A necessary and sufficient condition for the existence of a spanning tree with specified vertices having large degrees

Yoshimi Egawa

Department of Mathematical Information Science,
Tokyo University of Science,
Shinjuku-ku, Tokyo 162-8601, Japan

Kenta Ozeki

National Institute of Informatics,
2-1-2 Hitotsubashi, Chiyoda-ku,
Tokyo 101-8430, Japan
and
JST, ERATO, Kawarabayashi Large Graph Project
e-mail: ozeki@nii.ac.jp

Abstract

Let G be a connected simple graph, let $X \subseteq V(G)$ and let f be a mapping from X to the set of integers. When X is an independent set, Frank and Gyárfás, and independently, Kaneko and Yoshimoto gave a necessary and sufficient condition for the existence of a spanning tree T in G such that $d_T(x) \geq f(x)$ for all $x \in X$, where $d_T(x)$ is the degree of x in T. In this paper, we extend this result to the case where the subgraph induced by X has no induced path of order four, and prove that there exists a spanning tree T in G such that $d_T(x) \geq f(x)$ for all $x \in X$ if and only if for any nonempty subset $S \subseteq X$, $|N_G(S) - S| - f(S) + 2|S| - \omega_G(S) \geq 1$, where $\omega_G(S)$ is the number of components of the subgraph induced by S.

Keywords: Degree constrained spanning tree, Submodular function

1 Introduction

Throughout this paper, we consider only simple graphs.

Let G be a connected graph. For $x \in V(G)$, a vertex $v \in V(G) - \{x\}$ is called a *neighbor of* x *in* G if v is adjacent to x in G. We let $N_G(x)$ denote the set of neighbors of x in G. The degree of x is denoted by $d_G(x)$; thus $d_G(x) = |N_G(x)|$. For $S \subseteq V(G)$, let $N_G(S) := \bigcup_{x \in S} N_G(x)$ and $\Gamma_G(S) := N_G(S) - S$. Let $G[S]$ be the subgraph of G induced by S. We denote the number of components of $G[S]$ by $\omega_G(S)$. For $S, R \subseteq V(G)$, the set of edges in G connecting S and R is denoted by $E_G(S, R)$.

Let X be a subset of $V(G)$, and let f be a mapping from X to the set of integers. In this paper, we concentrate on the existence of a spanning tree T in G such that $d_T(x) \geq f(x)$ for all $x \in X$. For simplicity, we call such a spanning tree *an* (X, f)*-tree*. (In this context, the mapping f is referred to as a *lower capacity vector* or a *demand vector* in some papers.)

More specifically, we consider a necessary and sufficient condition for the existence of an (X, f)-tree. In general, it seems difficult to give such a condition. This is because the problem of determining whether a graph G has an (X, f)-tree or not is at least as hard as the problem of the existence of a Hamiltonian path, which is well-known as an NP-complete problem. Formally, when we want to determine whether a given graph G has a Hamiltonian path, we construct a graph G' by joining two new vertices to G, and let $X := V(G)$ and $f \equiv 2$. Then G has a Hamiltonian path if and only if G' has an (X, f)-tree.

However, when we confine ourselves to the case where X is an independent set, we can obtain a necessary and sufficient condition for the existence of an (X, f)-tree. Frank and Gyárfás, and independently, Kaneko and Yoshimoto gave the following result. For $S \subseteq X$, let $f(S) := \sum_{x \in S} f(x)$.

Theorem 1.1 (Frank and Gyárfás [1], Kaneko and Yoshimoto [2]) *Let G be a connected graph, let $X \subseteq V(G)$ and let f be a mapping from X to the set of integers. Then there exists an (X, f)-tree in G if and only if for any nonempty subset $S \subseteq X$,*

$$|N_G(S)| - f(S) + |S| \geq 1.$$

In the case where X is a non-independent set, we similarly obtain the following necessary condition for the existence of an (X, f)-tree. We prove Proposition 1.2 in Section 2. Let

$$g(S; G, f) := |\Gamma_G(S)| - f(S) + 2|S| - \omega_G(S).$$

Proposition 1.2 *Let G be a connected graph, let $X \subseteq V(G)$ and let f be a mapping from X to the set of integers. If there exists an (X, f)-tree in G, then for any nonempty subset $S \subseteq X$,*

$$g(S; G, f) \geq 1. \tag{1}$$

Note that for an independent set X, condition (1) is equivalent to the necessary condition in Theorem 1.1, because $\Gamma_G(S) = N_G(S) - S = N_G(S)$ and $\omega_G(S) = |S|$.

One might expect that condition (1) is also a sufficient condition. However, as mentioned above, it is impossible to have such a simple result (unless NP = co-NP). In fact, even when X induces a path consisting of four vertices, we can construct an example that satisfies condition (1) but has no (X, f)-tree as follows: let G_1 be the graph in Figure 1, let $X := \{x_1, x_2, x_3, x_4\}$, and define a mapping f by letting $f(x_i) = 2$ for each $1 \leq i \leq 4$. Note that for $S \subseteq X$, we have $|\Gamma_G(S)| \geq 2$ and $f(S) - 2|S| + \omega_G(S) = \omega_G(S) = 1$ if $|S| = 1$ or 4, and $|\Gamma_G(S)| \geq 3$ and $f(S) - 2|S| + \omega_G(S) = \omega_G(S) \leq 2$ if $|S| = 2$ or 3. Thus, for any nonempty subset $S \subseteq X$, $g(S; G, f) \geq 1$, but G_1 has no (X, f)-tree. More generally, when H is a graph having an induced path of order four, by considering a structure similar to G_1 in Figure 1, we can show the existence of an example of a graph G with $G[X] \cong H$ that satisfies condition (1) but has no (X, f)-tree.

