



HAL
open science

Connected subgraphs, matching, and partitions

Mohamed Didi Biha, Hervé L M Kerivin

► **To cite this version:**

Mohamed Didi Biha, Hervé L M Kerivin. Connected subgraphs, matching, and partitions. 2023.
hal-03963095

HAL Id: hal-03963095

<https://hal.science/hal-03963095>

Preprint submitted on 30 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Connected subgraphs, matching, and partitions

Mohamed Didi Biha ^{*} Hervé L.M. Kerivin [†]

Abstract

Given an undirected graph G and a real edge-weight vector, the connected subgraph problem consists of finding a maximum-weight subset of edges which induces a connected subgraph of G . In this paper, we establish a link between the complexity of the connected subgraph problem and the matching number. We study the separation problem associated with the Matching-partition inequalities which are introduced by Didi Biha et al. [4] for the connected subgraph polytope.

keywords connected subgraph, polytope, matching, partition, separation

1 Introduction

Let $G = (V, E)$ be a simple, connected, and undirected graph having at least one edge. A non-empty edge subset $F \subseteq E$ induces a *connected subgraph* of G if the subgraph $G[F] := (V[F], F)$ of G induced by F is connected, where $V[F]$ corresponds to all the vertices in V incident with an edge in F . Given an edge-weight vector $\mathbf{w} \in \mathbb{R}^E$, not necessarily non-negative, the *Maximum-Weight Connected-Subgraph Problem (MWCSP)* consists of finding a non-empty subset $F \subseteq E$ of edges so that the subgraph $G[F]$ is connected and $\mathbf{w}(F) = \sum_{e \in F} w_e$ is maximum.

^{*}Université de Caen Normandie, LMNO - CNRS UMR 6139, F-14000 Caen, France. Email: mohamed.didi-biha@unicaen.fr

[†]LIMOS - CNRS UMR 6158, Université Blaise Pascal, Clermont-Ferrand 2, 63177 Aubière Cedex, France. Email: kerivin@isima.fr

MWCSP was first considered by Kerivin and Ng [18] who focused on its complexity. They showed that it is a new NP-hard combinatorial optimization problem even on planar or bipartite graphs; they proceeded by reducing the Steiner tree problem to MWCSP. Kerivin and Ng [18] also proved that the MWCSP and the prize-collecting Steiner tree problem, as defined in Johnson et al. [16] and in Goemans and Williamson [11], are equivalent optimization problems. Using a similar approach as Feigenbaum et al. [9], they showed that it is NP-hard to approximate MWCSP to within a constant factor.

MWCSP may appear as an underlying problem in many optimization problems. As an illustration of the practical interest of MWCSP, one may mention a peer-to-peer (P2P) video-streaming problem in under-provisioned networks [2]. From a combinatorial-optimization viewpoint, this P2P problem consists of finding a packing of rooted trees (i.e., independent delivery trees from a specific source peer) which spans as many vertices (i.e., peers) as possible while satisfying some bounds on the cumulative vertex's degrees (i.e., the upload capacity of each peer is not exceeded). Since one does not know a priori the subsets of vertices spanned by each rooted tree, the problem aims at finding a packing of acyclic connected subgraphs spanning the source vertex.

The essence of MWCSP consists of having to simultaneously deal with positive and negative edge weights. This matter makes MWCSP be a nontrivial optimization problem on classes of graphs where the Steiner tree problem is trivial. For instance, the optimal solutions to STP on trees correspond to the subtrees which are induced by all the edges belonging to paths between terminal nodes. On a cycle C , solving STP reduces to picking out the minimum-weight paths among the $|T|$ paths obtained from C by removing all the edges between two consecutive terminal vertices. In [18], Kerivin and Ng derived quadratic algorithms, based on the purposely-introduced prize-collecting connected-subgraph problem, to solve MWCSP on trees and cycles. In [4], Didi Biha, Kerivin and Pe studied the connected subgraph polytope. They strengthened a cut-based formulation by considering some partition inequalities and studied the separation problem associated with these inequalities. They also gave a complete polyhedral characterization of the connected subgraph polytope on cycles and trees.

This article is organized as follows. In Section 2, we establish the link, in terms of complexity, between the MWCSPP and the matching number. Section 3 present the *matching-partition inequalities* which are introduced in [4] for the so-called Connected Subgraph Polytope. In Section 4, we investigate the separation problem associated with the Matching-partition inequalities. Finally, some concluding remarks are given in Section 5.

We conclude this introduction with some definition and notation, which have been mainly taken from [5] and [21].

Let $G = (V, E)$ be a simple, connected, and undirected graph. The *order* n of G is its number of vertices, that is, $n := |V|$. The number of edges of G is denoted by m . If $e \in E$ is an edge with extremities u and v , we also write uv to denote e .

A path, cycle, and, complete graph of order n are denoted P_n , C_n , and K_n , respectively.

Let U be a subset of V . The set of edges having one extremity in U and the other one in $\bar{U} := V \setminus U$ is called a *cut* and is denoted by $\delta(U)$. If $U = \{v\}$ for some $v \in V$, then we write $\delta(v)$ for $\delta(\{v\})$. We denote by $E[U]$ the set of edges having both extremities in U and $G[U]$ the subgraph induced by U (i.e., $G[U] = (U, E[U])$). Given $W \subset V$ with $W \cap U = \emptyset$, $[U, W]$ denotes the set of edges having one extremity in U and the other one in W . If $\pi = \{V_1, \dots, V_p\}$, $p \geq 2$, is a partition of V , then we denote by $E(\pi)$ the set of edges having their extremities in different classes of π . We may also write $\delta(V_1, \dots, V_p)$ for $E(\pi)$.

