Proof. If oct O is $\mathcal{H}$-factorized, then this factor contains an induced factor-critical graph from $\mathcal{F}$, because an oct cannot be factorized by induced stars. To see the other direction, suppose that oct O has an induced subgraph F belonging to $\mathcal{F}$. First observe that the two-connected blocks of a factor-critical graph have an odd number of vertices, therefore F contains either zero or an odd number of vertices from each block of O. We claim that any connected subgraph L with this property is nice in O. We prove this by induction on the number of vertices of O not covered by L. If this is zero then $L = O$. Otherwise there is a block B of O intersected, but not covered by L. As L has an odd number of vertices in B, let u, v be two vertices in B not covered by L. As B is a clique, $L' := O[V(L) \cup \{u, v\}]$ is connected and satisfies the conditions. From the inductive hypothesis we know that $O - V(L')$ has a perfect matching. This perfect matching together with the edge uv shows that L is nice. □

Thus we obtain the desired characterization:

Theorem 3.12. *If $\mathcal{H}$ consists of the $\leq k$-stars and any family $\mathcal{F}$ of factor-critical graphs, then the $\mathcal{H}$-critical graphs are the $\mathcal{F}$-oct-s. Consequently a maximum $\mathcal{H}$-factor of the graph G misses $\max_{X \subseteq V(G)}(oct_{\mathcal{F}}(G - X) - k|X|)$ vertices of the graph G.*

As a consequence we get the following theorem, which is the analogue of Theorem 3.3:

Theorem 3.13. *Let $k > 1$ be an integer. The graph G can be factorized by induced $\leq k$-stars and not necessarily induced circuits of length 5 if and only if for all $X \subseteq V(G)$ the graph $G - X$ has at most $k|X|$ $\triangle$-cactus components.*

If we allow that a two-star can occur non-induced in the factor, then the one vertex graph will be the only critical graph, because every non-trivial factor-critical graph has a nice two-star, namely two consecutive edges of a nice circuit will do. Hence we obtain the theorem of Las Vergnas [8], explicitly stated first by Hell and Kirkpatrick [5]:

Theorem 3.14. *G has a not necessarily induced $\leq k$-star-factor if and only if there is no $X \subseteq V(G)$, such that $G - X$ has more than $k|X|$ isolated vertices.*

We have developed a direct algorithm for solving the maximum $\mathcal{H}$-packing problem if $K_1 \in \mathcal{F}_i$ and $\mathcal{F}$ is fixed and finite. The detailed description of the algorithm will be found in the final version. This algorithm is much faster than the algorithm we get from Theorem 2.10. The difference comes from two tricks. On one hand, if $C = |\mathcal{F}|$ and c denote the maximum number of blocks in any graph $F \in \mathcal{F}$, then it is not hard to see, that checking whether an n-vertex graph is an $\mathcal{F}$-oct, can be done in time $O(Cn^c)$. On the other hand we do not *use* Edmonds' algorithm, we imitate its steps instead. It turns out, that for this case it becomes simpler. This is because, by Lemma 3.6, we need not contract odd cycles, all the ears of the ear-decomposition of an oct can be chosen to have two vertices.

Note that we may allow any graphs of $\mathcal{H}$ to occur non-induced in the factor, because a supergraph of a star is either factorizable by edges and by an induced star, or is factor-critical, and the same applies to the factor-critical graphs of $\mathcal{F}$. Though the size of $\mathcal{F}$ can increase exponentially, the running time of the algorithm remains the same.

On the basis of the algorithm an augmenting path theorem can be obtained as well, stating that if an $\mathcal{H}$-packing of G is not maximal then certain kind of augmenting paths exist. This is a generalization of a similar theorem of Kelmans [7].

The algorithm works also for a local version, where for each vertex v the maximum size of an induced star centered at v is prescribed.

References

[1] G. Cornuéjols, D. Hartvigsen, W. Pulleyblank Packing subgraphs in a graph *Oper. Res. Letter* (1981/82) **1**, no. 4, 139–143.

[2] G. Cornuéjols, D. Hartvigsen An Extension of Matching Theory *J. Combin. Theory Ser. B* (1986) **40**, no. 3, 285–296.

[3] G. Cornuéjols, W. Pulleyblank Critical graphs, matchings and tours or a hierarchy of relaxations for the traveling salesman problem *Combinatorica* (1983) **3**, no. 1, 35–52.

[4] G. Cornuéjols, W. Pulleyblank Perfect triangle-free 2-matchings *Combinatorial optimization II (Proc. Conf., Univ. East Anglia, Norwich, 1979). Math. Programming Stud.* (1980) **13**, 1–7.

[5] P. Hell, D. G. Kirkpatrick Packings by complete bipartite graphs *SIAM J. Algebraic Discrete Methods* (1986) **7**, no. 2, 199–209.

[6] A. Kaneko A necessary and sufficient condition for the existence of a path factor every component of which is a path of length at least two *submitted*

[7] A. K. Kelmans Optimal packing of induced stars in a graph *Discrete Math.* (1997) **173**, no. 1-3, 97–127.

[8] M. Las Vergnas An extension of Tutte's 1-factor theorem *Discrete Math.* (1978) **23**, no. 3, 241–255.

[9] L. Lovász, M. D. Plummer Matching Theory *Akadémiai Kiadó* (1986) ISBN 9630541688

[10] M. Loebl, S. Poljak Efficient Subgraph Packing *J. Combin. Theory Ser. B* (1993) **59**, no. 1, 106–121.

[11] M. Loebl, S. Poljak On Matroids Induced by Packing Subgraphs *J. Combin. Theory Ser. B* (1988) **44**, no. 3, 338–354.

[12] M. Loebl, S. Poljak Good Family Packing *Annals of Discrete Math* (1992) **51**, 181–186.

[13] J. Szabó The Generalized Kaneko Theorem *3rd Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications* (2003)

[14] W. T. Tutte The factors of graphs *Canadian J. Math.* (1952) **4**, 314–328.