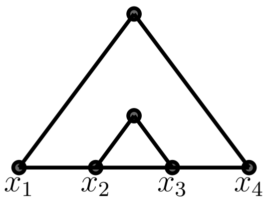

Figure 1: The graph G_1.

In view of the gap between the case of an independent set and the general case, it is natural to ask what causes the gap, in other words, what properties guarantee that condition (1) is also a sufficient condition. The main purpose of this paper is to give an answer to this question by showing the following result: when $G[X]$ has no induced path of order four, condition (1) is also a sufficient condition. As mentioned above, when X has an induced path of order four, there exists a graph satisfying condition (1) but has no (X, f)-tree. Thus the condition that "$G[X]$ has no induced path of order four" is best possible in this sense.

Theorem 1.3 *Let G be a connected graph and let $X \subseteq V(G)$, and suppose that $G[X]$ has no induced path of order four. Let f be a mapping from X to the set of integers. Then there exists an (X, f)-tree in G if and only if for any nonempty subset $S \subseteq X$,*

$$g(S; G, f) = \left|\Gamma_G(S)\right| - f(S) + 2|S| - \omega_G(S) \geq 1.$$

In Section 3, we show the submodularity of the function g, which plays a crucial role in our proof. After showing that, we prove Theorem 1.3 in Section 4.

2 A necessary condition for an (X, f)-tree

In this section, we prove Proposition 1.2. For reference in the proof of Theorem 1.3, we show a slightly stronger result, which includes a description of the case where equality holds. Before stating the result, we remark that under the assumption that G is connected, the existence of an (X, f)-tree is equivalent to the existence of a (not necessarily spanning) forest F with $X \subseteq V(F)$ such that $d_F(x) \geq f(x)$ for all $x \in X$.

Proposition 2.1 *Let G be a graph, let $X \subseteq V(G)$ and let f be a mapping from X to the set of integers. Suppose that there exists a forest F in G with $X \subseteq V(F)$ such that $d_F(x) \geq f(x)$ for all $x \in X$. Then for any nonempty subset $S \subseteq X$,*

$$g(S; G, f) = \big|\Gamma_G(S)\big| - f(S) + 2|S| - \omega_G(S) \geq 1.$$

Moreover, if equality holds for $S \subseteq X$, then $F[S \cup \Gamma_G(S)]$ is connected and $\Gamma_G(S) = \Gamma_F(S)$.

Proof of Proposition 2.1. Fix a subset $S \subseteq X$ and let l be the number of components of $F[S]$. Note that $l \geq \omega_G(S)$ and $\big|E(F[S])\big| = |S| - l$, because $F[S]$ has no cycle. Hence $\sum_{x \in S} \big|N_F(x) \cap S\big| = 2\big|E(F[S])\big| = 2\big(|S| - l\big)$.

Let H be the graph obtained from $F[S \cup \Gamma_F(S)]$ by contracting each component of $F[S]$ to one vertex. Then

$$\begin{aligned} |E(H)| &= \big|E_F\big(S, \Gamma_F(S)\big)\big| \\ &= \sum_{x \in S} \big|N_F(x)\big| - \sum_{x \in S} \big|N_F(x) \cap S\big| \\ &\geq f(S) - 2|S| + 2l. \end{aligned}$$

On the other hand, we obtain $|V(H)| = \big|\Gamma_F(S)\big| + l$. Since H is a forest, it follows from these inequalities that

$$\begin{aligned} \big|\Gamma_F(S)\big| + l &= |V(H)| \\ &\geq |E(H)| + 1 \\ &\geq f(S) - 2|S| + 2l + 1, \end{aligned}$$

or

$$\begin{aligned} g(S; G, f) &= \big|\Gamma_G(S)\big| - f(S) + 2|S| - \omega_G(S) \\ &\geq \big|\Gamma_F(S)\big| - f(S) + 2|S| - l \\ &\geq 1. \end{aligned}$$

Moreover when equality holds, H is connected and $\Gamma_G(S) = \Gamma_F(S)$, and hence $F[S \cup \Gamma_G(S)]$ is connected. □

3 Submodularity of the function g

In this section, we show the submodularity of the function g under the assumption that $G[X]$ has no induced path of order four. We actually prove the following slightly stronger result, which we need in the proof of Theorem 1.3. Throughout this section, we fix a graph G and a function f; so we simply write $\Gamma(S) := \Gamma_G(S)$, $\omega(S) := \omega_G(S)$ and $g(S) := g(S; G, f) = \big|\Gamma(S)\big| - f(S) + 2|S| - \omega(S)$.

Lemma 3.1 *Let G be a connected graph and let $X \subseteq V(G)$, and suppose that $G[X]$ has no induced path of order four. Let f be a mapping from X to the set of integers. Then for any $S, R \subseteq X$,*

$$g(S \cup R) + g(S \cap R) \le g(S) + g(R) - |A_{SR}| - \varepsilon_{SR},$$

where $A_{SR} := \Gamma(S) \cap \Gamma(R) - \Gamma(S \cap R)$,

and $$\varepsilon_{SR} := \begin{cases} 1 & \text{if } S \cap R = \emptyset \text{ and there exists an edge connecting } S \text{ and } R, \\ 0 & \text{otherwise.} \end{cases}$$

Before proving Lemma 3.1, we show that the assumption that $G[X]$ has no induced path of order four is needed for the submodularity of the function g. Let G_2 be the graph in Figure 2 and let $X := \{x_1, x_2, x_3, x_4\}$. Let $f(x_i) := 2$ for each $1 \le i \le 4$. Note that when $S := \{x_1, x_2, x_4\}$ and $R := \{x_1, x_3, x_4\}$, we have $g(S) = g(R) = g(S \cup R) = 1$ and $g(S \cap R) = 2$, so g does not satisfy submodularity.

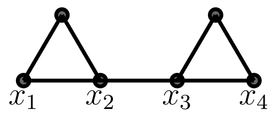

Figure 2: The graph G_2.