Let $F \subseteq E$. Given $\mathbf{x} \in R^E$, $\mathbf{x}(F)$ will denote $\sum_{e \in F} x(e)$.

2 MWCP and matching number

In the next theorem, we will show that MWCP is polynomial when the matching number of G (i.e., maximum size of a matching in G) is bounded by a fixed integer k . In order to do that, we will prove that MWCP can be solved using a sequence of polynomial time Steiner tree problems.

Let $G = (V, E)$ be a graph and weights $w(e)$ associated with the edges $e \in E$. Given a set of p distinguished vertices $S = \{s_1, \dots, s_p\} \subset V$, called *terminals*, the

Steiner tree problem (STP) is to find a minimum weight tree spanning S . Although this problem is NP-hard in general, it is solvable in polynomial time when p is fixed [6]

Theorem 1. *Let $G = (V, E)$ be a graph and an weight function w on E that associates with an edge $e \in E$ the weight $W(e) \in \mathbb{R}$. Let $k \geq 1$ be a fixed integer. If G has a matching number of value at most k , then MXCP is solvable in polynomial time.*

Proof. Let $E^- = \{e \in E : w(e) < 0\}$, $G^+ = (V, E \setminus E^-)$ and $G_1^+ = (V_1, E_1^+), \dots, G_q^+ = (V_q, E_q^+)$ be the non-trivial connected components of G^+ . Since the matching number of G is at most k , we have $q \leq k$.

Let $\bar{G} = (\bar{V}, \bar{E})$ be the graph obtained from G by contracting every subset vertex V_i into t_i , $i = 1, \dots, q$, adding q new vertices s_1, \dots, s_q and q new edges $e_i = (s_i t_i)$. If F is a solution of MWCP, then we have $E_i^+ \cap F = \emptyset$ or $E_i^+ \cap F = E_i^+$ for $i = 1, \dots, q$. Also it is easy to see that if $e = (uv) \in E^-$ such that $u, v \in V_i$ for some i , then $e \notin F$. Thus, MWCP can be solved by solving for every fixed $S \subset \{s_1, \dots, s_q\}$ the instance $I(S) = (\bar{G}, S, \bar{w})$ of STP where $\bar{w}(e_i) = w(E_i)$ for $i = 1, \dots, q$, and $\bar{w}(e) = -w(e)$ if $e \in \bar{E} \cap E^-$. So, to solve MWCP, we need to solve $(2^q - 1)$ polynomial time instances of TSP. \square

3 Matching-partition inequalities

Given any edge set $F \subseteq E$, its *incidence vector* is the vector \mathbf{x}^F in $\{0, 1\}^E$ so that $x_e^F = 1$ if and only if $e \in F$. The convex hull of the incidence vectors of edge sets inducing connected subgraphs of G is called the *connected subgraph polytope* and is denoted $\text{CSP}(G)$, that is,

$$\text{CSP}(G) := \text{conv}\{\mathbf{x}^F \in \{0, 1\}^E : (V[F], F) \text{ is connected}\}.$$

The maximum-weight connected subgraph then consists of solving $\max\{\mathbf{w}^T \mathbf{x} : \mathbf{x} \in \text{CSP}(G)\}$, where $\mathbf{w} \in \mathbb{R}^E$ is a weight vector on the edges of G .

The concept of partition has played an important role towards designing facet-defining inequalities of many polyhedra. (See Baiou et al. [1] for a recent survey on

variants of partition inequalities.)

Let $M = \{e_1, e_2, \dots, e_p\}$ be a matching of G with $p \geq 1$ (i.e., a set of pairwise non-adjacent edges in E). For any partition $\pi = \{V_1, V_2, \dots, V_p\}$ of V so that $e_i \in E[V_i]$ for $i = 1, 2, \dots, p$, the pair (M, π) is called a *matching-partition* of G . Recall that the elements V_1, V_2, \dots, V_p are called the *classes* of the partition. With any matching-partition (M, π) of G , we associate the following *matching-partition inequality*

$$\mathbf{x}(E(\pi)) \geq \mathbf{x}(M) - 1. \tag{1}$$

Any matching-partition inequality (1) intuitively expresses the property that to connect n vertices together, one needs at least $n - 1$ edges. In [4], Didi Biha et al. proved the validity of the inequality (1) for the polytope $\text{CSP}(G)$ and gave necessary and sufficient conditions to be facet-defining. They also proved that when G is a cycle or a tree, $\text{CSP}(G)$ is given by the trivial inequalities (i.e., $0 \leq x(e) \leq 1$ for all $e \in E$) and the Matching-partition inequalities.

4 Separation of the matching-partition inequalities

The *Separation Problem for the Matching-Partition Inequalities* (1) (SPMPI for short) consists of deciding whether a given vector $\mathbf{x} \in \mathbb{R}^E$ satisfies the set of inequalities (1), and if not, followed by finding an inequality in (1) which is violated by \mathbf{x} . In [4], Didi Biha, Kerivin and Pe showed that SPMI is NP-complete even when the matching is fixed. They proved that it becomes polynomial if the cardinality of the matching is restricted to be 2. In this section, we prove that SPMI is polynomial on cycles and trees. We prove also that it polynomial under some restrictions on the matching.

4.1 Technical Lemmas

We give three technical lemmas related to the inequalities (1). Those lemmas will be useful for the forthcoming polynomial transformation.

