



Relationship of two formulations for shortest bibranchings

Kazuo Murota¹ · Kenjiro Takazawa²

Received: 24 January 2020 / Revised: 29 April 2020 / Accepted: 9 July 2020 / Published online: 22 July 2020
© The JJIAM Publishing Committee and Springer Japan KK, part of Springer Nature 2020

Abstract

The shortest bibranching problem is a common generalization of the minimum-weight edge cover problem in bipartite graphs and the minimum-weight arborescence problem in directed graphs. For the shortest bibranching problem, an efficient primal-dual algorithm is given by Keijsper and Pendavingh (J Comb Theory Ser B 73:130–145, 1998), and the tractability of the problem is ascribed to total dual integrality in a linear programming formulation by Schrijver (Ann Discret Math 16:261–280, 1982). Another view on the tractability of this problem is afforded by a valuated matroid intersection formulation by Takazawa (Jpn J Ind Appl Math 29:561–573, 2012). In the present paper, we discuss the relationship between these two formulations for the shortest bibranching problem. We first demonstrate that the valuated matroid intersection formulation can be derived from the linear programming formulation through the Benders decomposition, where integrality is preserved in the decomposition process and the resulting convex programming is endowed with discrete convexity. We then show how a pair of primal and dual optimal solutions of one formulation is constructed from that of the other formulation, thereby providing a connection between polyhedral combinatorics and discrete convex analysis.

Keywords Polyhedral combinatorics · Discrete convex analysis · Benders decomposition · Integrality · Duality

Mathematics Subject Classification 52A41 · 90C10 · 90C25 · 90C27 · 90C35

This work is supported by CREST, JST, Grant numbers JPMJCR14D2, JPMJCR1402, Japan, and JSPS KAKENHI Grant numbers JP26280004, JP16K16012, JP20K11697, JP20K11699.

✉ Kenjiro Takazawa
takazawa@hosei.ac.jp

Kazuo Murota
murota@tmu.ac.jp

¹ Department of Economics and Business Administration, Tokyo Metropolitan University, Tokyo 192-0397, Japan

² Department of Industrial and Systems Engineering, Faculty of Science and Engineering, Hosei University, Tokyo 184-8584, Japan

1 Introduction

The shortest bibranching problem, introduced in [12] (see also [14]), is a common generalization of the minimum-weight edge cover problem in bipartite graphs and the minimum-weight arborescence problem in directed graphs. In a directed graph $D = (V, A)$ with vertex set V and arc set A , an arc subset $B \subseteq A$ is called a *branching* if B does not contain a directed cycle and every vertex v has at most one arc in B entering v . For a vertex $r \in V$, a branching B is called an *r -arborescence* if every vertex $v \in V \setminus \{r\}$ has an arc in B entering v . In an undirected graph $G = (V, E)$ with vertex set V and edge set E , an edge subset $F \subseteq E$ is an *edge cover* if the union of the end vertices of the edges in F is equal to V .

The *shortest bibranching problem* is described as follows. Let $D = (V, A)$ be a directed graph $D = (V, A)$, and $\{S, T\}$ be a (nontrivial) partition of the vertex set V , that is, S and T are nonempty disjoint subsets of V such that $S \cup T = V$. A subset $B \subseteq A$ of arcs is called an *S - T bibranching* if, in the subgraph (V, B) , every vertex in S reaches T and every vertex in T is reachable from S . We denote the set of nonnegative integers by \mathbb{Z}_+ .

Instance: A directed graph (V, A) , a partition $\{S, T\}$ of V , and a nonnegative integer arc-weight $w \in \mathbb{Z}_+^A$.

Objective: Find an S - T bibranching B minimizing $w(B) = \sum_{a \in B} w(a)$.

We denote an arc leaving u and entering v by uv . We also denote $A[S] = \{uv \in A : u, v \in S\}$, $A[T] = \{uv \in A : u, v \in T\}$, and $A[S, T] = \{uv \in A : u \in S, v \in T\}$. Throughout this paper, we assume, without loss of generality, that there is no arc uv with $u \in T$ and $v \in S$, which implies that $A = A[S] \cup A[T] \cup A[S, T]$.

The shortest S - T bibranching problem includes, as special cases, the minimum-weight edge cover problem in bipartite graphs and the minimum-weight r -arborescence problem in directed graphs. If $A[S] = A[T] = \emptyset$, then $D = (V, A)$ is a bipartite graph with color classes S and T , and an S - T bibranching corresponds exactly to an edge cover in this bipartite graph (the underlying undirected bipartite graph, to be more precise). If $S = \{r\}$, an inclusion-wise minimal S - T bibranching is exactly an r -arborescence, and hence the minimum-weight r -arborescence problem is reduced to the shortest S - T bibranching problem.

There are several methods to solve the shortest bibranching problem in polynomial time. First, the total dual integrality of a linear programming formulation is proved by Schrijver [12], and hence the ellipsoid method works. Second, based on this formulation, Keijsper and Pendavingh [6] designed a much faster primal-dual algorithm, which is followed by a scaling algorithm by Babenko [1]. Third, the shortest bibranching problem can be described as the shortest strong connector problem in a source-sink connected digraph, which can be reduced to the weighted matroid intersection problem (see [14] for details). The most recent method by Takazawa [16] is a polynomial reduction of the shortest bibranching problem to the valuated matroid intersection problem [7, 8], and hence

any valuated matroid intersection algorithm can solve the shortest bibranching problem. Among the above four approaches, the linear-programming and valuated-matroid approaches are extended to a further generalization, the shortest b -bibranching problem [18].

These results demonstrate that the shortest bibranching problem can be understood through the standard framework of polyhedral combinatorics [14], and a relatively new framework of discrete convex analysis [10] as well. In the present paper, we discuss the relationship between these two approaches to the shortest bibranching problem. First, we demonstrate that the valuated matroid intersection formulation can be derived from the linear programming formulation through the Benders decomposition [2, 3], where integrality is preserved in the decomposition process and the resulting convex programming is endowed with discrete convexity. In this view the valuated matroid intersection formulation corresponds to the master problem and the subproblems¹ are instances of the minimum-weight r -arborescence problem. This understanding naturally leads us to a solution algorithm analogous to the Bender decomposition. The concave functions representing the objective values of the subproblems are replaced by valuated matroids, which are discrete analogues of concave functions. Next we discuss the relationship between the two duality theorems associated with the linear programming and valuated matroid intersection formulations, and show how a pair of primal and dual optimal solutions of one formulation is constructed from that of the other formulation.

The organization of this paper is as follows. In Sect. 2, we recapitulate the two formulations for the shortest S - T bibranching problem, a linear programming formulation and a valuated matroid intersection formulation, where the emphasis is laid on a clear-cut presentation of the existing derivation of the latter formulation. In Sect. 3, we point out that the valuated matroid intersection formulation can also be derived from the linear programming formulation through the Benders decomposition, which turns out to be compatible with integrality and discrete convexity. In Sect. 4, we exhibit how to construct a pair of primal and dual optimal solutions for the valuated matroid intersection formulation from a pair of primal and dual optimal solutions for the linear programming formulation. Section 5 shows the converse, i.e., how to construct a pair of primal and dual optimal solutions for the linear programming formulation from a pair of primal and dual optimal solutions for the valuated matroid intersection formulation.

2 Existing two formulations

2.1 Linear programming formulation

In this section, we review the system of linear inequalities describing the shortest S - T bibranching problem [12, 14]. This system of inequalities is a common generalization of that for the minimum-weight edge cover problem in bipartite graphs and that for

¹ These subproblems correspond to *recourse problems* in stochastic programming.

the minimum-weight r -arborescence problem. The total dual integrality of this system forms the basis of our understanding of the shortest S – T bibranching problem in the framework of polyhedral combinatorics [14].