Proof of Lemma 3.1. Let

$$\begin{aligned} & B_S := \Gamma(S) \cap R - \Gamma(S \cap R), \\ \text{and} \quad & B_R := \Gamma(R) \cap S - \Gamma(S \cap R). \end{aligned}$$

Since

$$\begin{aligned} |\Gamma(S \cup R)| + |\Gamma(S \cap R)| &= |\Gamma(S)| + |\Gamma(R)| - |A_{SR}| - |B_S| - |B_R|, \\ f(S \cup R) + f(S \cap R) &= f(S) + f(R), \\ \text{and} \qquad |S \cup R| + |S \cap R| &= |S| + |R|, \end{aligned}$$

we obtain

$$\begin{aligned} & g(S) + g(R) - |A_{SR}| - \varepsilon_{SR} - g(S \cup R) - g(S \cap R) \\ &= \big(|\Gamma(S)| - f(S) + 2|S| - \omega(S)\big) + \big(|\Gamma(R)| - f(R) + 2|R| - \omega(R)\big) - |A_{SR}| - \varepsilon_{SR} \\ &\quad - \Big(|\Gamma(S \cup R)| - f(S \cup R) + 2|S \cup R| - \omega(S \cup R)\Big) \\ &\quad - \Big(|\Gamma(S \cap R)| - f(S \cap R) + 2|S \cap R| - \omega(S \cap R)\Big) \\ &= \omega(S \cup R) + \omega(S \cap R) + |B_S| + |B_R| - \omega(S) - \omega(R) - \varepsilon_{SR}. \end{aligned}$$

Thus it suffices to show that

$$\omega(S \cup R) + \omega(S \cap R) + |B_S| + |B_R| \geq \omega(S) + \omega(R) + \varepsilon_{SR}.$$

Now we construct two bipartite graphs $\tilde{H}$ and H as follows. One partite set of $\tilde{H}$ is $\mathcal{C}(S)$ and the other is $\mathcal{C}(R)$, where $\mathcal{C}(S)$ is the set of components of $G[S]$ and $\mathcal{C}(R)$ is that of $G[R]$. For $S_i \in \mathcal{C}(S)$ and for $R_j \in \mathcal{C}(R)$, we let $S_iR_j \in E(\tilde{H})$ if and only if (I) $S_i \cap R_j \neq \emptyset$ or (II) there exists an edge in G connecting S_i and R_j such that at least one of the end vertices is not contained in $\Gamma(S \cap R)$. Edges of $\tilde{H}$ satisfying (I) are said to be of *Type I*, and other edges, that is, edges not satisfying (I) and satisfying (II) are said to be of *Type II*. Let H be a spanning forest of $\tilde{H}$ such that the number of components of H is equal to that of $\tilde{H}$. Let $E_I := \{e \in E(H) : e \text{ is of Type I}\}$ and $E_{II} := \{e \in E(H) : e \text{ is of Type II}\}$. Note that $|V(H)| = \omega(S) + \omega(R)$.

Claim 1 *The number of components of H is at most $\omega(S \cup R)$.*

Proof. Since the number of components of H is equal to that of $\tilde{H}$, it suffices to show that the number of components of $\tilde{H}$ is at most $\omega(S \cup R)$.

By the definition of edges of $\tilde{H}$, each component of $\tilde{H}$ corresponds to some component of $G[S \cup R]$. Formally, if we let $S_1, \dots, S_p \in \mathcal{C}(S)$ and $R_1, \dots, R_q \in \mathcal{C}(R)$ be such that they constitute a component of $\tilde{H}$, $\bigcup_{i=1}^{p} S_i \cup \bigcup_{j=1}^{q} R_j$ is contained in a single component of $G[S \cup R]$; we consider the correspondence defined in this way. Now it suffices to show that no two components of $\tilde{H}$ are mapped to the same component of $G[S \cup R]$ by the correspondence.

Suppose that there exist two components of $\tilde{H}$ which correspond to the same component of $G[S \cup R]$. Again by the definition of edges of $\tilde{H}$, there exists an edge e in G connecting a component of $G[S]$, say S_i, and a component of $G[R]$, say R_j, such that S_i and R_j belong to distinct components of $\tilde{H}$ and both end vertices of e are contained in $\Gamma(S \cap R)$. Let x_1 be the end vertex of e contained in S_i and let x_2 be the other end vertex of e. Since $x_1, x_2 \in \Gamma(S \cap R)$, there exist two vertices $y_1, y_2 \in S \cap R$ such that $x_i \in N_G(y_i)$ for $i = 1, 2$. Note that $y_1 \neq y_2$ and $y_1x_2, y_1y_2, x_1y_2 \notin E(G)$, because $y_1, x_1 \in S_i$, $x_2, y_2 \in R_j$, and S_i and R_j belong to distinct components of $\tilde{H}$. Hence $\{y_1, x_1, x_2, y_2\}$ induces a path of order four, a contradiction. □

Claim 2 $|E_I| \leq \omega(S \cap R)$.

Proof. Let $e = S_iR_j$ be an edge of Type I. Then $S_i \cap R_j \neq \emptyset$. We associate with e a component of $G[S \cap R]$ contained in $S_i \cap R_j$. Note that $S_i \cap R_j$ intersects with no component in $\mathcal{C}(S) \cup \mathcal{C}(R) - \{S_i, R_j\}$. Thus, this correspondence is injective. Hence $|E_I| \leq \omega(S \cap R)$. □

Claim 3 $|E_{II}| \leq |B_S| + |B_R| - \varepsilon_{SR}$.

Proof. We show that there exists an injective mapping h from E_{II} to $B_S \cup B_R$; moreover, we show that when $S \cap R = \emptyset$ and there exists an edge in G connecting S and R, there exists $v \in B_S \cup B_R$ such that v is not contained in the image of h.