Let $\mathcal{MP}(G)$ be the set composed of all the pairs (M, π) where M is a matching of G , π is a partition of V , and (M, π) is a matching-partition of G . Given a

vector \mathbf{x} in \mathbb{R}^E , consider the function $g_{\mathbf{x}} : \mathcal{MP}(G) \rightarrow \mathbb{R}$ which associates with each matching-partition $(M, \pi := \{V_1, V_2, \dots, V_p\})$, $p \geq 1$, the real value $g_{\mathbf{x}}(M, \pi) := \mathbf{x}(\delta(V_1, V_2, \dots, V_p)) - \mathbf{x}(M)$. The first lemma states that, in order to lower the value of $g_{\mathbf{x}}$ over $\mathcal{MP}(G)$, it is more beneficial to compose the matching of the largest-valued edges within the classes of the partition.

Lemma 1. *Let $(M := \{e_1, e_2, \dots, e_p\}, \pi := \{V_1, V_2, \dots, V_p\})$, $p \geq 1$, be a matching-partition of G and \mathbf{x} be a vector in \mathbb{R}^E . If there exists an edge f in $E[V_i]$ so that $x_f \geq x_{e_i}$, for some $i \in \{1, 2, \dots, p\}$, we then have $g_{\mathbf{x}}(M \setminus \{e_i\} \cup \{f\}, \pi) \leq g_{\mathbf{x}}(M, \pi)$; the inequality is strict if $x_f > x_{e_i}$.*

Proof. Substituting f for e_i in M straightforwardly preserves a matching-partition. Because of $x_f \geq x_{e_i}$, we then obtain $g_{\mathbf{x}}(M \setminus \{e_i\} \cup \{f\}, \pi) = g_{\mathbf{x}}(M, \pi) - x_{e_i} + x_f \leq g_{\mathbf{x}}(M, \pi)$. \square

Still keeping in mind a decrease of $g_{\mathbf{x}}$, the next lemma asserts that the total \mathbf{x} -value of the edges between two classes of a partition should not exceed the lowest x -value of the matching edges of those classes.

Lemma 2. *Let $(M := \{e_1, e_2, \dots, e_p\}, \pi := \{V_1, V_2, \dots, V_p\})$, $p \geq 2$, be a matching-partition of G and \mathbf{x} be a vector in \mathbb{R}^E . If there exist two distinct classes V_i and V_j of π so that $\mathbf{x}[V_i, V_j] \geq \min\{x_{e_i}, x_{e_j}\}$, we then have $g_{\mathbf{x}}(M', \pi') \leq g_{\mathbf{x}}(M, \pi)$, where $M' := M \setminus \{\operatorname{argmin}\{x_{e_i}, x_{e_j}\}\}$ and π' is the partition of V obtained from π by combining V_i and V_j into a single class; the inequality is strict if $\mathbf{x}[V_i, V_j] > \min\{x_{e_i}, x_{e_j}\}$.*

Proof. (M', π') clearly is a matching-partition of G . We have

$$\begin{aligned} g_{\mathbf{x}}(M', \pi') &= \mathbf{x}(\delta(V_1, V_2, \dots, V_p)) - \mathbf{x}[V_i, V_j] - \mathbf{x}(M) + \min\{x_{e_i}, x_{e_j}\} \\ &\leq g_{\mathbf{x}}(M, \pi), \end{aligned}$$

the inequality coming from $\mathbf{x}[V_i, V_j] \geq \min\{x_{e_i}, x_{e_j}\}$. \square

Lemma 3 can be viewed as a converse of Lemma 2. In fact, splitting a class which contains an edge set separating its matching edge from any edge whose x -value is larger than the one of the edge set would reduce the value of $g_{\mathbf{x}}$.

Lemma 3. Let $(M := \{e_1, e_2, \dots, e_p\}, \pi := \{V_1, V_2, \dots, V_p\})$, $p \geq 1$, be a matching-partition of G and \mathbf{x} be a vector in \mathbb{R}^E . If there exists a class V_j of π which can be partitioned into V_j^l and V_j^r so that $e_j \in E[V_j^l]$ and $\mathbf{x}[V_j^l, V_j^r] \leq \max\{x_e : e \in E[V_j^r]\}$, we then have $g_{\mathbf{x}}(M', \pi') \leq g_{\mathbf{x}}(M, \pi)$, where $M' := M \cup \{\operatorname{argmax}\{x_e : e \in E[V_j^r]\}\}$ and π' is the partition of V obtained from π by splitting V_j into V_j^l and V_j^r ; the inequality is strict if $\mathbf{x}[V_j^l, V_j^r] < \max\{x_e : e \in E[V_j^r]\}$.

Proof. Without loss of generality, assume $j = p$ and let $e_{p+1} := \operatorname{argmax}\{x_e : e \in E[V_p^r]\}$. Both matching $M' := \{e_i : i = 1, 2, \dots, p+1\}$ and partition $\pi' := \{V_1, V_2, \dots, V_{p-1}, V_p^l, V_p^r\}$ clearly form a matching-partition of G . We therefore have

$$\begin{aligned} g_{\mathbf{x}}(M', \pi') &= \mathbf{x}(\delta(V_1, V_2, \dots, V_p)) + \mathbf{x}[V_p^l, V_p^r] - \mathbf{x}(M) - x_{e_{p+1}} \\ &\leq g_{\mathbf{x}}(M, \pi), \end{aligned}$$

the inequality coming from $\mathbf{x}[V_j^l, V_j^r] \leq \max\{x_e : e \in E[V_j^r]\}$. \square

4.2 SPMPI on cycles and trees

On trees and cycles, if a class of a vertex partition induces a connected subgraph, then this subgraph is a tree. This straightforward fact actually happens to be very useful while studying the optimal solutions to the optimization problem

$$\min\{g_{\mathbf{x}}(M, \pi) : (M, \pi) \in \mathcal{MP}(G)\}, \quad (2)$$

where \mathbf{x} is a vector in $[0, 1]^E$. Notice that if a polynomial-time algorithm was known to solve (2), the separation problem for the matching-partition inequalities (1) would be polynomially solvable.