Let $D = (V, A)$ be a directed graph, $\{S, T\}$ be a (nontrivial) partition of V , and $w \in \mathbb{Z}_+^A$ be a nonnegative integer arc-weight vector. For $X \subseteq V$, let $\delta^+X = \{uv \in A : u \in X, v \in V \setminus X\}$ and $\delta^-X = \{uv \in A : u \in V \setminus X, v \in X\}$. The following linear program (P) in variable $x \in \mathbb{R}^A$ represents the shortest S – T bibranching problem:

$$(P) \quad \text{Minimize} \quad \sum_{a \in A} w(a)x(a)$$

$$\text{subject to} \quad \sum_{a \in \delta^+S'} x(a) \geq 1 \quad (\emptyset \neq S' \subseteq S), \tag{1}$$

$$\sum_{a \in \delta^-T'} x(a) \geq 1 \quad (\emptyset \neq T' \subseteq T), \tag{2}$$

$$x(a) \geq 0 \quad (a \in A). \tag{3}$$

Described below is the dual program (D) of (P), whose variables are $y \in \mathbb{R}^{2^S \setminus \{\emptyset\}}$ and $z \in \mathbb{R}^{2^T \setminus \{\emptyset\}}$:

$$(D) \quad \text{Maximize} \quad \sum_{\emptyset \neq S' \subseteq S} y(S') + \sum_{\emptyset \neq T' \subseteq T} z(T')$$

$$\text{subject to} \quad \sum_{S' \subseteq S, a \in \delta^+S'} y(S') + \sum_{T' \subseteq T, a \in \delta^-T'} z(T') \leq w(a) \quad (a \in A), \tag{4}$$

$$y(S') \geq 0 \quad (\emptyset \neq S' \subseteq S), \tag{5}$$

$$z(T') \geq 0 \quad (\emptyset \neq T' \subseteq T). \tag{6}$$

A vertex subset X satisfying either $X \subseteq S$ or $X \subseteq T$ is called a *bicut*. Intuitively, the dual program (D) represents the problem of packing bicuts $S' \subseteq S$ and $T' \subseteq T$ so that the number of bicuts “crossing” an arc a does not exceed $w(a)$ for each $a \in A$.

The complementary slackness conditions for (P) and (D) are as follows:

$$x(a) > 0 \implies \sum_{S' : a \in \delta^+S'} y(S') + \sum_{T' : a \in \delta^-T'} z(T') = w(a), \tag{7}$$

$$y(S') > 0 \implies \sum_{a \in \delta^+S'} x(a) = 1, \tag{8}$$

$$z(T') > 0 \implies \sum_{a \in \delta^-T'} x(a) = 1, \tag{9}$$

where $a \in A$ in (7), $\emptyset \neq S' \subseteq S$ in (8), and $\emptyset \neq T' \subseteq T$ in (9).

Theorem 1 (Schrijver [12], see also [14]) *For an arbitrary integer vector $w \in \mathbb{Z}_+^A$, (P) and (D) have integral optimal solutions.*

2.2 M-convex submodular flow formulation

Another formulation of the shortest S - T bibrancheding problem, given in [16], falls in the framework of valuated matroid intersection [7, 8]. This formulation provides a new insight into the shortest S - T bibrancheding problem through discrete convex analysis [10]. In this paper we adopt a formulation by the M^{\natural} -convex submodular flow problem [9], which does not differ essentially from the valuated matroid intersection formulation [16], but offers a clearer correspondence to the linear programming formulation in Sect. 2.1.

We begin with some definitions. For a finite set X and an integer vector $\eta \in \mathbb{Z}^X$, we define $\text{supp}^+(\eta) = \{u \in X : \eta(u) > 0\}$ and $\text{supp}^-(\eta) = \{u \in X : \eta(u) < 0\}$. For $Y \subseteq X$, $\chi_Y \in \mathbb{Z}^X$ is the characteristic vector of Y defined by $\chi_Y(u) = 1$ if $u \in Y$ and $\chi_Y(u) = 0$ if $u \in X \setminus Y$. For $u \in X$, $\chi_{\{u\}}$ is abbreviated as χ_u . For a function $f : \mathbb{Z}^X \rightarrow \overline{\mathbb{Z}}$, where $\overline{\mathbb{Z}} = \mathbb{Z} \cup \{+\infty\}$, the effective domain $\text{dom} f$ of f is defined by $\text{dom} f = \{\eta \in \mathbb{Z}^X : f(\eta) < +\infty\}$. A function $f : \mathbb{Z}^X \rightarrow \overline{\mathbb{Z}}$ is called an M^{\natural} -convex function [10, 11] if it satisfies the following exchange property:

For each $\eta, \zeta \in \mathbb{Z}^X$ and $u \in \text{supp}^+(\eta - \zeta)$, it holds that

$$f(\eta - \chi_u) + f(\zeta + \chi_u) \leq f(\eta) + f(\zeta) \tag{10}$$

or there exists $v \in \text{supp}^-(\eta - \zeta)$ such that

$$f(\eta - \chi_u + \chi_v) + f(\zeta + \chi_u - \chi_v) \leq f(\eta) + f(\zeta). \tag{11}$$

A set $D \subseteq \mathbb{Z}^X$ is called an M^{\natural} -convex set if its indicator function $\delta_D : \mathbb{Z}^X \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$\delta_D(\eta) = \begin{cases} 0 & (\eta \in D), \\ +\infty & (\eta \notin D) \end{cases}$$

is an M^{\natural} -convex function. Equivalently, a set $D \subseteq \mathbb{Z}^X$ is an M^{\natural} -convex set if and only if it satisfies the following exchange property:

For each $\eta, \zeta \in D$ and $u \in \text{supp}^+(\eta - \zeta)$, it holds that

$$\eta - \chi_u \in D \quad \text{and} \quad \zeta + \chi_u \in D \tag{12}$$

or there exists $v \in \text{supp}^-(\eta - \zeta)$ such that

$$\eta - \chi_u + \chi_v \in D \quad \text{and} \quad \zeta + \chi_u - \chi_v \in D. \tag{13}$$

It is pointed out in Takazawa [15, 17] that discrete convexity inherent in branchedings follows from the arguments in Schrijver [13]. A further connection of S - T bibranchedings to discrete convex analysis is revealed in [16]. In the following, we summarize the

arguments in [15–17] and exhibit an M^h -convex submodular flow formulation to highlight the discrete convexity in the shortest S – T bibranching problem.

For the M^h -convex submodular flow formulation, it is convenient to regard a (shortest) S – T bibranching as a discrete system consisting of three components, a branching, a cobranching, and a bipartite edge cover, where a *cobranching* means an arc subset such that the reversal of its arcs is a branching. For a precise formulation, we need some notations.

For a digraph $D = (V, A)$ and a partition $\{S, T\}$ of V , denote the subgraphs induced by S and T , respectively, as $D[S]$ and $D[T]$, that is, $D[S] = (S, A[S])$ and $D[T] = (T, A[T])$. For $B \subseteq A$, denote $B[S] = \{uv \in B : u, v \in S\}$, $B[T] = \{uv \in B : u, v \in T\}$, and $B[S, T] = \{uv \in B : u \in S, v \in T\}$. For an arc set $F \subseteq A[S, T]$, define $\partial^+ F \in \mathbb{Z}^S$ and $\partial^- F \in \mathbb{Z}^T$ by

$$\begin{aligned} \partial^+ F(u) &= |F \cap \delta^+ u| \quad (u \in S), \\ \partial^- F(v) &= |F \cap \delta^- v| \quad (v \in T), \end{aligned}$$

respectively, where $\delta^+ u = \{uv \in A : v \in V \setminus \{u\}\}$ and $\delta^- v = \{uv \in A : u \in V \setminus \{v\}\}$. For a branching B_T in $D[T]$, let $R(B_T)$ denote the set of vertices in T which no arc in B_T enters. For a cobranching B_S in $D[S]$, let $R^*(B_S)$ denote the set of vertices in S which no arc in B_S leaves.

Then we can say that an arc subset $B \subseteq A$ is an S – T bibranching if $B[S]$ is a cobranching with $R^*(B[S]) = \text{supp}^+(\partial^+ B[S, T])$ and $B[T]$ is a branching with $R(B[T]) = \text{supp}^+(\partial^- B[S, T])$. Equivalently, $B \subseteq A$ is an S – T bibranching if $B[S]$ is a cobranching in $D[S]$, $B[T]$ is a branching in $D[T]$, and $B[S, T]$ is an edge cover in the graph $D[R^*(B[S]), R(B[T])]$. This definition slightly differs from that in [12]: here $B[S]$ should be a cobranching and $B[T]$ should be a branching, which is not necessarily the case in the definition in [12]. However, we may naturally adopt this alternative definition as long as we consider the shortest S – T bibranching problem.

If we first specify $F \subseteq A[S, T]$ as the intersection of $A[S, T]$ and our S – T bibranching, then arcs in $A[T]$ to be added to F should form a branching B_T in $D[T]$ such that $R(B_T) = \text{supp}^+(\partial^- F)$. Similarly, a cobranching $B_S \subseteq A[S]$ satisfying $R^*(B_S) = \text{supp}^+(\partial^+ F)$ should be added to F . Then an S – T bibranching B is obtained as $B = F \cup B_S \cup B_T$.