Let $e = S_i R_j \in E_{II}$. Then $S_i \cap R_j = \emptyset$, and there exists an edge in G connecting S_i and R_j such that at least one of its end vertices is not contained in $\Gamma(S \cap R)$. Let $\tilde{h}(e)$ be such an edge, and let $h(e)$ be an end vertex of $\tilde{h}(e)$ not contained in $\Gamma(S \cap R)$. Note that h is a mapping from E_{II} to $B_S \cup B_R$. Note also that even if we fix $\tilde{h}(e)$ for every e, h is not uniquely determined, because if neither end vertex of $\tilde{h}(e)$ is contained in $\Gamma(S \cap R)$, we may choose either of the two end vertices of $\tilde{h}(e)$ as $h(e)$. Choose h so that the image of h is as large as possible.

Suppose that h is not injective; that is, there exist two edges $e = S_i R_j$ and $e' = S_{i'} R_{j'}$ of Type II such that $h(e) = h(e')$. Let $u = h(e) = h(e')$, and let v and v' be the other end vertices of $\tilde{h}(e)$ and $\tilde{h}(e')$, respectively. By symmetry, we may assume that $u \in S$, so $S_i = S_{i'}$, $v \in R_j$ and $v' \in R_{j'}$.

Take a longest path $P = v_0 v_1 v_2 \dots v_p$ of G with $v_0 = u$ and $v_1 = v$ such that for any $0 \le l \le p-1$, $\tilde{h}(e_l) = v_l v_{l+1}$ and $h(e_l) = v_l$ for some $e_l \in E_{II}$. For convenience, let $v_{-1} = v'$. Note that $p \ge 1$, and $v_i \in R - S$ for any odd integer i with $1 \le i \le p$ and $v_i \in S - R$ for any even integer i with $1 \le i \le p$.

Assume for the moment that p is even. Suppose that $v_p \in \Gamma(S \cap R)$. Then there exists a vertex $w \in S \cap R$ such that $v_p \in N_G(w)$. Note that v_p and v_{p-2} are contained in distinct components of $G[S]$. Hence w and v_{p-2} are contained in distinct components of $G[S]$. Since $e_{p-1} \notin E_I$, we also see that w and v_{p-1} are contained in distinct components of $G[R]$. These imply that $\{w, v_p, v_{p-1}, v_{p-2}\}$ induces a path of order four, a contradiction. Thus $v_p \notin \Gamma(S \cap R)$. Suppose that $v_p = h(e'')$ for some $e'' \in E_{II}$, and let v_{p+1} be the other end vertex of $\tilde{h}(e'')$. Then $v_{p+1} \ne v_l$ for any $-1 \le l \le p$ because H has no cycle. Hence $v_0 v_1 v_2 \dots v_p v_{p+1}$ is a path longer than P, a contradiction. Thus, $v_p \ne h(e'')$ for any $e'' \in E_{II}$, and hence by redefining $h(e_l) = v_{l+1}$ for $0 \le l \le p-1$, we get a contradiction to the maximality of the image of h. When p is odd, we can similarly get a contradiction.

Consequently, h is an injective mapping from E_{II} to $B_S \cup B_R$. Now we consider the case where $S \cap R = \emptyset$ and there exists an edge in G connecting S and R. Note that $\Gamma(S \cap R) = \emptyset$ and $E_I = \emptyset$, and hence $E_{II} \ne \emptyset$ by the assumption that $E_G(S, R) \ne \emptyset$. By the above argument, we can find an injective mapping h from E_{II} to $B_S \cup B_R$. Take $e = S_i R_j \in E_{II}$. Since $\Gamma(S \cap R) = \emptyset$, neither end vertex of $\tilde{h}(e)$ is contained in $\Gamma(S \cap R)$. Let $u = h(e)$ and let v be the other end vertex of $\tilde{h}(e)$. Again take a longest path $P = v_0 v_1 v_2 \dots v_p$ of G with $v_0 = u$ and $v_1 = v$ such that for any $0 \le l \le p-1$, $\tilde{h}(e_l) = v_l v_{l+1}$ and $h(e_l) = v_l$ for some $e_l \in E_{II}$. Since $\Gamma(S \cap R) = \emptyset$, we clearly have $v_p \in B_S \cup B_R$. By the same argument as above, we can also show that v_p is not contained in the image of h. This implies that $|E_{II}| \le |B_S| + |B_R| - \varepsilon_{SR}$.

□

Note that $|V(H)| = \omega(S)+\omega(R)$. Thus by Claim 1, $|E(H)| \geq |V(H)|-\omega(S\cup R) = \omega(S) + \omega(R) - \omega(S \cup R)$. Therefore it follows from Claims 2 and 3 that

$$\begin{aligned} \omega(S) + \omega(R) - \omega(S \cup R) &\leq |E(H)| \\ &\leq \omega(S \cap R) + |B_S| + |B_R| - \varepsilon_{SR}, \end{aligned}$$

or $$\omega(S \cup R) + \omega(S \cap R) + |B_S| + |B_R| \geq \omega(S) + \omega(R) + \varepsilon_{SR}.$$

This completes the proof of Lemma 3.1.

4 Proof of Theorem 1.3

In this section, we will prove Theorem 1.3 but, before proving the theorem, we mention an algorithmic aspect. As mentioned before, the submodularity of the function g plays a key role in our proof, and it is well-known that the problem of minimizing a submodular function can be solved in polynomial time. (For example, see Theorem 45.1 on Page 791 of [3].) Based on this fact and the following nature of our proof, we believe that there is a polynomial time algorithm to find an (X, f)-tree in a graph satisfying condition (1). Note that at each step in the proof, for a given graph G, a given set $X \subseteq V(G)$ such that $G[X]$ has no induced path of order four, and a given mapping f, we delete a vertex from X or an edge from G. To find an appropriate vertex or edge, we make use of a "tight set" (see the definition made immediately before Claim 4), which can be found by using the problem of minimizing the submodular function g.