We first give two technical lemmas which will allow us to conveniently consider optimal solutions to (2) having partition whose classes all induce connected subgraphs.

Lemma 4. Let $(M := \{e_1, e_2, \dots, e_p\}, \pi := \{V_1, V_2, \dots, V_p\})$, $p \geq 2$, be an optimal solution to (2) so that the subgraph $G[V_j]$ contains an isolated vertex v , for some $j \in \{1, 2, \dots, p\}$. We then have

$$(i) \quad \mathbf{x}(\delta(v)) = 0 \text{ and}$$

(ii) the matching-partition (M, π') also is an optimal solution to (2), where π' is obtained from π by moving v to a class containing one of its neighbors in G .

Proof. Without loss of generality, suppose $j = 1$ and vertex v has a neighbor u in V_2 so that $x_{uv} = \max\{x_e : e \in \delta(v)\}$. Consider the partition $\pi' := \{V_1 \setminus \{v\}, V_2 \cup \{v\}, V_3, \dots, V_p\}$. It is obvious that (M, π') is a matching-partition of G . Since \mathbf{x} is non-negative, we have $g_{\mathbf{x}}(M, \pi') \leq g_{\mathbf{x}}(M, \pi) - x_{uv}$. The optimality of (M, π) then implies $x_{uv} = 0$. We therefore deduce $\mathbf{x}(\delta(v)) = \mathbf{x}[\{v\}, V_2] = 0$, and our proof is complete. \square

Lemma 5. Let $(M := \{e_1, e_2, \dots, e_p\}, \pi := \{V_1, V_2, \dots, V_p\})$, $p \geq 2$, be an optimal solution to (2) so that the subgraph $G[V_j]$ contains a connected component (U, F) with $F \neq \emptyset$ and $e_j \notin F$ for some $j \in \{1, 2, \dots, p\}$. We then have

(i) $\mathbf{x}(F) = 0$ and

(ii) the matching-partition (M', π') also is an optimal solution to (2), where $M' := M \cup \{e_{p+1}\}$ for any edge $e_{p+1} \in F$ and π' is obtained from π by replacing V_j by $V_j \setminus U$ and adding a new class $V_{p+1} := U$.

Proof. Let $e_{p+1} := \operatorname{argmax}\{x_e : e \in F\}$. Since (U, F) is a connected component of V_j not containing e_j , we have $V_j \setminus U \neq \emptyset$ and $\mathbf{x}[V_j \setminus U, U] = 0$. It is clear that (M', π') is a matching-partition of G , and then $g_{\mathbf{x}}(M', \pi') = g_{\mathbf{x}}(M, \pi) - x_{e_{p+1}}$. The optimality of (M, π) implies $x_{e_{p+1}} = 0$, which combines with the non-negativity of \mathbf{x} gives $\mathbf{x}(F) = 0$, and consequently (ii). \square

A direct consequence of Lemmas 4 and 5 is the following when non-connected graphs are considered in problem (2).

Corollary 1. Suppose graph G contains $k \geq 0$ connected components of size at least two $(U_1, F_1), (U_2, F_2), \dots, (U_k, F_k)$ and $q \geq 0$ isolated vertices $W := \{v_1, v_2, \dots, v_q\}$. If (M_i, π_i) represents an optimal solution to (2) with respect to (U_i, F_i) , $i = 1, 2, \dots, k$, an optimal solution to (2) with respect to G is $(\cup_{i=1}^k M_i, (\cup_{i=1}^{k-1} \pi_i) \cup \pi')$, where π' is the partition obtained from π_k by adding W to any of its classes.

In the following lemma, we focus on the connected classes of a matching-partition minimizing (2) whenever G is a tree or a cycle. We state that if V_j is such a class, any

path in $G[V_j]$ between e_j and some edge g in $\delta(V_j)$ is a sequence of non-increasing edges (with respect to their x -values), the one adjacent to g having a x -value at least x_g .

Lemma 6. *Suppose that G is a tree or a cycle. Let $(M := \{e_1, e_2, \dots, e_p\}, \pi := \{V_1, V_2, \dots, V_p\})$, $p \geq 2$, be an optimal solution to (2) so that the subgraph $G[V_j]$ is connected for some $j \in \{1, 2, \dots, p\}$. Consider any edge g in $\delta(V_j)$ and let $P(e_j, g) := (e_j, f_1, f_2, \dots, f_t, g)$, $t \geq 0$, be the unique path between e_j and g in the subgraph $G[E_j \cup \{g\}]$. We then have $x_{e_j} \geq x_{f_1} \geq x_{f_2} \geq \dots \geq x_{f_t} \geq x_g$.*