The minimum weights of B_T and B_S are expressed respectively by the functions $g_T : \mathbb{Z}^T \rightarrow \overline{\mathbb{Z}}$ and $g_S : \mathbb{Z}^S \rightarrow \overline{\mathbb{Z}}$ defined as follows. The effective domain $\text{dom } g_T$ is defined as

$$\text{dom } g_T = \{\eta \in \mathbb{Z}_+^T : \text{there is a branching } B_T \text{ in } D[T] \text{ with } R(B_T) = \text{supp}^+(\eta)\},$$

and, for $\eta \in \text{dom } g_T$, the function value $g_T(\eta)$ is defined as

$$g_T(\eta) = \min\{w(B_T) : B_T \text{ is a branching in } D[T], R(B_T) = \text{supp}^+(\eta)\}. \quad (14)$$

Similarly, we define $g_S : \mathbb{Z}^S \rightarrow \overline{\mathbb{Z}}$ by

$$\begin{aligned} \text{dom } g_S = & \{ \eta \in \mathbb{Z}_+^S : \text{there is a cobranching } B_S \text{ in } D[S] \text{ with } R^*(B_S) = \text{supp}^+(\eta), \\ & g_S(\eta) = \min\{w(B_S) : B_S \text{ is a cobranching in } D[S], R^*(B_S) = \text{supp}^+(\eta)\} \\ & (\eta \in \text{dom } g_S). \end{aligned} \tag{15}$$

We remark here that g_T and g_S are not linear functions. For example, $g_T(\eta_1 + \eta_2) \neq g_T(\eta_1) + g_T(\eta_2)$ in general.

For $\xi \in \{0, 1\}^{A[S,T]}$, define $\partial\xi \in \mathbb{Z}^{S \cup T}$ by

$$\partial\xi(v) = |\{a : \xi(a) = 1, a \in \delta^+v\}| - |\{a : \xi(a) = 1, a \in \delta^-v\}| \quad (v \in S \cup T).$$

The restrictions of $\partial\xi \in \mathbb{Z}^{S \cup T}$ to S and T are denoted by $\partial\xi|_S \in \mathbb{Z}^S$ and $\partial\xi|_T \in \mathbb{Z}^T$, respectively. Now the shortest S - T bibranching problem is formulated as the following 0-1 integer program in variable $\xi \in \{0, 1\}^{A[S,T]}$ representing $F \subseteq A[S, T]$:

$$\text{(MSF) Minimize } w(\xi) + g_S(\partial\xi|_S) + g_T(-\partial\xi|_T), \tag{16}$$

where $w(\xi) = \sum_{a \in A[S,T]} w(a)\xi(a)$.

Notice that the functions $g_S(\partial\xi|_S)$ and $g_T(\partial\xi|_T)$ are not linear functions. However, this formulation reveals the discrete convexity inherent in the shortest S - T bibranching problem, which is described as follows.

Theorem 2 (Takazawa [16]) *Functions g_S in (15) and g_T in (14) are M^{\natural} -convex functions. Thus, the shortest S - T bibranching problem is formulated as the M^{\natural} -convex submodular flow problem (MSF) in (16).*

We often refer to $\xi \in \{0, 1\}^{A[S,T]}$ as a *flow*, and a flow ξ is said to be *feasible* if $\partial\xi|_S \in \text{dom } g_S$ and $-\partial\xi|_T \in \text{dom } g_T$. That is, $\xi \in \{0, 1\}^{A[S,T]}$ is feasible if there exist a cobranching B_S in $D[S]$ with $R^*(B_S) = \text{supp}^+(\partial\xi|_S)$ and a branching B_T in $D[T]$ with $R(B_T) = \text{supp}^+(-\partial\xi|_T)$.

Remark 1 Note that $\partial\xi$ may not be a $\{0, 1\}$ -vector, though ξ itself is a $\{0, 1\}$ -vector. Hence the domains of g_S and g_T should not be restricted to sets of $\{0, 1\}$ -vectors, but they are sets of integers. With some further argument Takazawa [16] reduced the formulation (MSF) to the *valuated matroid intersection problem* [7, 8] so that both of the original shortest S - T bibranching problem and the resulting valuated matroid intersection problem can be defined on $\{0, 1\}$ -vectors. In this paper, however, we adopt the M^{\natural} -convex submodular flow formulation (MSF) in order to make the whole logic clearer. □

We now show the proof of Theorem 2 by clarifying the arguments scattered in [15–17]. The matroidal nature of branchings (M^{\natural} -convexity of $\text{dom } g_T$, to be specific) is first noted in [15]. For a digraph $D = (V, A)$, a *source component* K in D is a strong component such that no arc in A enters K , where we identify a component K and its vertex set and denote either of them by K . It is not difficult to see that, for $U \subseteq V$, there exists a branching B with $R(B) = U$ if and only if $U \cap K \neq \emptyset$ for every source component K , where $R(B)$ denotes the set of vertices without

entering arcs in B . Hence, $\{V \setminus R(B) : B \text{ is a branching in } D\}$ is an independent set of a partition matroid, and thus $\{\eta \in \mathbb{Z}^V : B \text{ is a branching in } D, R(B) = \text{supp}^+(\eta)\}$ is an M^\natural -convex set (g-matroid).

To prove Theorem 2, we need a stronger exchange property of branchings: the arc sets of branchings also have an exchange property. First, the following lemma is derived from Edmonds' disjoint branchings theorem [5].

Lemma 1 [13] *Let $D = (V, A)$ be a digraph, and B_1, B_2 be branchings partitioning A . For $R'_1, R'_2 \subseteq V$ satisfying $R'_1 \cup R'_2 = R(B_1) \cup R(B_2)$ and $R'_1 \cap R'_2 = R(B_1) \cap R(B_2)$, the arc set A can be partitioned into branchings B'_1 and B'_2 such that $R(B'_1) = R'_1$ and $R(B'_2) = R'_2$ if and only if $K \cap R'_1 \neq \emptyset$ and $K \cap R'_2 \neq \emptyset$ for every source component K .*

The next lemma, which follows from Lemma 1, describes the exchange property of the arc sets of branchings.

Lemma 2 ([13], See also [16, 17]) *Let $D = (V, A)$ be a digraph, B_1 and B_2 be branchings partitioning A , and $s \in R(B_1) \setminus R(B_2)$. Then, there exist branchings B'_1 and B'_2 which partition A and satisfy that*

- $R(B'_1) = R(B_1) \setminus \{s\}$ and $R(B'_2) = R(B_2) \cup \{s\}$, or
- there exists $t \in R(B_2) \setminus R(B_1)$ such that $R(B'_1) = (R(B_1) \setminus \{s\}) \cup \{t\}$ and $R(B'_2) = (R(B_2) \cup \{s\}) \setminus \{t\}$.

Proof Let K be the strong component containing s . If K is a source component, then let t be the root of the directed tree in B_2 containing s , and define $R'_1 = (R(B_1) \cup \{s\}) \setminus \{t\}$ and $R'_2 = (R(B_2) \setminus \{s\}) \cup \{t\}$. Note that $t \in K$ and $t \in R(B_2) \setminus R(B_1)$. Otherwise, define $R'_1 = R(B_1) \cup \{s\}$ and $R'_2 = R(B_2) \setminus \{s\}$. Then the claim follows from Lemma 1.

We are now ready to show a proof for Theorem 2.

Proof for Theorem 2 It suffices to deal with g_T , since the M^\natural -convexity of g_S is proved similarly. Let $\eta, \zeta \in \text{dom } g_T$, and let $u \in \text{supp}^+(\eta - \zeta)$.

If $\zeta(u) \geq 1$, then $\text{supp}^+(\eta - \chi_u) = \text{supp}^+(\eta)$ and $\text{supp}^+(\zeta + \chi_u) = \text{supp}^+(\zeta)$, which imply $g_T(\eta - \chi_u) = g_T(\eta)$ and $g_T(\zeta + \chi_u) = g_T(\zeta)$. Hence $g_T(\eta - \chi_u) + g_T(\zeta + \chi_u) \leq g_T(\eta) + g_T(\zeta)$ in (10) holds with equality.