In order to prove Theorem 1.3, it suffices to show the following theorem (see the remark made immediately before the statement of Proposition 2.1).

Theorem 4.1 *Let G be a graph and let $X \subseteq V(G)$, and suppose that $G[X]$ has no induced path of order four. Let f be a mapping from X to the set of integers, and suppose that for any nonempty subset $S \subseteq X$,*

$$g(S, G, f) \geq 1.$$

Then there exists a forest F in G with $X \subseteq V(F) \subseteq X \cup \Gamma_G(X)$ and $E(F) \subseteq E(G[X]) \cup E_G\big(X, \Gamma_G(X)\big)$ such that $d_F(x) \geq f(x)$ for all $x \in X$.

Proof of Theorem 4.1. Since removing an edge connecting two vertices of $V(G)-X$ never destroys the assumption of Theorem 4.1, we may assume that $V(G) - X$ is an independent set. We prove Theorem 4.1 by simultaneous induction on $|X|$ and $\big|E_G\big(X, \Gamma_G(X)\big)\big|$. If $|X| = 0$, there is nothing to prove; so we may assume that $X \neq \emptyset$.

Suppose that G has no forest satisfying the properties in Theorem 4.1. We first show that $\big|E_G\big(X, \Gamma_G(X)\big)\big| \neq 0$. Suppose that $\big|E_G\big(X, \Gamma_G(X)\big)\big| = 0$. This implies that $\Gamma_G(X) = \emptyset$, and hence $g(X; G, f) = -f(X) + 2|X| - \omega_G(X) \geq 1$, or $f(X) \leq 2|X| - \omega_G(X) - 1 \leq 2|X| - 2$. Consequently, there exists a vertex $x \in X$ with $f(x) \leq 1$. By the induction hypothesis for $X' := X - \{x\}$, G has a forest F' with $E(F') \subseteq E(G[X']) \cup E_G\big(X', \Gamma_G(X')\big)$ such that $d_{F'}(x') \geq f(x')$ for all $x' \in X'$. When $f(x) \leq 0$ or $d_{F'}(x) \geq 1$, F' is a forest as desired. On the other hand, when $f(x) = 1$ and $x \notin V(F')$, $F := F' \cup \{xu\}$ is a forest as desired, where $u \in N_G(x)$ (note that there exists such a vertex u because $d_G(x) = |\Gamma_G(x)| \geq 1 + f(x) - 2 + \omega_G(\{x\}) = 1$). This contradicts the assumption that G has no forest satisfying the properties in Theorem 4.1. Thus, $\big|E_G\big(X, \Gamma_G(X)\big)\big| \neq 0$.

For a nonempty subset $S \subseteq X$, we say that S is a *tight set* if $g(S; G, f) = 1$. Let

$$\mathcal{S} := \{S \subseteq X : S \text{ is a tight set with } S \neq X\}.$$

Claim 4 *For any $x \in X - \bigcup_{S \in \mathcal{S}} S$, $N_G(x) - X = \emptyset$.*

Suppose that there exists a vertex $x \in X - \bigcup_{S \in \mathcal{S}} S$ such that $N_G(x) - X \neq \emptyset$, say $u \in N_G(x) - X$. Since $x \notin S$ for any $S \in \mathcal{S}$, there exists no tight set containing x except for X. Hence for any nonempty subset $R \subseteq X$ with $x \in R$ and $R \neq X$, $g(R; G, f) \geq 2$. If u is a neighbor of some vertex in $X - \{x\}$, then let $f' = f$; otherwise let

$$f'(y) := \begin{cases} f(y) - 1 & \text{if } y = x, \\ f(y) & \text{otherwise.} \end{cases}$$

Let G' be the graph obtained from G by deleting the edge xu. Then $\big|E_{G'}\big(X, \Gamma_{G'}(X)\big)\big| < \big|E_G\big(X, \Gamma_G(X)\big)\big|$. For any nonempty subset $R \subseteq X$ with $x \in R$ and $R \neq X$, $\Gamma_{G'}(R) \supseteq \Gamma_G(R) - \{u\}$, and hence $g(R; G', f') \geq g(R; G', f) \geq g(R; G, f) - 1 \geq 1$. For any nonempty subset $R \subseteq X$ with $x \notin R$, $f'(R) = f(R)$ and $\Gamma_{G'}(R) = \Gamma_G(R)$, and hence $g(R; G', f') = g(R; G, f) \geq 1$.

When u is a neighbor of some vertex in $X - \{x\}$, $\Gamma_{G'}(X) = \Gamma_G(X)$, and hence $g(X; G', f') = g(X; G, f) \geq 1$; otherwise $|\Gamma_{G'}(X)| = |\Gamma_G(X)| - 1$ and $f'(X) = f(X) - 1$, and hence $g(X; G', f') = g(X; G, f) \geq 1$. Thus, in each case, we obtain $g(X; G', f') \geq 1$. Therefore G', X and f' satisfy the assumption of Theorem 4.1. By the induction hypothesis, G' has a forest F' with $E(F') \subseteq E(G'[X]) \cup E_{G'}\big(X, \Gamma_{G'}(X)\big)$ such that $d_{F'}(x') \geq f'(x')$ for all $x' \in X$. When u is a neighbor of some vertex in $X - \{x\}$, F' is also a forest as desired for G, X and f; otherwise the graph F obtained from F' by adding the edge xu is a forest (because u is no neighbor of $X - \{x\}$), and hence F is a forest as desired. This contradiction completes the proof of the claim. □

Claim 5 *Let $S \in \mathcal{S}$. Then the following hold.*

(i) *There exists a tree* F_S *with* $V(F_S) = S \cup \Gamma_G(S)$ *and* $E(F_S) \subseteq E(G[S]) \cup E_G(S, \Gamma_G(S))$ *such that* $d_{F_S}(x) \geq f(x)$ *for all* $x \in S$.

