Node-based valid inequalities for
the optimal transmission switching problem

Santanu S. Dey,* Burak Kocuk,† Nicole Redder‡

December 28, 2021

Abstract

The benefits of transmission line switching are well-known in terms of reducing operational cost and improving system reliability of power systems. However, finding the optimal power network configuration is a challenging task due to the combinatorial nature of the underlying optimization problem. In this work, we identify a certain "node-based" set that appears as substructure of the optimal transmission switching problem and then conduct a polyhedral study of this set. We construct an extended formulation of the integer hull of this set and present the inequality description of the integer hull in the original space in some cases. These inequalities in the original space can be used as cutting-planes for the transmission line switching problem. Finally, we present the results of our computational experiments using these cutting-planes on difficult test cases from the literature.

1 Introduction

The study of short-term electric power systems planning offers a wide range of interesting optimization problems for the operations research community [9]. One of the fundamental optimization problems from this area is called the alternating current optimal power flow (AC-OPF) problem. The main goal in this problem is to determine a minimum cost flow across the power network that satisfies demand while following physics laws such as the Ohm's law and Kirchhoff's law. However, the AC-OPF problem is very challenging to solve since the power flow constraints are nonlinear and nonconvex, the underlying networks are very large, and finally AC-OPF needs to be solved rapidly for changing power demands. Therefore, the constraints are typically approximated by direct current (DC) power equations, a linearization of the AC equations, as in [2], [1], [4], [15], and the resulting model is called as the DC-OPF problem.

Because of the underlying physics of electricity, removing lines from a network may result in improved network efficiency. It is therefore possible that switching off lines may reduce the

---

*santanu.dey@isye.gatech.edu, H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

†burak.kocuk@sabanciuniv.edu, Industrial Engineering Program, Sabancı University, Istanbul, Turkey 34956.

‡nredder@gatech.edu, H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

generation cost [12]. In this way, we arrive at the optimal transmission switching (OTS) problem, formalized by Fisher et al. [2]. The OTS problem is the AC-OPF problem modified by the additional choice of switching lines off. With the DC approximation DC-OPF, OTS can be modeled as a mixed-integer linear program (MILP), in which binary variables model the switching status of the network lines; and power flow constraints are only applied to active lines. Henceforth, in the paper we refer to the DC version of optimal switching problem as DC-OTS.

DC-OTS is NP-hard to solve [10, 8]. Since the linear relaxations are often weak (involving many big-M constraints), DC-OTS model with many binary variables are often very hard for modern computational integer programming software. To the best of our knowledge the only paper that conducts a polyhedral study of a substructure of DC-OTS in order to improve the quality of linear programming relaxation is [8]. We emphasize that finding high quality feasible solutions is also quite difficult (see, for example, a large number of heuristics proposed for this problem [6, 11, 13, 3, 7]). Therefore, the DC-OTS problem is a very challenging problem to solve exactly.

In this paper, we identify a certain "node-based" set that appears as substructure of the DC-OTS problem and then conduct a polyhedral study of this set. We construct an extended formulation of the integer hull of this set. Then, we present the inequality description of the integer hull in the original space in some special (but commonly occurring in practice) cases. These inequalities in the original space can be used as cutting-planes for the DC-OTS problem. Finally, we present the results of our computational experiments using these cutting-planes on difficult test cases from the literature.

The rest of the paper is organized as follows: In Section 2 we formally introduce the MILP formulation of the DC-OTS problem. In Section 3, we identify the "node-based" set that appears as substructure of the DC-OTS problem and show that optimizing on this set is NP-hard. Then in Section 4, we present an extended formulation of the integer hull of the "node-based" set and identify conditions under which this extended formulation is compact. Since the extended formulation is exponential in size in general, we present a compact outer approximation of the integer hull in Section 5. We then show that the outer approximation yields the convex hull in some special cases. In Section 6, we present out computational results. Finally, we make some concluding remarks in Section 7.

Notation: Given a set S, we denote its convex hull, (affine) dimension, set of extreme points as $\mathrm{conv}(S)$, $\dim(S)$, $\mathrm{extr}(S)$, respectively. We will use $[m:n]$ as a shorthand notation for the set $\{m, m+1, \ldots, n\}$, where $m, n \in \mathbb{Z}$, the set of integers.

2 DC optimal transmission switching problem

A power network is a set of nodes called *buses* and edges called *transmission lines* within a power system. Consider a power network $\mathcal{N} = (\mathcal{B}, \mathcal{L})$, where $\mathcal{B}$ represents the node set and $\mathcal{L}$ represents the edge set. The subset of buses $\mathcal{G} \subseteq \mathcal{B}$ represents buses connected to generation units (power generators), and we assume that every bus holds some electric demand, called load. The goal of

DC-OTS is to satisfy demand at all buses with the minimum total production costs of generators such that the solution obeys the physical laws (e.g., Ohm's Law) and other operational restrictions (e.g. transmission line flow limit constraints).

We now describe the parameters of the DC-OTS problem:

- For each bus $i \in \mathcal{B}$, let p_i^d and be the real power load.

- For each generator located at bus $i \in \mathcal{G}$, active output is restricted to be in the interval $[p_i^{\min}, p_i^{\max}]$. We set $p_i^{\min} = p_i^{\max} = 0$ for $i \in \mathcal{B} \setminus \mathcal{G}$.