Proof. Assume without loss of generality that $j = 1$ and $g = u_1u_2$ with $u_1 \in V_1$ and $u_2 \in V_2$. Since G is a tree or a cycle, the subgraph $G[V_1]$ is a tree. We remark that by Lemma 1 we must have $x_{e_1} \geq x_e$ for all $e \in E_1$, and by Lemma 2 we also must have $x_{e_1} \geq x_g$. We then need to only consider the case $t > 0$. Suppose first there exist i_1 and i_2 in $\{1, 2, \dots, t\}$ so that $i_1 < i_2$ and $x_{f_{i_1}} < x_{f_{i_2}}$. Let $U \subseteq V_1$ be the vertex set of the connected component of $G[V_1] \setminus \{f_{i_1}\}$ containing f_{i_2} (and not e_1). We then can partition V_1 into $V_1 \setminus U$ and U where $e_1 \in E[V_1 \setminus U]$, $f_{i_2} \in E[U]$, and $f_{i_1} = [V_1 \setminus U, U]$. Since $x_{f_{i_1}} < x_{f_{i_2}} \leq \max\{x_e : e \in U\}$, Lemma 3 with respect to the foregoing partition of V_1 provides a contradiction to the optimality of (M, π) to (2). Consequently, we have $x_{e_j} \geq x_{f_1} \geq x_{f_2} \geq \dots \geq x_{f_t}$. Finally, suppose $x_{f_t} < x_g$ and consider the partition $\pi' := \{V_1 \setminus \{u_1\}, V_2 \cup \{u_1\}, V_3, \dots, V_p\}$. Pair (M, π') clearly is a matching-partition of G , and $g(M, \pi') = g(M, \pi) - x_g + x_{f_t} < g(M, \pi)$, a contradiction to the optimality of (M, π) to (2). \square

In Lemma 6 lies the essence of our linear-time algorithm which solves the separation problem for the matching-partition inequalities (1) on trees or cycles. This algorithm works by first seeking the (inclusionwise) maximal subtree which contains a largest x -valued edge and whose paths having this edge as an extremity satisfies the order specified in Lemma 6; then the algorithm proceeds in the similar way on the graph induced by the remaining edges, and so on.

Theorem 2. *Let $G = (V, E)$ be a tree or a cycle and let $\bar{x} \in [0, 1]^E$. The SPMPPI can be solved in linear time.*

Proof. we consider the problem

$$\min\{\bar{\mathbf{x}}(\delta(V_1, V_2, \dots, V_p)) - \bar{\mathbf{x}}(M) : (M, \delta(V_1, V_2, \dots, V_p)) \in \mathcal{MP}(G)\}, \quad (3)$$

that is, problem (2) with respect to $\bar{\mathbf{x}}$. If the optimal value to problem (3) is lower than -1 , any optimal solution to (3) induces a most-violated inequality (1) at $\bar{\mathbf{x}}$. Let $E^+(\bar{\mathbf{x}}) := \{e \in E : \bar{x}_e > 0\}$ and $G^+ := G[E^+(\bar{\mathbf{x}})]$. From Corollary 1, we may suppose that G^+ is connected, for otherwise one may decompose the separation problem for (1) to as many problems (3) as connected components of size at least 2 are in G^+ . Let e_1 denote a largest \bar{x} -valued edge in E , that is, $\bar{x}_{e_1} \geq \bar{x}_e$ for all $e \in E$. Notice first that $g_{\bar{\mathbf{x}}}(\{e_1\}, \{V\}) = x_{e_1} \geq 0$; in other words, one can only focus on matching-partition whose matching is of cardinality at least 2. Moreover, problem (3) clearly has an optimal solution, say $(M^* := \{e_1^*, e_2^*, \dots, e_p^*\}, \pi^* := \{V_1^*, V_2^*, \dots, V_p^*\})$. Without loss of generality, suppose $e_1^* = e_1$.

From Lemmas 4 and 5, (M^*, π^*) can be chosen so that each subgraph $G^+[V_i]$ is connected, $i = 1, 2, \dots, p$. We can also add the assumption that for any $i = 1, 2, \dots, p$ and any edge g in $\delta(V_i^*)$, the path $P(e_i^*, g)$ in $G^+[E_i^* \cup \{g\}]$ satisfies the order specified in Lemma 6. Consequently, for any edge $g = uv$ with $u \in V_1^*$ and $v \notin V_1^*$, we have $\bar{x}_g \leq \max\{\bar{x}_e : e \in \delta(v) \setminus \{g\}\}$. We now claim that (M^*, π^*) can be chosen so that for any edge $g = uv$ with $u \in V_1^*$ and $v \notin V_1^*$, we have $\bar{x}_g < \max\{\bar{x}_e : e \in \delta(v)\}$. Assume, to the contrary, $\bar{x}_g = \bar{x}_f$ where $f := \operatorname{argmax}\{\bar{x}_e : e \in \delta(v) \setminus \{g\}\}$. Without loss of generality, we suppose $v \in V_2^*$ and $f \in E^+[V_2^*]$, the latter due to Lemma 6. We consider two cases.

Case 1. If $\bar{x}_f = \bar{x}_{e_2^*}$, the matching-partition $(M \setminus \{e_2^*\}, \{V_1^* \cup V_2^*, V_3^*, \dots, V_p^*\})$ then is an optimal solution to problem (3) by Lemma 2.