(ii) *Let* F_S *be as in (i), and suppose that* F_S *is chosen so that* $\big|E(F_S) \cap E_G\big(S, \Gamma_G(X)\big)\big|$ *is as small as possible. Let* D *be a component of* $G[X]$ *with* $S \cap V(D) \neq \emptyset$. *Then* $F_S\big[\big(S \cup \Gamma_G(S)\big) \cap V(D)\big]$ *is connected.*

(iii) *Let* F_S *be as in (ii). Then* $E_G\big(S, \Gamma_G(X)\big) \subseteq E(F_S)$.

Proof. (i) Since $|S| < |X|$, we can apply the induction hypothesis to G, S and f, to obtain a forest F_S with $V(F_S) \subseteq S \cup \Gamma_G(S)$ and $E(F_S) \subseteq E(G[S]) \cup E_G\big(S, \Gamma_G(S)\big)$ such that $d_{F_S}(x) \geq f(x)$ for all $x \in S$. By Proposition 2.1, F_S is a tree and $\Gamma_G(S) \subseteq V(F_S)$. Thus, $V(F_S) = S \cup \Gamma_G(S)$. $\square$

(ii) We take a tree F_S as in (i) so that $\big|E(F_S) \cap E_G\big(S, \Gamma_G(X)\big)\big|$ is as small as possible. Suppose that $F_S[\big(S \cup \Gamma_G(S)\big) \cap V(D)]$ has at least two components, say C_1 and C_2. Since F_S is connected, there exists a path P connecting C_1 and C_2 in F_S with $|V(P) \cap V(C_1)| = 1$ and $|V(P) \cap V(C_2)| = 1$. For each $i = 1, 2$, write $V(P) \cap V(C_i) = \{x_i\}$ and $N_P(x_i) = \{u_i\}$. Note that $u_1, u_2 \notin X$.

Since $G[X]$ has no induced path of order four, we have $x_1x_2 \in E(G)$ or there exists a vertex $y \in X - S$ such that $y \in N_G(x_1) \cap N_G(x_2)$. In the second case, y cannot reach x_1 in F_S without passing through u_1 or cannot reach x_2 without passing through u_2, since otherwise, these two paths and P together contain at least one cycle, a contradiction. Thus, by symmetry, we may assume that y cannot reach x_1 in F_S without passing through u_1. In either case, let F'_S be the graph obtained from F_S by deleting x_1u_1, and adding x_1x_2 or x_1y. Then $d_{F'_S}(x) \geq f(x)$ for all $x \in X$ and $\big|E(F'_S) \cap E_G\big(S, \Gamma_G(X)\big)\big|$ is smaller than $\big|E(F_S) \cap E_G\big(S, \Gamma_G(X)\big)\big|$, contradicting the choice of F_S. $\square$

(iii) Let G_S be the graph obtained from G by deleting all edges in $E_G\big(S, \Gamma_G(X)\big)$ not contained in $E(F_S)$. We will show that for G_S, X and f, the assumption of Theorem 4.1 holds. Take an arbitrary nonempty subset $R \subseteq X$. If $S \cap R = \emptyset$, then $\Gamma_{G_S}(R) = \Gamma_G(R)$, and hence $g(R; G_S, f) = g(R; G, f) \geq 1$. Thus, we may assume that $S \cap R \neq \emptyset$. Since the tree F_S satisfies $d_{F_S}(x) \geq f(x)$ for all $x \in S$, it follows from Proposition 2.1 that $g(S \cap R; G_S, f) \geq 1$. On the other hand, $\Gamma_{G_S}(S) = \Gamma_{F_S}(S) = \Gamma_G(S)$, and hence we have $g(S \cup R; G_S, f) = g(S \cup R; G, f) \geq 1$ and $g(S; G_S, f) = g(S; G, f) = 1$. These inequalities and Lemma 3.1 imply that $g(R; G_S, f) \geq g(S \cup R; G_S, f) + g(S \cap R; G_S, f) - g(S; G_S, f) \geq 1$. Therefore G_S, X and f satisfy the assumption of Theorem 4.1. Hence if there exists at least one edge in $E_G\big(S, \Gamma_G(X)\big)$ not contained in $E(F_S)$, we can use the induction hypothesis, and obtain a forest as desired, a contradiction. Thus, all edges in $E_G\big(S, \Gamma_G(X)\big)$ are used in F_S. $\square$

For each $S \in \mathcal{S}$, we let F_S be a tree as in Claim 5 (i) and we assume that F_S is chosen so that $\big|E(F_S) \cap E_G\big(S, \Gamma_G(X)\big)\big|$ is as small as possible. By Claim 5, we obtain the following claim.

Claim 6 *Let $S \in \mathcal{S}$. Then there exists no cycle in $G[S \cup \Gamma_G(S)] - E_G\big(X - S, \Gamma_G(X)\big)$ containing at least one vertex in $\Gamma_G(X)$.*

Proof. Suppose that there exists a cycle C in $G[S \cup \Gamma_G(S)] - E_G\big(X - S, \Gamma_G(X)\big)$ containing at least one vertex in $\Gamma_G(X)$. Let $E_0 := E(C) \cap E_G\big(\Gamma_G(X), V(G)\big)$. Since $\Gamma_G(X)$ is independent and C does not use an edge in $E_G\big(\Gamma_G(X), X - S\big)$, we obtain $E_0 \subseteq E_G\big(S, \Gamma_G(X)\big)$. Hence $E_0 \subseteq E(F_S)$ by Claim 5 (iii).