If $\eta(u) \geq 2$ and $\zeta(u) = 0$, then $\text{supp}^+(\eta - \chi_u) = \text{supp}^+(\eta)$ and $\text{supp}^+(\zeta + \chi_u) = \text{supp}^+(\zeta) \cup \{u\}$, which imply $g_T(\eta - \chi_u) = g_T(\eta)$ and $g_T(\zeta + \chi_u) \leq g_T(\zeta)$. The latter is derived as follows. Let B_ζ be a branching in $D[T]$ yielding $g_T(\zeta)$, i.e., $R(B_\zeta) = \text{supp}^+(\zeta)$ and $w(B_\zeta) = g_T(\zeta)$. Now $\zeta(u) = 0$ implies $u \in T \setminus R(B)$, i.e., B_ζ has an arc a entering u . Then, $B'_\zeta = B_\zeta \setminus \{a\}$ is a branching with $R(B'_\zeta) = \text{supp}^+(\zeta + \chi_u)$, and thus $g_T(\zeta + \chi_u) \leq w(B'_\zeta) = w(B_\zeta) - w(a) \leq w(B_\zeta) = g_T(\zeta)$, where the latter inequality follows from the nonnegativity of w . Therefore $g_T(\eta - \chi_u) + g_T(\zeta + \chi_u) \leq g_T(\eta) + g_T(\zeta)$ in (10) holds.

If $\eta(u) = 1$ and $\zeta(u) = 0$, then there exist branchings B_η and B_ζ in $D[T]$ such that

$$\begin{aligned} R(B_\eta) &= \text{supp}^+(\eta), & w(B_\eta) &= g_T(\eta), \\ R(B_\zeta) &= \text{supp}^+(\zeta), & w(B_\zeta) &= g_T(\zeta). \end{aligned}$$

It is understood that in digraph $(T, B_\eta \cup B_\zeta)$, an arc a contained in both B_η and B_ζ has multiplicity two in $B_\eta \cup B_\zeta$. We have $u \in R(B_\eta) \setminus R(B_\zeta)$. By Lemma 2 applied to $(T, B_\eta \cup B_\zeta)$, there exist branchings B'_η and B'_ζ which partition $B_\eta \cup B_\zeta$ and satisfy that

$$R(B'_\eta) = R(B_\eta) \setminus \{u\} \quad \text{and} \quad R(B'_\zeta) = R(B_\zeta) \cup \{u\}$$

or

$$R(B'_\eta) = (R(B_\eta) \setminus \{u\}) \cup \{v\} \quad \text{and} \quad R(B'_\zeta) = (R(B_\zeta) \cup \{u\}) \setminus \{v\}$$

for some $v \in R(B_\zeta) \setminus R(B_\eta)$. Then, in the former case we obtain

$$g_T(\eta - \chi_u) + g_T(\zeta + \chi_u) \leq w(B'_\eta) + w(B'_\zeta) = w(B_\eta) + w(B_\zeta) = g_T(\eta) + g_T(\zeta),$$

which shows (10), and in the latter case,

$$\begin{aligned} g_T(\eta - \chi_u + \chi_v) + g_T(\zeta + \chi_u - \chi_v) &\leq w(B'_\eta) + w(B'_\zeta) \\ &= w(B_\eta) + w(B_\zeta) = g_T(\eta) + g_T(\zeta), \end{aligned}$$

which shows (11). This proves M^\natural -convexity of g_T .

3 M^\natural -convex submodular flow formulation via Benders decomposition

In this section, we demonstrate that the M^\natural -convex submodular flow formulation (MSF) can be obtained from the linear program (P) through the Benders decomposition, where integrality is preserved in the decomposition process and the resulting convex programming is endowed with discrete convexity.

We denote by $x_{S,T}$, x_S , and x_T the restrictions $x|_{A[S,T]}$, $x|_{A[S]}$, and $x|_{A[T]}$ of x to $A[S, T]$, $A[S]$, and $A[T]$, respectively. Similarly, we use abbreviations $w_{S,T} = w|_{A[S,T]}$, $w_S = w|_{A[S]}$, and $w_T = w|_{A[T]}$. Then the linear program (P) is rewritten as

$$\begin{aligned} \text{(LP)} \quad & \text{Minimize} \quad \sum_{a \in A[S,T]} w_{S,T}(a)x_{S,T}(a) + \sum_{a \in A[S]} w_S(a)x_S(a) + \sum_{a \in A[T]} w_T(a)x_T(a) \\ & \text{subject to} \quad \sum_{a \in A[S,T]} x_{S,T}(a) \geq 1, \end{aligned} \tag{17}$$

$$\sum_{a \in \delta^+ S' \cap A[S, T]} x_{S, T}(a) + \sum_{a \in \delta^+ S' \cap A[S]} x_S(a) \geq 1 \quad (\emptyset \neq S' \subsetneq S), \tag{18}$$

$$\sum_{a \in \delta^- T' \cap A[S, T]} x_{S, T}(a) + \sum_{a \in \delta^- T' \cap A[T]} x_T(a) \geq 1 \quad (\emptyset \neq T' \subsetneq T), \tag{19}$$

$$x_{S, T}, x_S, x_T \geq 0. \tag{20}$$

The Benders decomposition proceeds in the following manner. The *master problem*, in variable $x_{S, T}$, is described as

$$\begin{aligned} \text{(MASTER)} \quad & \text{Minimize} \quad \sum_{a \in A[S, T]} w_{S, T}(a)x_{S, T}(a) + h_S(x_{S, T}) + h_T(x_{S, T}) \\ & \text{subject to} \quad \sum_{a \in A[S, T]} x_{S, T}(a) \geq 1, \end{aligned} \tag{21}$$

$$x_{S, T} \geq 0, \tag{22}$$

where the functions h_S and h_T respectively represent the optimal values of the following subproblems (SUB(S)) and (SUB(T)) parametrized by $x_{S, T}$:

$$\begin{aligned} \text{(SUB}(S)) \quad & \text{Minimize} \quad \sum_{a \in A[S]} w_S(a)x_S(a) \\ & \text{subject to} \quad \sum_{a \in \delta^+ S' \cap A[S]} x_S(a) \geq 1 - \sum_{a \in \delta^+ S' \cap A[S, T]} x_{S, T}(a) \quad (\emptyset \neq S' \subsetneq S), \end{aligned} \tag{23}$$

$$x_S \geq 0; \tag{24}$$

$$\begin{aligned} \text{(SUB}(T)) \quad & \text{Minimize} \quad \sum_{a \in A[T]} w_T(a)x_T(a) \\ & \text{subject to} \quad \sum_{a \in \delta^- T' \cap A[T]} x_T(a) \geq 1 - \sum_{a \in \delta^- T' \cap A[S, T]} x_{S, T}(a) \quad (\emptyset \neq T' \subsetneq T), \end{aligned} \tag{25}$$

$$x_T \geq 0. \tag{26}$$

The subproblems (SUB(S)) and (SUB(T)) are linear programs, whereas the master problem (MASTER) is a convex program.

We are concerned with a $\{0, 1\}$ -valued optimal solution $x \in \{0, 1\}^A$. Theorem 1 guarantees the existence of an integer optimal solution for (LP), and then the constraints (17)–(20) imply that it is $\{0, 1\}$ -valued. This implies that the master problem (MASTER) and the subproblems (SUB(S)) and (SUB(T)) are also equipped with discreteness.

The combinatorial (or matroidal) nature of the subproblems can be seen as follows. Fix $x_{S,T} = \xi \in \{0, 1\}^{A[S,T]}$ satisfying (21) and (22). We first consider (SUB(T)). On noting that (25) can be rewritten as

$$\sum_{a \in \delta^{-} T' \cap A[T]} x_T(a) \geq 1 + \partial \xi|_T \quad (\emptyset \neq T' \subsetneq T)$$

and x_T may be assumed to be a $\{0, 1\}$ -vector, we can see that (SUB(T)) is nothing other than the problem of finding the minimum-weight branching $B_T \subseteq A[T]$ in $D[T]$ with $R(B_T) = \text{supp}^+(-\partial \xi|_T)$. Thus, the optimal value of (SUB(T)), denoted $h_T(\xi)$, is in fact equal to $g_T(-\partial \xi|_T)$ for the function g_T defined in (14), i.e., $h_T(\xi) = g_T(-\partial \xi|_T)$. In addition, the function g_T is M^{\natural} -convex by Theorem 2. This shows the matroidal property of (SUB(T)). Similarly, we have $h_S(\xi) = g_S(\partial \xi|_S)$ for the other subproblem (SUB(S)), where g_S is also an M^{\natural} -convex function by Theorem 2.

With the above observations the master problem (MASTER) can be rewritten as:

$$\begin{aligned} &\text{Minimize} && \sum_{a \in A[S,T]} w_{S,T}(a)\xi(a) + g_S(\partial \xi|_S) + g_T(-\partial \xi|_T) \\ &\text{subject to} && \xi \in \{0, 1\}^{A[S,T]}, \end{aligned}$$

where the constraint (21) in (MASTER) is deleted since it is implied by $\partial \xi|_S \in \text{dom } g_S$ and $-\partial \xi|_T \in \text{dom } g_T$. Thus, the master problem (MASTER) in the Benders decomposition is equivalent to the M^{\natural} -convex submodular formulation (MSF) in (16).