Case 2. If $\bar{x}_f < \bar{x}_{e_2^*}$, because of Lemma 3, f can then be chosen so that it belongs to the only path $P(v, e_2^*)$ in $G^+[V_2^*]$ between v and e_2^* , that is, $P(v, e_2^*) := (f := f_0, f_1, \dots, f_t, e_2^*)$ with $t \geq 0$. Let τ be the largest integer in $\{0, 1, \dots, t\}$ so that $\bar{x}_{f_\tau} = \bar{x}_f$. Consider the partition $\{U, V_2\}$ of V_2^* so that U is the vertex set of the connected component of $G^+[E_2^+ \setminus \{f_\tau\}]$ containing v . We obviously have $e_2^* \in E^+[V_2]$. The matching-partition $(M, \{V_1^* \cup U, V_2, V_3^*, \dots, V_p^*\})$ clearly is also an optimal solution to problem (3).

We remark that the order specified in Lemma 6 is preserved in all the classes of the partitions considered in both cases. Consequently, the separation problem for the matching-partition (1) on trees and cycles can be first reduced to looking for V_1^* so that $e_1^* = e_1$, V_1^* satisfies Lemma 6, and for any edge g in $\delta(V_1^*)$ there exists an adjacent edge, not in $E[V_1^*]$ whose \bar{x} -value is larger than \bar{x}_g . An algorithm to find such a

Step 0 Set $V_1^* \leftarrow V[\{e_1\}]$; $F \leftarrow \delta(V_1^*)$;

Step 1 While $F \neq \emptyset$

set V_1^* is as follows.

Choose $f := uv$ in F with $v \notin V_1^*$;

Set $F \leftarrow F \setminus \{f\}$;

If $(\delta(v) = \{f\})$ or $(\bar{x}_f \geq \bar{x}_e$ for all $e \in \delta(v))$ then

Set $V_1^* \leftarrow V_1^* \cup \{v\}$; $F \leftarrow F \cup (\delta(v) \setminus \{f\})$;

Using a breadth-first-search type approach on G^+ from either endvertex of e_1 , Steps 0 and 1 can be implemented so that the only considered edges are those in $E^+[V_1^*] \cup \delta(V_1^*)$. Moreover, those edges are considered at most twice, one time in the conditional test, and a second time for those belonging to F .

The separation problem for (1) now is equivalent to solving problem (3) with respect to graph $G^+[V \setminus V_1^*]$. This graph may not be connected, yet using Corollary 1, the latter problem (3) can be solved by considering it with respect to each connected component of $G^+[V \setminus V_1^*]$. Notice that, because of the conditional test in Step 1, all the connected components of $G^+[V \setminus V_1^*]$ are of size at least 2. A subproblem (3) with respect to a connected component (U, F) can be solved in a similar way as described above, that is, by first looking for a maximal vertex set V_j^* containing a largest \bar{x} -valued edge in F , and then by solving problems (3) with respect to the connected components of $G^+[U \setminus V_j^*]$. Consequently, the separation problem for the matching-partition inequalities (1) can be solved in $O(n)$ time on trees and cycles. \square

4.3 SPMPI when the matching is maximal

Since the SPMPI is NP-complete, we have to settle a heuristic approach for the SPMPI if we want to use the matching-partition inequalities within a cutting-plane framework. The following propositions can be used to devise such heuristic.

Proposition 1. *If M is required to be a perfect matching, then the SPMPI can be solved in polynomial time.*

Proof. Let $\mathbf{x} \in [0, 1]^E$ and $M := \{e_1 = u_1w_1, \dots, e_p = u_pw_p\}$, $p \geq 3$, be a perfect matching. Let $V_i := \{u_i, w_i\}$ for $i = 1, \dots, p$. The partition $\pi_M := \{V_1, \dots, V_p\}$ is the unique partition of V so that (M, π_M) is a matching-partition. We have $\mathbf{x}(E(\pi_M)) - \mathbf{x}(M) = \mathbf{x}(E) - 2\mathbf{x}(M)$. Thus, finding a perfect matching M so that $\mathbf{x}(E(\pi_M)) - \mathbf{x}(M)$ is minimum, is equivalent to finding a perfect matching M so that $\mathbf{x}(M)$ is maximum. This latter problem is well-known to be polynomial [7, 8]. \square

Proposition 2. *When M is a fixed maximal matching, then the SPMPI can be solved in polynomial time.*

Proof.

Let $\mathbf{x} \in [0, 1]^E$ and $M := \{e_1 = u_1w_1, \dots, e_p = u_pw_p\}$, $p \geq 3$, be a fixed maximal matching. We can find a partition $\pi_M := \{V_1, \dots, V_p\}$ so that $e_i \in E[V_i]$ for $i = 1, \dots, p$, and $\mathbf{x}(E(\pi_M)) - \mathbf{x}(M)$ is minimum. In fact, since M is maximal, the vertices not covered by M form a stable set S_M . We can then construct $\pi_M := \{V_1, \dots, V_p\}$ as follows: set $V_i := \{u_i, w_i\}$ for $i = 1, \dots, p$; for every $v \in S_M$ choose $i \in \{1, \dots, p\}$ so that $\mathbf{x}(\{\{v\}, \{u_i, w_i\}\}) = c_v^M$, where $c_v^M = \max\{\mathbf{x}(\{\{v\}, \{u_i, w_i\}\}) : i = 1, \dots, p\}$, and add v into V_i .