On the other hand, for a component of D of $G[X]$, let $E_D := E(F_S[V(D)])$ if $V(C) \cap V(D) \neq \emptyset$; otherwise let $E_D = \emptyset$. By the definition of E_0 and E_D and by Claim 5 (ii), $E_0 \cup \bigcup_{D \in \mathcal{C}(X)} E_D$ contains at least one cycle, where $\mathcal{C}(X)$ is the set of components of $G[X]$. This contradicts the fact that F_S has no cycle. $\square$

Claim 7 *Let $S \in \mathcal{S}$ and let D be a component of $G[X]$ with $S \cap V(D) \neq \emptyset$. Then $S \cap V(D) \in \mathcal{S}$.*

Proof. If $S \subseteq V(D)$, then there is nothing to prove. Thus we may assume that S intersects with at least two components of $G[X]$. Let D_1, D_2, …, D_l be the components of $G[X]$ such that $S \cap V(D_i) \neq \emptyset$, and let $S_i = S \cap V(D_i)$. If we have $\big|\Gamma_G(S_i) \cap \bigcup_{j \neq i} \Gamma_G(S_j)\big| \geq 2$ for every $1 \leq i \leq l$, then $\bigcup_{i=1}^{l} \big(S_i \cup \Gamma_G(S_i)\big)$ contains a cycle, contradicting Claim 6. Hence there exists an index i, say $i = 1$, such that $\big|\Gamma_G(S_1) \cap \bigcup_{j=2}^{l} \Gamma_G(S_j)\big| \leq 1$. Let $\bar{S} := \bigcup_{j=2}^{l} S_j$. Then

$$
\begin{aligned}
g(S; G, f) &= \big|\Gamma_G(S)\big| - f(S) + 2|S| - \omega_G(S) \\
&= \big|\Gamma_G(S_1)\big| + \big|\Gamma_G(\bar{S})\big| - \big|\Gamma_G(S_1) \cap \Gamma_G(\bar{S})\big| \\
&\quad - f(S_1) - f(\bar{S}) + 2|S_1| + 2|\bar{S}| - \omega_G(S_1) - \omega_G(\bar{S}) \\
&\geq g(S_1; G, f) + g(\bar{S}; G, f) - 1.
\end{aligned}
$$

Since $g(S; G, f) = 1$, $g(S_1; G, f) \geq 1$ and $g(\bar{S}; G, f) \geq 1$, we obtain $g(S_1; G, f) = 1$ and $g(\bar{S}; G, f) = 1$, and hence S_1, $\bar{S} \in \mathcal{S}$. By applying the above argument recursively, we can prove that $S \cap V(D_i) \in \mathcal{S}$ for all $1 \leq i \leq l$. $\square$

Claim 8 *There exists no cycle containing at least one vertex in $\Gamma_G(X)$.*

Proof. Suppose that there exists a cycle C containing at least one vertex in $\Gamma_G(X)$. We choose C so that $|V(C) \cap \Gamma_G(X)|$ is as small as possible. Let $\{S_1, S_2, \ldots, S_k\}$ be a family of sets in $\mathcal{S}$ such that $V(C) \cap S_i \neq \emptyset$ for each $1 \leq i \leq k$ and $V(C) \cap \bigcup_{S \in \mathcal{S}} S \subseteq \bigcup_{i=1}^{k} S_i$. We call such a family $\{S_1, S_2, \ldots, S_k\}$ *a covering family of C*. By Claim 7, we may assume that each S_i is contained in one component of $G[X]$. By Claim 4, we have $k \geq 1$.

Let $u \in V(C) \cap \Gamma_G(X)$. If $|N_G(u) \cap S_i| \geq 2$ for some i, then we can find a cycle in $G[S_i \cup \Gamma_G(S_i)] - E_G\big(X - S_i, \Gamma_G(X)\big)$ containing u, contradicting Claim 6.

Thus, $|N_G(u) \cap S_i| \leq 1$ for each $1 \leq i \leq k$. In particular, we have $k \geq 2$, because $|N_G(u)| \geq |N_C(u)| = 2$ and $N_G(u) \subseteq \bigcup_{i=1}^{k} S_i$ by Claim 4.

Suppose that $G[X]$ consists of only one component. Since $|N_G(u)| \geq 2$ and $|N_G(u) \cap S_i| \leq 1$ for each $1 \leq i \leq k$, there exist two sets S_i and S_j, say S_1 and S_2, such that $N_G(u) \cap S_1 \neq \emptyset$, $N_G(u) \cap S_2 \neq \emptyset$ and $N_G(u) \cap S_1 \cap S_2 = \emptyset$. Thus $|A_{S_1 S_2}| \geq 1$, where $A_{S_1 S_2} := \Gamma_G(S_1) \cap \Gamma_G(S_2) - \Gamma_G(S_1 \cap S_2)$. On the other hand, since S_1 and S_2 are contained in the same component of $G[X]$ and $G[X]$ has no induced path of order four, one of the following holds:

(i) $S_1 \cap S_2 \neq \emptyset$;

(ii) $\Gamma_G(S_1) \cap \Gamma_G(S_2) \cap X \neq \emptyset$; or

(iii) there exists an edge connecting S_1 and S_2.

If (i) holds, $g(S_1 \cap S_2; G, f) \geq 1$. On the other hand, when (i) does not hold, we have $|A_{S_1 S_2}| \geq 2$ if (ii) holds, and $\varepsilon_{S_1 S_2} = 1$ if (iii) holds. In any case, it follows from Lemma 3.1 that

$$\begin{aligned} g(S_1 \cup S_2; G, f) &\leq g(S_1; G, f) + g(S_2; G, f) - g(S_1 \cap S_2; G, f) - |A_{S_1 S_2}| - \varepsilon_{S_1 S_2} \\ &\leq 0, \end{aligned}$$

a contradiction. Therefore $G[X]$ has at least two components.