We remark that this observation implies that the linear program (P) can be solved by the Benders decomposition, in which the subproblems are the minimum-weight r -arborescence problem and hence can be solved efficiently. This method requires neither the ellipsoid method to solve (P) directly, nor the sophisticated techniques used in the previous combinatorial algorithms [1, 6] other than finding a minimum-weight branching.

It is emphasized that the formulation in the M^{\natural} -convex submodular problem (MSF) in Sect. 2.2 is based on purely combinatorial arguments, without directly relying on the linear programming formulation (P) or (LP). In contrast, in this section we have started with the linear programming formulation (P) and its integrality (Theorem 1), and derived (MSF) therefrom.

4 Optimal flow and potential from optimal LP solutions

According to the theory of M -convex submodular flows in discrete convex analysis [9, 10], the M^{\natural} -convex submodular flow formulation (MSF) admits an optimality criterion in terms of potentials (dual variables). The objective of this section is to show that an optimal flow and an optimal potential for (MSF) can be constructed from the optimal solutions of the primal-dual pair of linear programs (P) and (D).

4.1 Optimality criterion for (MSF)

The optimality criterion for M^{\natural} -convex submodular flows [9, 10], when tailored to (MSF), is given in Theorem 3 below. For vectors $p \in \mathbb{Z}^S$ and $q \in \mathbb{Z}^T$, define functions $g_S[+p] : \mathbb{Z}^S \rightarrow \overline{\mathbb{Z}}$ and $g_T[+q] : \mathbb{Z}^T \rightarrow \overline{\mathbb{Z}}$ by

$$g_S[+p](\eta) = g_S(\eta) + \sum_{u \in S} p(u)\eta(u) \quad (\eta \in \mathbb{Z}^S),$$

$$g_T[+q](\zeta) = g_T(\zeta) + \sum_{v \in T} q(v)\zeta(v) \quad (\zeta \in \mathbb{Z}^T),$$

where g_S and g_T are given in (15) and (14), respectively.

Theorem 3 *A feasible flow $\xi \in \{0, 1\}^{A[S,T]}$ is an optimal solution for (MSF) if and only if there exist $p \in \mathbb{Z}^S$ and $q \in \mathbb{Z}^T$ satisfying the following (i)–(iii):*

(i) for $a = uv \in A[S, T]$,

$$\xi(a) = 1 \implies w(a) + p(u) - q(v) \leq 0, \tag{27}$$

$$\xi(a) = 0 \implies w(a) + p(u) - q(v) \geq 0. \tag{28}$$

(ii) $\partial\xi|_S \in \arg \min (g_S[-p])$.

(iii) $-\partial\xi|_T \in \arg \min (g_T[+q])$.

We refer to $(p, q) \in \mathbb{Z}^{S \cup T}$ satisfying (i)–(iii) in Theorem 3 for some $\xi \in \{0, 1\}^{A[S,T]}$ as an *optimal potential* for (MSF). Intuitively Theorem 3 is understood in the following way. Given $p \in \mathbb{Z}^S$ and $q \in \mathbb{Z}^T$, define a reduced weight $w'(a)$ for each $a = uv \in A$ by

$$w'(a) = \begin{cases} w(a) + p(u) - q(v) & (a = uv \in A[S, T]), \\ w(a) + p(u) & (a = uv \in A[S]), \\ w(a) - q(v) & (a = uv \in A[T]). \end{cases}$$

Roughly speaking, Conditions (i)–(iii) mean that $\xi \in \{0, 1\}^{A[S,T]}$ corresponds to optimal solutions for the three independent problems of minimizing

$$\sum_{a \in A[S,T]} w'_{S,T}(a)x_{S,T}(a), \quad g'_S(\partial\xi|_S), \quad g'_T(-\partial\xi|_T),$$

where g'_S and g'_T are defined by (15) and (14), respectively, with w replaced by w' . To be precise, we need a certain care of the cases $\partial\xi(u) \geq 2$ for $u \in S$ and $\partial\xi(v) \leq -2$ for $v \in T$, because in those cases $g'_S(\partial\xi|_S)$ and $g'_T(-\partial\xi|_T)$ do not exactly coincide with the minimum-weights of a cobranchings B_S with $R^*(S) = \text{supp}^+(\partial\xi|_S)$ and a branchings B_T with $R(T) = \text{supp}^+(-\partial\xi|_T)$, respectively.

4.2 Construction of an optimal flow and potential for (MSF)

We will show how to construct an optimal flow $\xi^* \in \{0, 1\}^{A[S, T]}$ and an optimal potential $(p^*, q^*) \in \mathbb{Z}^{S \cup T}$ for (MSF) from the optimal solutions $x \in \{0, 1\}^A$ and $(y, z) \in \mathbb{Z}^{2^S \setminus \{\emptyset\}} \times \mathbb{Z}^{2^T \setminus \{\emptyset\}}$ of the linear programs (P) and (D). Recall from Theorem 1 that both (P) and (D) have integer optimal solutions.

Given x and (y, z) , define ξ^* and (p^*, q^*) by

$$\xi^*(a) = x(a) \quad (a \in A[S, T]), \tag{29}$$

$$p^*(u) = - \sum_{S' \subseteq S, u \in S'} y(S') \quad (u \in S), \tag{30}$$

$$q^*(v) = \sum_{T' \subseteq T, v \in T'} z(T') \quad (v \in T). \tag{31}$$

In other words, for $a \in A[S, T]$, the flow value $\xi^*(a)$ is one if and only if a is an arc in the shortest bibrancheding represented by $x \in \{0, 1\}^A$. For $u \in S$, the optimal potential value $p^*(u)$ is equal to the negation of the number of bicuts $S' \subseteq S$ containing u . Similarly, for $v \in T$, the optimal potential value $q^*(v)$ is equal to the number of bicuts $T' \subseteq T$ containing v .

Below we show that ξ^* and (p^*, q^*) satisfy Condition (i)–(iii) in Theorem 3, to prove that they are an optimal flow and an optimal potential for (MSF). The above intuitive understanding of Theorem 3 also works here, but the proof is more involved due to the possibilities of $\partial \xi^*(u) \geq 2$ for $u \in S$ and $\partial \xi^*(v) \leq -2$ for $v \in T$.

Theorem 4 *Let $x \in \{0, 1\}^A$ and $(y, z) \in \mathbb{Z}^{2^S} \times \mathbb{Z}^{2^T}$ be optimal solutions for (P) and (D), respectively. Then, ξ^* and (p^*, q^*) defined in (29)–(31) are an optimal flow and an optimal potential for (MSF), respectively.*

Proof In the following we show (i)–(iii) in Theorem 3. We first show (i). For $a = uv \in A[S, T]$, it holds that

$$\begin{aligned} -p^*(u) + q^*(v) &= \sum_{S' \subseteq S, u \in S'} y(S') + \sum_{T' \subseteq T, v \in T'} z(T') \\ &= \sum_{S' \subseteq S, a \in \delta^+ S'} y(S') + \sum_{T' \subseteq T, a \in \delta^- T'} z(T') \\ &\leq w(a), \end{aligned}$$

where the last inequality is due to (4). Moreover, if $\xi(a) = 1$, the inequality turns into an equality by (7), and therefore (27) and (28) follow.