Let $\pi := \{V'_1, \dots, V'_p\}$ be a partition of V such that $e_i \in E[V_i]$ for $i = 1, \dots, p$, and (M, π) is a matching-partition of G . We will prove that $\mathbf{x}(E(\pi_M)) \leq \mathbf{x}(E(\pi))$. For every $v \in S_M$, let $i(v) \in \{1, \dots, p\}$ such that $v \in V'_{i(v)}$. Let $F = E[V[M]] \setminus M = \{e = u_iw_j \in E : i, j \in \{1, \dots, p\}, i \neq j\}$. We have

$$\begin{aligned} \mathbf{x}(E(\pi)) &= \mathbf{x}(F) + \sum_{v \in S_M} (\mathbf{x}(\delta(v)) - \mathbf{x}(\{\{v\}, \{u_{i(v)}, w_{i(v)}\}\})) \\ &= \mathbf{x}(F) + \sum_{v \in S_M} \mathbf{x}(\delta(v)) - \sum_{v \in S_M} \mathbf{x}(\{\{v\}, \{u_{i(v)}, w_{i(v)}\}\}) \\ &\geq \mathbf{x}(F) + \sum_{v \in S_M} \mathbf{x}(\delta(v)) - \sum_{v \in S_M} c_v^M \\ &= \mathbf{x}(E(\pi_M)), \end{aligned}$$

implying that $\mathbf{x}(E(\pi_M)) - \mathbf{x}(M) \leq \mathbf{x}(E(\pi)) - \mathbf{x}(M)$. \square

Now, we will show that for all possible pairs of maximum \mathbf{x} -valued matching and vertex partition in a graph G , the SPMPPI can be solved in polynomial time. Let $\mathbf{x} \in \mathbb{R}_+^E$ and $\mathcal{M}_{\mathbf{x}} := \{M \subseteq E : M \text{ is a matching, } \mathbf{x}(M) \text{ is maximum}\}$. Consider then the following optimization problem:

$$\min\{\mathbf{x}(E(\pi)) - \mathbf{x}(M) : (M, \pi) \in \mathcal{MP}(G) \text{ with } M \in \mathcal{M}_{\mathbf{x}}\}. \quad (4)$$

Proposition 3. *The optimization problem (4) is polynomial.*

Proof. We shall use the same notation as in the proof of the previous proposition. We can suppose, without loss of generality, that $\mathbf{x} > \mathbf{0}$. Let $\alpha := \max\{\mathbf{x}(M) : M \text{ is a matching}\}$ and $M := \{e_1 = u_1w_1, \dots, e_p = u_pw_p\}$, $p \geq 3$, be a fixed matching with $\mathbf{x}(M) = \alpha$. Since $\mathbf{x}(M)$ is maximum and $\mathbf{x} > \mathbf{0}$, M is a maximal matching. As we have seen in Proposition 2, $\mathbf{x}(E(\pi_M))$ is minimum, implying that $\mathbf{x}(E(\pi_M)) - \mathbf{x}(M) = \mathbf{x}(E(\pi_M)) - \alpha$ is minimum. We have

$$\mathbf{x}(E(\pi_M)) = \mathbf{x}(E) - \mathbf{x}(M) - \sum_{v \in S_M} c_v^M,$$

and hence

$$\mathbf{x}(E(\pi_M)) - \mathbf{x}(M) = \mathbf{x}(E) - 2\mathbf{x}(M) - \sum_{v \in S_M} c_v^M = \mathbf{x}(E) - 2\alpha - \sum_{v \in S_M} c_v^M.$$

Thus, the optimisation problem (4) is equivalent to

$$\max\left\{\sum_{v \in S_M} c_v^M : M \text{ is a matching with } \mathbf{x}(M) = \alpha\right\}.$$

For every $v \in V$ define $c_v = \max\{\mathbf{x}([\{v\}, V(\{e\})]) : e \notin \delta(v)\}$. By convention, $\mathbf{x}([\{v\}, V(\{e\})]) = 0$ if $[\{v\}, V(\{e\})] = \emptyset$. Thus, if M is a matching so that $\mathbf{x}(M) = \alpha$ and $v \in S_M$ then $c_v^M = c_v$. Let B be a sufficiently large positive number. The above discussion and the characterization of matching polytope [7] lead to the following linear program:

$$(LP) \left\{ \begin{array}{ll} \max B \sum_{e \in E} x_e x_e + \sum_{v \in V} c_v y_v & \\ \text{s.t.} & \\ \mathbf{x}(\delta(v)) + y_v = 1 & \text{for all } v \in V, \\ \mathbf{x}(E(S)) \leq \frac{|S|-1}{2} & \text{for all } S \subset V \text{ with } |S| \text{ odd,} \\ (\mathbf{x}, \mathbf{y}) \geq \mathbf{0}. & \end{array} \right.$$

It is clear that if we add integrality constraints (i.e., $(\mathbf{x}, \mathbf{y}) \in \mathbb{N}^E \times \mathbb{N}^V$) to (LP) , then \mathbf{x} corresponds to a matching and \mathbf{y} corresponds to a stable set so that no edge in the matching is adjacent to a vertex in the stable set. The optimal solution of the resulting integer-programming problem will correspond to a matching M so that $\mathbf{x}(M) = \alpha$ and $\mathbf{x}((E(\pi_M)) - \mathbf{x}(M))$ is minimum.