Suppose that there exists a component of $G[X]$ containing two sets S_i and S_j, say S_1 and S_2. Since $G[X]$ has no induced path of order four, $S_1 \cap S_2 \neq \emptyset$, or $\Gamma_G(S_1) \cap \Gamma_G(S_2) \neq \emptyset$, or there exists an edge connecting S_1 and S_2. This means that $g(S_1 \cap S_2; G, f) \geq 1$, or $|A_{S_1 S_2}| \geq 1$, or $\varepsilon_{S_1 S_2} = 1$. Consequently we obtain

$$\begin{aligned} g(S_1 \cup S_2; G, f) &\leq g(S_1; G, f) + g(S_2; G, f) - g(S_1 \cap S_2; G, f) - |A_{S_1 S_2}| - \varepsilon_{S_1 S_2} \\ &\leq 1, \end{aligned}$$

and hence $S_1 \cup S_2$ is also a tight set. Since $G[X]$ has at least two components and S_1 and S_2 are contained in the same component of $G[X]$, we have $S_1 \cup S_2 \neq X$, and hence $S_1 \cup S_2 \in \mathcal{S}$. This implies that $\{S_1 \cup S_2, S_3, \ldots, S_k\}$ is also a covering family of C. By applying the above argument recursively, we may assume that any component of $G[X]$ contains at most one set in $\{S_1, S_2, \ldots, S_k\}$. This choice implies that $S_i \cap S_j = \emptyset$ for any $1 \leq i < j \leq k$. Since $\{S_1, S_2, \ldots, S_k\}$ is a covering family of C, changing the order of $S_1, S_2, \ldots, S_k$ if necessary, we may also assume that there exist k distinct vertices $u_1, u_2, \ldots, u_k \in \Gamma_G(X)$ such that $u_i \in \Gamma_G(S_i) \cap \Gamma_G(S_{i+1})$ for each $1 \leq i \leq k$, where $S_{k+1} = S_1$ (recall that we have chosen C so that $|V(C) \cap \Gamma_G(X)|$ is minimum).

For $1 \leq i \leq k-1$, let $R_i := \bigcup_{j=1}^{i} S_j$. By inductive argument, we will show that $R_i \in \mathcal{S}$ for each $1 \leq i \leq k-1$. When $i = 1$, definitely $R_1 = S_1 \in \mathcal{S}$. Thus let $i \geq 2$ and suppose that $R_{i-1} \in \mathcal{S}$. By the properties of $\{S_1, S_2, \ldots S_k\}$ mentioned at the end of the preceding paragraph, we have $R_{i-1} \cap S_i = \emptyset$ and $\Gamma_G(R_{i-1}) \cap \Gamma_G(S_i) \neq \emptyset$. Thus, $|A_{R_{i-1}S_i}| \geq 1$, where $A_{R_{i-1}S_i} := \Gamma_G(R_{i-1}) \cap \Gamma_G(S_i) - \Gamma_G(R_{i-1} \cap S_i)$, and hence it follows from Lemma 3.1 that $g(R_i; G, f) \leq g(R_{i-1}; G, f) + g(S_i; G, f) - |A_{R_{i-1}S_i}| \leq 1$. Since $g(R_i; G, f) \geq 1$, we obtain $g(R_i; G, f) = 1$. Hence $R_i \in \mathcal{S}$.

In particular, $R_{k-1} \in \mathcal{S}$. Moreover, $R_{k-1} \cap S_k = \emptyset$ and $|A_{R_{k-1}S_k}| \geq 2$, where $A_{R_{k-1}S_k} := \Gamma_G(R_{k-1}) \cap \Gamma_G(S_k) - \Gamma_G(R_{k-1} \cap S_k)$. Again by Lemma 3.1, this implies that $g(R_{k-1} \cup S_k; G, f) \leq g(R_{k-1}; G, f) + g(S_k; G, f) - |A_{R_{k-1}S_k}| \leq 0$, contradicting the assumption. $\square$

We are now in a position to complete the proof of Theorem 4.1. Since $\left|E_G(X, \Gamma_G(X))\right| \neq 0$, there exists an edge $xu \in E(G)$ with $x \in X$ and $u \in \Gamma_G(X)$. Let

$$f'(y) := \begin{cases} f(y) - 1 & \text{if } y = x, \\ f(y) & \text{otherwise.} \end{cases}$$

Let G' be the graph obtained from G by deleting the edge xu. Then $\left|E_{G'}\big(X, \Gamma_G(X)\big)\right| < \left|E_G\big(X, \Gamma(X)\big)\right|$. For any nonempty subset $R \subseteq X$ with $x \in R$, $\Gamma_{G'}(R) \supseteq \Gamma_G(R) - \{u\}$ and $f'(R) = f(R) - 1$, and hence $g(R; G', f') = g(R; G', f) + 1 \geq g(R; G, f) \geq 1$. For any nonempty subset $R \subseteq X$ with $x \notin R$, $\Gamma_{G'}(R) = \Gamma_G(R)$ and $f'(R) = f(R)$, and hence $g(R; G', f') = g(R; G', f) = g(R; G, f) \geq 1$. Therefore G', X and f' satisfy the assumption of Theorem 4.1, and by the induction hypothesis, G' has a forest F' with $E(F') \subseteq E(G'[X]) \cup E_{G'}\big(X, \Gamma_{G'}(X)\big)$ such that $d_{F'}(x') \geq f'(x')$ for all $x' \in X$. By Claim 8, the graph obtained from F' by adding the edge xu is a forest satisfying the condition required in Theorem 4.1, which contradicts the assumption that there is no such forest. This completes the proof of Theorem 4.1. $\square$

Acknowledgments

The authors are grateful to the referee for useful comments.

References

[1] A. Frank and A. Gyárfás, How to orient the edges of a graph? *Colloq. Math. Soc. Jànos Bolyai* **18** (1976) 353–364.

[2] A. Kaneko and K. Yoshimoto, On spanning trees with restricted degrees, *Inform. Process. Lett.* **73** (2000) 163–165.

[3] A. Schrijver, Combinatorial Optimization, *Springer* 2003.