Next we show (iii) (rather than (ii)). Let $w'(a) = w(a) - q^*(v)$ for $a = uv \in A[T]$. For an arbitrary $\eta \in \text{dom } g_T$, it holds that

$$\begin{aligned}
 &g_T[+q^*](\eta) \\
 &= \min\{w(B) : B \text{ is a branching in } D[T], R(B) = \text{supp}^+(\eta)\} + \sum_{v \in T} q^*(v)\eta(v) \\
 &= \min\{w'(B) : B \text{ is a branching in } D[T], R(B) = \text{supp}^+(\eta)\} \\
 &\quad + q^*(T) + \sum_{v \in \text{supp}^+(\eta)} q^*(v)(\eta(v) - 1).
 \end{aligned} \tag{32}$$

A lower bound for the right-hand side of (32) is provided as follows. For the first term we have

$$\begin{aligned}
 &\min\{w'(B) : B \text{ is a branching in } D[T], R(B) = \text{supp}^+(\eta)\} \\
 &\geq - \sum_{\emptyset \neq T' \subseteq T} (|T'| - 1)z(T'),
 \end{aligned} \tag{33}$$

since, for any branching B in $D[T]$ with $R(B) = \text{supp}^+(\eta)$, it holds that

$$\begin{aligned}
 w'(B) &= \sum_{uv \in B} (w(uv) - q^*(v)) \\
 &= \sum_{uv \in B} \left(w(uv) - \sum_{T' \subseteq T, v \in T'} z(T') \right) \\
 &\geq \sum_{uv \in B} \left(\sum_{T' \subseteq T, uv \in \delta^- T'} z(T') - \sum_{T' \subseteq T, v \in T'} z(T') \right)
 \end{aligned} \tag{34}$$

$$\begin{aligned}
 &= \sum_{uv \in B} \left(- \sum_{T' \subseteq T, u, v \in T'} z(T') \right) \\
 &= - \sum_{\emptyset \neq T' \subseteq T} |B[T']| \cdot z(T') \\
 &\geq - \sum_{\emptyset \neq T' \subseteq T} (|T'| - 1)z(T'),
 \end{aligned} \tag{35}$$

where the first inequality is by (4). In addition, the last term of the right-hand side of (32) is nonnegative, i.e.,

$$\sum_{v \in \text{supp}^+(\eta)} q^*(v)(\eta(v) - 1) \geq 0, \tag{36}$$

since $q^*(v) \geq 0$ by (31). From (32), (33), and (36), we obtain

$$g_T[+q^*](\eta) \geq - \sum_{\emptyset \neq T' \subseteq T} (|T'| - 1)z(T') + q^*(T),$$

where the right-hand side is a constant for a fixed z . Hence, in order to prove $-\partial\xi^*|_T \in \arg \min g_T[+q^*]$, it suffices to show that the three inequalities (34), (35), and (36) in the above turn into equalities when $\eta = -\partial\xi^*|_T$.

For the first and second inequalities (34) and (35), let $B^* = \text{supp}^+(x)$ be the shortest S - T bibranching corresponding to x . Then $B^*[T]$ is a branching in $D[T]$ such that $R(B^*[T]) = \text{supp}^+(-\partial\xi^*|_T)$, and the first inequality (34) holds with equality for $B^*[T]$ by (7). Moreover, $|B^* \cap \delta^-T'| = 1$ for every nonempty $T' \subseteq T$ with $z(T') > 0$ by (9). Thus, $|B^*[T']| = |T'| - |B^* \cap \delta^-T'| = |T'| - 1$ if $z(T') > 0$, and hence the equality in (35) follows. For the third inequality (36), suppose $q^*(v) > 0$ and let $T' \subseteq T$ contribute to $q^*(v)$ in (31), i.e., $v \in T'$ and $z(T') > 0$. Since $v \in \text{supp}^+(-\partial\xi^*|_T)$, there exists at least one arc $a^* = uv \in A[S, T]$ such that $x(a^*) = 1$. Then we have that $a^* \in \delta^-T'$. We also have $\sum_{a \in \delta^-T'} x(a) = 1$ by (9), and hence such a^* is unique. Therefore $-\partial\xi^*|_T(v) = 1$ follows. Hence all terms in the summation in (36) are equal to zero.

Finally, condition (ii) is proved similarly to (iii).

5 Optimal LP solutions from an optimal flow and potential

In this section, we describe how to construct optimal solutions for (P) and (D) of the linear programming formulation from an optimal flow $\xi \in \{0, 1\}^{A[S, T]}$ and an optimal potential $(p, q) \in \mathbb{Z}^{S \cup T}$ for the M^{\natural} -convex submodular flow formulation (MSF).

5.1 Optimal flow and potential with a stronger property

We first establish the following lemma, in which (p, q) need not be an optimal potential but an arbitrary pair of vectors.

Lemma 3 *For arbitrary $p \in \mathbb{Z}^S$ and $q \in \mathbb{Z}^T$, the following hold.*

- If $\arg \min g_S[-p] \neq \emptyset$, then $p(u) \leq 0$ for every $u \in S$. Moreover, $p(u) = 0$ if $\eta^*(u) \geq 2$ for some $\eta^* \in \arg \min g_S[-p]$.
- If $\arg \min g_T[+q] \neq \emptyset$, then $q(v) \geq 0$ for every $v \in T$. Moreover, $q(v) = 0$ if $\eta^*(v) \geq 2$ for some $\eta^* \in \arg \min g_T[+q]$.

Proof It suffices to prove the latter assertion. Suppose that $q(v) < 0$ for some $v \in T$. Note that $\chi_T \in \text{dom } g_T$. Then, for an arbitrary positive integer α , we have that

$$g_T[+q](\chi_T + \alpha\chi_v) = g_T(\chi_T) + q(T) + \alpha q(v),$$

which tends to $-\infty$ as $\alpha \rightarrow +\infty$. Therefore, $\arg \min g_T[+q] \neq \emptyset$ implies $q(v) \geq 0$ for every $v \in T$.

Now suppose that $\eta^* \in \arg \min g_T[+q]$ and $\eta^*(v) \geq 2$. Then $\text{supp}^+(\eta^*) = \text{supp}^+(\eta^* - \chi_v)$, and hence $g_T(\eta^*) = g_T(\eta^* - \chi_v)$, whereas $g_T[+q](\eta^*) \leq g_T[+q](\eta^* - \chi_v)$ by $\eta^* \in \arg \min g_T[+q]$. Therefore, we have

$$\begin{aligned} g_T[+q](\eta^*) &\leq g_T[+q](\eta^* - \chi_v) \\ &= g_T(\eta^* - \chi_v) + q \cdot (\eta^* - \chi_v) \\ &= g_T[+q](\eta^*) - q(v), \end{aligned}$$

which implies $q(v) \leq 0$. Therefore, $q(v) = 0$ follows.

We next show the existence of an optimal potential satisfying a property stronger than (27).

Lemma 4 *For an optimal flow $\xi \in \{0, 1\}^{A[S, T]}$, there exists an optimal potential $(p, q) \in \mathbb{Z}^{S \cup T}$ such that*

$$\xi(a) = 1 \implies w(a) + p(u) - q(v) = 0 \tag{37}$$

holds for every $a = uv \in A[S, T]$.

Proof Let (p°, q°) be a given optimal potential and assume that (37) fails for $a^* = u^*v^* \in A[S, T]$. This means, by (27), that $\xi(a^*) = 1$ and $w(a^*) + p^\circ(u^*) - q^\circ(v^*) < 0$.

By Lemma 3, it holds that $p^\circ(u^*) \leq 0$ and $q^\circ(v^*) \geq 0$. Then, there exist $\alpha, \beta \in \mathbb{Z}$ such that

$$p^\circ(u^*) \leq \alpha \leq 0, \quad 0 \leq \beta \leq q^\circ(v^*), \quad w(a^*) + \alpha - \beta = 0.$$

With such α, β we modify (p°, q°) to $(p', q') \in \mathbb{Z}^S \times \mathbb{Z}^T$ as

$$p'(u) = \begin{cases} p^\circ(u) & (u \in S \setminus \{u^*\}), \\ \alpha & (u = u^*), \end{cases} \quad q'(v) = \begin{cases} q^\circ(v) & (v \in T \setminus \{v^*\}), \\ \beta & (v = v^*). \end{cases}$$

Note that (37) holds for $a^* = u^*v^*$ with respect to the modified potential (p', q') .

Claim (p', q') is an optimal potential.

Proof for Claim We prove that ξ and (p', q') satisfy (i)–(iii) in Theorem 3. Note that (i)–(iii) in Theorem 3 hold for ξ and (p°, q°) .

We first show (i). Inequality (28) follows from $p' \geq p^\circ$ and $q' \leq q^\circ$. As for (27), it is obvious that (27) holds for a^* . Let $\hat{a} = \hat{u}\hat{v} \in A[S, T] \setminus \{a^*\}$ be such that $\xi(\hat{a}) = 1$. If \hat{a} is not adjacent to a^* , then $w(\hat{a}) + p'(\hat{u}) - q'(\hat{v}) = w(\hat{a}) + p^\circ(\hat{u}) - q^\circ(\hat{v}) \leq 0$. Suppose that \hat{a} is adjacent to a^* , i.e., $\hat{v} = v^*$ or $\hat{u} = u^*$. If $\hat{v} = v^*$, then $-\partial\xi(\hat{v}) \geq 2$, and $q^\circ(\hat{v}) = 0$ follows from Lemma 3. Therefore, $q'(\hat{v}) = q^\circ(\hat{v}) = 0$, and hence (27) holds for \hat{a} . The other case of $\hat{u} = u^*$ can be treated similarly.