Let P be the polyhedron defined by the feasible solution set of (LP) . We now claim that P is integral. In fact, let (x, y) be an extreme point of P . If x is integer, then obviously y is integer. Suppose that \mathbf{x} is not integer. Since $\mathbf{x} \in P_{\text{matching}}(G)$, where $P_{\text{matching}}(G)$ is the matching polytope, and $P_{\text{matching}}(G)$ is integral [7], then there exists an extreme point $\bar{\mathbf{x}} \in P_{\text{matching}}(G)$ so that every constraint of $P_{\text{matching}}(G)$ binding at \mathbf{x} is also binding at $\bar{\mathbf{x}}$. Note that if $y_v = 0$, then $\mathbf{x}(\delta(v)) = 1$ and hence $\bar{\mathbf{x}}(\delta(v)) = 1$. Let $\bar{\mathbf{y}} \in \mathbb{R}^V$ defined as $\bar{y}_v = 1 - \bar{\mathbf{x}}(\delta(v))$ for all $v \in V$. It is clear that every constraint of P binding at (\mathbf{x}, \mathbf{y}) is also binding at $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$. Since $\bar{\mathbf{x}}$ is integer [7], we have $(\bar{\mathbf{x}}, \bar{\mathbf{y}}) \neq (\mathbf{x}, \mathbf{y})$. This is a contradiction with the fact that (\mathbf{x}, \mathbf{y}) is an extreme point of P . Thus, problem (4) is equivalent to the linear program (LP) . And, so P is integral.

Since the separation problem associated with P is polynomially solvable [20], so is the linear program (LP) , and hence the optimization problem (4) can be solved in polynomial time. \square

5 Conclusions

In this article, we have considered the maximum-weight connected subgraph problem (MWCSP). We have showed that this problem is polynomial when the matching number is less or equal some fixed integer. We have studied the separation problem associated with the so-called Matching-partition inequalities introduced in [4]. In particular, we have showed that it is polynomial on cycles and trees. It would be interesting first to identify more classes of graphs on which the separation problem for the matching-partition inequalities can be solved in polynomial time, and second to devise approximate separation routines. The latter may be handled through additional properties on the matchings (e.g., perfect, maximal, maximum) which we have already considered in this article. It would be interesting to use the results of

this article in the framework of a branch-and-cut algorithm for the NWCSPP. MWCSPP is closely related to the Steiner tree problem, and any algorithm for the first problem can be used to resolve the second one. An obvious question is whether our results can be used to improve the resolution of the Steiner tree problem.

References

- [1] M. Baïou, F. Barahona, and A.R. Mahjoub, Partition inequalities: Separation, Extensions, and Network design, to appear in *Progress in Combinatorial Optimization*, A. R. Mahjoub (ED.), ISTE-WILEY(2011).
- [2] J. Chakareski, P. Frossard, H.L.M. Kerivin, J. Leblet, and G. Simon, A note on the data-driven capacity of P2P networks, Technical Report EPFL-LTS-2009-008, École Polytechnique Fédérale de Lausanne, Switzerland (2009).
- [3] E. Dahlhaus, D.S. Johnson, C.H. Papadimitriou, P.D. Seymour, and M. Yannakakis, The complexity of multiterminal cuts, *SIAM J. Comput.* 23 (1994), 864-894.
- [4] M. Didi Biha, H. Kerivin, and P. Ng, Polyhedral study of the connected subgraph problem, *Discrete Mathematics* 338(1) (2015), pp. 80–92.
- [5] R. Diestel, *Graph Theory*, Third ed., Springer-Verlag, Berlin Heidelberg (2006).
- [6] S.E. Dreyfus, R.A. Wagner, The Steiner Problems in Graphs, *Networks* 1(1971), 195-207.
- [7] J. Edmonds, Maximum matching and polyhedron with 0,1-vertices, *Journal of research of the National Bureau of Standards* 69 (1965), 125-130.
- [8] J. Edmonds, paths, Trees and Flowers, *Canadian Journal of Mathematics* 17 (1965), 449-467.
- [9] J. Feigenbaum, C. Papadimitriou, S. Shenker, Sharing the cost of multicast transmissions, *Journal of Computer and System Sciences* 63 (2001), 21-41.
- [10] M. Goemans, The Steiner tree polytope and related polyhedra, *Mathematical Programming* 63 (1994), 157-182.
- [11] M. Goemans, D. Williamson, A general approximation technique for constrained forest problems, *SIAM J. Comput.* 24 (1995), 296-317.

- [12] A.V. Goldberg and R.E. Tarjan, A new approach to the maximum flow problem, *J. Assoc. Comput. Mach.* 35 (1988), 921–940.
- [13] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin (1988).
- [14] F.K. Hwang, D.S. Richards, and P. Winter, The Steiner tree problem, *Annals of Discrete Mathematics* 53, North-Holland, Amsterdam (1992).
- [15] T. Ideker, O. Ozier, B. Schwikowski, and A.F. Siegel, Discovering regulatory and signalling circuits in molecular interaction networks, *Bioinformatics* 18, Suppl. 1 (2002), S233-240.
- [16] D.S. Johnson, M. Minkoff, S. Phillips, The prize collecting Steiner tree problem: theory and practice, *Proc. 11th Ann. ACM-SIAM Symp. on Discrete Algorithms* (2000) 760-769.
- [17] R.M. Karp, Maximum-weight connected subgraph, [HTTP://WWW.CYTOSCAPE.ORG/ISMB2002](http://www.cytoscape.org/ISMB2002) (2002).
- [18] H.L.M. Kerivin and P.H. Ng, Maximum-weight connected-subgraph problems, *Technical Report* (2011).
- [19] K. Menger, Zur allgemeinen Kurventheorie, *Fund. Math.* 10 (1927) 96-115.
- [20] M.W. Padberg and M.R. Rao, Odd minimum cut-sets and b-matchings, *Math. Oper. Res.* 7 (1982), 67-80.
- [21] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Springer-Verlag, Berlin Heidelberg (2003).