We next show (iii), while noting that (ii) can be proved similarly as (iii). Suppose, to the contrary, that $-\partial\xi|_T \notin \arg \min(g_T[+q'])$. That is, $g_T[+q'](\eta) < g_T[+q'](-\partial\xi|_T)$ holds for some $\eta \in \mathbb{Z}_+^T$. Here, we claim the following:

$$-\partial\xi(v^*) = 1, \tag{38}$$

$$\eta(v^*) \geq 2. \tag{39}$$

Proof for (38). Since $-\partial\xi|_T \in \arg \min(g_T[+q^\circ])$ and $-\partial\xi|_T \notin \arg \min(g_T[+q'])$, we have that $q' \neq q^\circ$ and consequently $0 \leq q'(v^*) < q^\circ(v^*)$. Then, (38) follows from Lemma 3.

Proof for (39). Denote $\Delta = q^\circ(v^*) - q'(v^*) > 0$. Since

$$\begin{aligned} 0 &< g_T[+q'](-\partial\xi|_T) - g_T[+q'](\eta) \\ &= \left(g_T[+q^\circ](-\partial\xi|_T) - g_T[+q^\circ](\eta) \right) + \Delta \cdot \left(\eta(v^*) + \partial\xi(v^*) \right) \\ &\leq \Delta \cdot \left(\eta(v^*) + \partial\xi(v^*) \right), \end{aligned}$$

we have that $\eta(v^*) \geq -\partial\xi(v^*) + 1 = 2$.

For $\hat{\eta} \in \mathbb{Z}_+^T$ defined by

$$\hat{\eta}(v) = \begin{cases} 1 & (v = v^*), \\ \eta(v) & (v \in T \setminus \{v^*\}), \end{cases}$$

it holds that

$$\begin{aligned} g_T[+q^\circ](\hat{\eta}) &= g_T[+q'](\eta) - q'(v^*)(\eta(v^*) - 1) + \Delta \\ &\leq g_T[+q'](\eta) + \Delta \\ &< g_T[+q'](-\partial\xi|_T) + \Delta \\ &= g_T[+q^\circ](-\partial\xi|_T), \end{aligned}$$

where (38) and (39) are used. This contradicts $-\partial\xi|_T \in \arg \min(g_T[+q^\circ])$. Thus we have shown $-\partial\xi|_T \in \arg \min(g_T[+q'])$ in (iii). This completes the proof of the claim.

By the above claim, we can reduce the number of arcs violating (37) by modifying $(p, q) = (p^\circ, q^\circ)$ to $(p, q) = (p', q')$, while maintaining the optimality. By repeating such modifications we eventually arrive at the situation where (37) holds for every $a = uv \in A[S, T]$. This completes the proof for Lemma 4.

5.2 Construction of optimal primal and dual solutions for the LP formulation

In what follows, we assume that ξ is an optimal flow and (p, q) is an optimal potential satisfying the condition (37) in Lemma 4. We construct optimal solutions for (P) and

(D) by combining the primal and dual solutions of the minimum-weight arborescence problems in auxiliary directed graph defined by $D[S]$ and p , and that by $D[T]$ and q .

Let $D_T = (V_T, A_T)$ be a directed graph with arc weight $w' \in \mathbb{Z}^{A_T}$ defined as follows:

$$V_T = \{r_T\} \cup T, \quad A_T = \{r_T v : v \in T\} \cup A[T],$$

$$w'(uv) = \begin{cases} q(v) & (u = r_T), \\ w(uv) & (u \in T), \end{cases}$$

where r_T is a newly introduced additional vertex. For any r_T -arborescence \tilde{B}_T in D_T , $B_T = \tilde{B}_T \cap A[T] = \tilde{B}_T[T]$ is a branching in $D[T]$ with $R(B_T) = \{v \in T : r_T v \in \tilde{B}_T\}$. Conversely, for any branching B_T in $D[T]$, $\tilde{B}_T = B_T \cup \{r_T v : v \in R(B_T)\}$ is an r_T -arborescence in D_T .

Lemma 5 *There exists in D_T a minimum-weight r_T -arborescence \tilde{B}_T such that $R(\tilde{B}_T[T]) = \text{supp}^+(-\partial\xi|_T)$.*

Proof By the correspondence between r_T -arborescences in D_T and branchings in $D[T]$ described above, the minimum-weight r_T -arborescence problem in D_T with respect to w' is equivalent to minimizing $w(B_T) + \sum_{v \in R(B_T)} q(v)$ over branchings B_T in $D[T]$. On the other hand, in minimizing $g_T[+q](\eta)$, we may assume $\eta \in \{0, 1\}^T$ by Lemma 3, and for $\eta = \chi_X$ with $X \subseteq T$, the value of $g_T[+q](\chi_X)$ is equal to the minimum of $w(B_T) + \sum_{v \in X} q(v)$ for a branching B_T in $D[T]$ satisfying $R(B_T) = X$. Since $-\partial\xi|_T \in \arg \min g_T[+q]$, there exists a minimum-weight branching B_T in $D[T]$ satisfying $R(B_T) = \text{supp}^+(-\partial\xi|_T)$. Then the corresponding r_T -arborescence $\tilde{B}_T = B_T \cup \{r_T v : v \in R(B_T)\}$ is a minimum-weight r_T -arborescence such that $R(\tilde{B}_T[T]) = \text{supp}^+(-\partial\xi|_T)$.

The following problems (P') and (D'), whose variables are $x' \in \mathbb{R}^{A_T}$ and $\rho \in \mathbb{R}^{2^T}$, are a linear programming formulation of the minimum-weight r_T -arborescence problem in D_T and its dual program, respectively [4, 14]:

$$3(P') \quad \text{Minimize } \sum_{a \in A_T} w'(a)x'(a)$$

$$\text{subject to } \sum_{a \in \delta^-v} x'(a) = 1 \quad (v \in T), \tag{40}$$

$$\sum_{a \in \delta^-T'} x'(a) \geq 1 \quad (T' \subseteq T, |T'| \geq 2), \tag{41}$$

$$x'(a) \geq 0 \quad (a \in A_T). \tag{42}$$

$$\begin{aligned}
 \text{(D')} \quad & \text{Maximize } \sum_{v \in T} \rho(v) + \sum_{T' \subseteq T, |T'| \geq 2} \rho(T') \\
 \text{subject to } & \rho(v) + \sum_{T': |T'| \geq 2, a \in \delta^- T'} \rho(T') \leq w'(a) \quad (a = uv \in A_T), \\
 & \rho(T') \geq 0 \quad (T' \subseteq T, |T'| \geq 2).
 \end{aligned}
 \tag{43}$$

$$\rho(T') \geq 0 \quad (T' \subseteq T, |T'| \geq 2).
 \tag{44}$$

The complementary slackness conditions for (P') and (D') are as follows:

$$x'(a) > 0 \implies \rho(v) + \sum_{T': |T'| \geq 2, a \in \delta^- T'} \rho(T') = w'(a),
 \tag{45}$$

$$\rho(T') > 0 \implies \sum_{a \in \delta^- T'} x'(a) = 1,
 \tag{46}$$

where $a = uv \in A_T$ in (45) and $T' \subseteq T$ with $|T'| \geq 2$ in (46).

It is known [4, 14] that there exists an integer optimal solution ρ^* for (D') such that $\rho^*(v)$ is nonnegative for all $v \in T$, i.e.,

$$\rho^*(v) \geq 0 \quad (v \in T).
 \tag{47}$$

For example, the arborescence algorithm of Edmonds [4] finds an optimal solution ρ^* such that $\rho^*(v) = \min\{w'(a) : a = uv\}$ for every $v \in T$. Let $\rho^* \in \mathbb{Z}_+^{2T}$ be an integral optimal solution for (D') satisfying (47). Also let \tilde{B}_T be a minimum-weight r_T -arborescence in D_T such that $R(\tilde{B}_T[T]) = \text{supp}^+(-\partial\xi|_T)$ and x' be the characteristic vector of this \tilde{B}_T ; cf. Lemma 5.

Similarly, on the S -side, we consider another directed graph $D_S = (V_S, A_S)$ with arc weight $w'' \in \mathbb{Z}^{A_S}$ defined as

$$\begin{aligned}
 V_S &= \{r_S\} \cup S, \quad A_S = \{ur_S : u \in S\} \cup A[S], \\
 w''(uv) &= \begin{cases} -p(u) & (v = r_S), \\ w(uv) & (v \in S) \end{cases}
 \end{aligned}$$

with a new vertex r_S . We consider an arc subset such that the reversal of its arcs is an r_S -arborescence. Let \tilde{B}_S be such an arc subset of minimum weight that satisfies $R^*(\tilde{B}_S[S]) = \text{supp}^+(\partial\xi|_S)$. Also let $\pi^* \in \mathbb{Z}_+^{2S}$ be an integral optimal solution for the associated dual problem satisfying $\pi^*(u) \geq 0$ for all $u \in S$.

Using π^* and ρ^* above as well as $F = \{a \in A[S, T] : \xi(a) = 1\}$, define $x^* \in \{0, 1\}^A$, $y^* \in \mathbb{Z}^{2S}$, and $z^* \in \mathbb{Z}^{2T}$ by

$$x^* = \chi_{F \cup \tilde{B}_S[S] \cup \tilde{B}_T[T]},
 \tag{48}$$

$$y^*(S') = \pi^*(S') \quad (\emptyset \neq S' \subseteq S),
 \tag{49}$$

$$z^*(T') = \rho^*(T') \quad (\emptyset \neq T' \subseteq T).
 \tag{50}$$

We prove that x^* and (y^*, z^*) are optimal solutions for (P) and (D), respectively.

Lemma 6 x^* and (y^*, z^*) defined in (48), (49), and (50), respectively, are feasible for (P) and (D), respectively.

Proof Since the arc set $F \cup \tilde{B}_S[S] \cup \tilde{B}_T[T]$ is a bibranching in $D = (V, A)$ by $R(\tilde{B}_T[T]) = \text{supp}^+(-\partial\xi|_T)$ and $R^*(\tilde{B}_S[S]) = \text{supp}^+(\partial\xi|_S)$, it is clear that $x^* = \chi_{F \cup \tilde{B}_S[S] \cup \tilde{B}_T[T]}$ is feasible for (P). As for (y^*, z^*) , we first show that it satisfies (4). For $a = uv \in A[S, T]$, by (28) and Lemma 4, we have that $w(a) + p(u) - q(v) \geq 0$, and hence

$$\begin{aligned} \sum_{S': a \in \delta^+ S'} y^*(S') + \sum_{T': a \in \delta^- T'} z^*(T') &= \sum_{S': a \in \delta^+ S'} \pi^*(S') + \sum_{T': a \in \delta^- T'} \rho^*(T') \\ &\leq -p(u) + q(v) \\ &\leq w(a), \end{aligned}$$

where $\sum_{T': a \in \delta^- T'} \rho^*(T') \leq w'(a) = q(v)$ by (43) and the definition of w' , and similarly $\sum_{S': a \in \delta^+ S'} \pi^*(S') \leq w''(a) = -p(u)$. For $a \in A[T]$, it follows from (43) that

$$\sum_{S': a \in \delta^+ S'} y^*(S') + \sum_{T': a \in \delta^- T'} z^*(T') = \sum_{T': a \in \delta^- T'} \rho^*(T') \leq w'(a) = w(a).$$

The case of $a \in A[S]$ can be treated similarly.

Constraint (6) is satisfied by (44) and (47). Similarly (5) is satisfied.

Theorem 5 x^* and (y^*, z^*) defined in (48), (49), and (50), respectively, are optimal solutions for (P) and (D), respectively.

Proof By Lemma 6, it suffices to prove that x^* and (y^*, z^*) satisfy the complementary slackness conditions (7)–(9). To show (7), assume $x^*(a) > 0$. For $a \in A[S, T]$, $x^*(a) > 0$ means $x'(a) = \xi(a) = 1$. Then, it follows from (45), its counterpart for the S -side, and (37) that

$$\sum_{S': a \in \delta^+ S'} y^*(S') + \sum_{T': a \in \delta^- T'} z^*(T') = w''(a) + w'(a) = -p(u) + q(v) = w(a).$$

For $a \in A[T]$, $x'(a) = x^*(a) > 0$ implies that

$$\sum_{S': a \in \delta^+ S'} y^*(S') + \sum_{T': a \in \delta^- T'} z^*(T') = \sum_{T': a \in \delta^- T'} \rho^*(T') = w'(a) = w(a)$$

by (45). The case of $a \in A[S]$ can be treated similarly.

We next consider (9), while noting that (8) can be shown similarly. To show (9), let $z^*(T') > 0$, where $\emptyset \neq T' \subseteq T$. We are to show $x^*(\delta^- T') = 1$.

If $|T'| \geq 2$, (46) with $\rho^*(T') = z^*(T') > 0$ implies $|\tilde{B}_T \cap \delta^- T'| = 1$ in D_T . Denote the unique arc in $\tilde{B}_T \cap \delta^- T'$ by uv . For $v \in T \setminus \text{supp}^+(-\partial\xi|_T)$, it is clear that $x^*(\delta^- T') = 1$, and hence (9) holds. For $v \in \text{supp}^+(-\partial\xi|_T)$, $x^*(\delta^- T') = -\partial\xi(v) \geq 2$

would imply $q(v) = 0$ by Lemma 3, whereas $0 < z^*(T') = \rho^*(T') \leq w'(a) = q(v)$; a contradiction. Hence $x^*(\delta^-T') = 1$ must hold.

When $|T'| = 1$, we have $T' = \{v\}$ for some $v \in T$. If $v \notin \text{supp}^+(-\partial\xi|_T)$, then $x^*(\delta^-v) = 1$ holds since $\tilde{B}_T[T]$ is a branching in $D[T]$. If $v \in \text{supp}^+(-\partial\xi|_T)$, then again $x^*(\delta^-T') = -\partial\xi(v) \geq 2$ would imply $z^*(T') \leq q(v) = 0$ by Lemma 3, a contradiction. Hence $x^*(\delta^-T') = 1$ must hold.

References

1. Babenko, M.A.: An efficient scaling algorithm for the minimum weight bibranching problem. *Algorithmica* **61**, 898–922 (2011)
2. Birge, J.R., Louveaux, F.: *Introduction to Stochastic Programming*, 2nd edn. Springer, New York (2011)
3. Dantzig, G.B., Thapa, M.N.: *Linear Programming, 2: Theory and Extensions*. Springer, New York (2003)
4. Edmonds, J.: Optimum branchings. *J. Res. Natl. Bur. Stand. Sect. B* **71**, 233–240 (1967)
5. Edmonds, J.: Edge-disjoint branchings. In: Rustin, R. (ed.) *Combinatorial Algorithms*, pp. 91–96. Algorithmics Press, New York (1973)
6. Keijsper, J., Pendavingh, R.: An efficient algorithm for minimum-weight bibranching. *J. Comb. Theory Ser. B* **73**, 130–145 (1998)
7. Murota, K.: Valuated matroid intersection I: optimality criteria. *SIAM J. Discret. Math.* **9**, 545–561 (1996)
8. Murota, K.: Valuated matroid intersection II: algorithms. *SIAM J. Discret. Math.* **9**, 562–576 (1996)
9. Murota, K.: Submodular flow problem with a nonseparable convex function. *Combinatorica* **19**, 87–109 (1999)
10. Murota, K.: *Discrete Convex Analysis*. Society for Industrial and Applied Mathematics, Philadelphia (2003)
11. Murota, K., Shioura, A.: M-convex function on generalized polymatroid. *Math. Oper. Res.* **24**, 95–105 (1999)
12. Schrijver, A.: Min–max relations for directed graphs. *Ann. Discret. Math.* **16**, 261–280 (1982)
13. Schrijver, A.: Total dual integrality of matching forest constraints. *Combinatorica* **20**, 575–588 (2000)
14. Schrijver, A.: *Combinatorial Optimization-Polyhedra and Efficiency*. Springer, Heidelberg (2003)
15. Takazawa, K.: Optimal matching forests and valuated delta-matroids. In: Günlük, O., Woeginger, G.J. (eds.) *Integer Programming and Combinatorial Optimization: Proceedings of the 15th IPCO*. LNCS, vol. 6655. Springer, Berlin, pp. 404–416 (2011)
16. Takazawa, K.: Shortest bibranchings and valuated matroid intersection. *Jpn. J. Ind. Appl. Math.* **29**, 561–573 (2012)
17. Takazawa, K.: Optimal matching forests and valuated delta-matroids. *SIAM J. Discret. Math.* **28**, 445–467 (2014)
18. Takazawa, K.: The b -bibranching problem: TDI system, packing, and discrete convexity. *CoRR* (2018). [arXiv:1802.03235](https://arxiv.org/abs/1802.03235)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.