- For each transmission line $(i, j) \in \mathcal{L}$, the susceptance is denoted as B_{ij} and the thermal limit is denoted as $\overline{s}_{ij}$.

We define the following decision variables to model the DC-OTS problem:

- For each bus $i \in \mathcal{B}$, let θ_i and be the phase angle.

- For each generator located at bus $i \in \mathcal{G}$, let p_i^g be its production output.

- For each transmission line $(i, j) \in \mathcal{L}$, let p_{ij} be its real power flow and x_{ij} be its on/off status.

The DC-OTS problem is modeled as follows:

$$\min \; \sum_{i \in \mathcal{G}} C_i(p_i^g) \tag{1a}$$

$$\text{s.t.} \; p_i^g - p_i^d = \sum_{j \in \delta(i)} p_{ij} \qquad i \in \mathcal{B} \tag{1b}$$

$$p_{ij} = [B_{ij}(\theta_i - \theta_j)]x_{ij} \qquad (i,j) \in \mathcal{L} \tag{1c}$$

$$|p_{ij}| \leq \overline{s}_{ij} x_{ij} \qquad (i,j) \in \mathcal{L} \tag{1d}$$

$$p_i^{\min} \leq p_i^g \leq p_i^{\max} \qquad i \in \mathcal{B}$$

$$x_{ij} \in \{0,1\} \qquad (i,j) \in \mathcal{L}$$

Here, objective function (1a) minimizes the total cost of production, where C_i is a linear or convex quadratic cost function of generator i. Constraint (1b) enforces flow conservation at each bus i. Real power across line (i, j), namely p_{ij}, is expressed in terms of the phase angles at buses i and j through constraint (1c), which is an approximation of the true physical function. Finally, the flow limit on line (i, j) is imposed by constraint (1d).

Note that constraint (1c) can be linearized using a big-M formulation to obtain a MILP formulation of DC-OTS as follows:

$$p_{ij} \leq [B_{ij}(\theta_i - \theta_j)] + 2\pi B_{ij}(1 - x_{ij}) \; (i,j) \in \mathcal{L}$$

$$p_{ij} \geq [B_{ij}(\theta_i - \theta_j)] - 2\pi B_{ij}(1 - x_{ij}) \; (i,j) \in \mathcal{L}.$$

Here, we use an upper bound of 2π on $\theta_i - \theta_j$.

3   A "node-based" substructure of DC-OTS

In this section, we start our polyhedral study of a "node-based" substructure of DC-OTS. Let $\bar{f}_j \in \mathbb{R}_+$, $\underline{f}_j = -\bar{f}_j$, for $j \in [1:n]$, $\underline{f}_0 \leq \bar{f}_0$ and $d \in \mathbb{R}$. For convenience, let $x_0 = 1$. Consider the set

$$\mathcal{S} = \left\{ (x, f) \in \{0,1\}^n \times \mathbb{R}^{n+1} : \sum_{j=0}^{n} f_j = d, \ \underline{f}_j x_j \leq f_j \leq \bar{f}_j x_j \ j \in [0:n] \right\}. \tag{2}$$

This mixed-integer set appears as a substructure in the DC-OTS problem as follows: Consider a bus in the power network with a generator[1] and n incident transmission lines. Let d represent the load at node 0, f_0 represent the dispatch variable with bounds $\underline{f}_0$ and $\bar{f}_0$, f_j represent the flow variable of line j with a bound of $\bar{f}_j$, and x_j represent the on/off status of line j, $j \in [1:n]$. See Figure 1 for an illustration.

Figure 1: Illustration with $n = 4$.

Our aim is to understand the complexity of optimizing a linear function over the set $\mathcal{S}$ and to study its polyhedral properties (e.g. obtain its convex hull or a tight polyhedral outer-approximation).

Proposition 1. *Optimizing a linear function over the set $\mathcal{S}$ defined in* (2) *is NP-Hard.*

Proof. Consider the following decision problem, which we will call DC − Node: Does there exist $(x, f) \in \mathcal{S}$ such that $\alpha^\top x + \beta^\top f \leq \gamma$? We claim that DC − Node is NP-Complete, from which the statement of the proposition follows.

We prove the NP-Completeness of DC − Node by a reduction from Subset − Sum, which is known to be NP-Complete [5]. Consider an instance of Subset − Sum as follows: Given $a \in \mathbb{Z}_{++}^n$ and $b \in \mathbb{Z}_{++}$, does there exist a subset $J \subseteq [1:n]$ such that $\sum_{j \in J} a_j = b$? We construct an instance of DC − Node as below:

$$\underline{f}_0 = \bar{f}_0 = 0, \ \beta_0 = 0, \ d = b, \ \gamma = 0, \ \bar{f}_i = -\underline{f}_i = a_i, \ \alpha_i = a_i, \ \beta_i = -1 \quad i \in [1:n].$$

---

[1]A bus without a generator can be simply modeled by setting $\underline{f}_0 = \bar{f}_0 = 0$.

We first note that the size of the DC − Node instance is polynomial in the size of the Subset − Sum instance. We now verify that DC − Node is feasible if and only if Subset − Sum is feasible.

($\Rightarrow$): Let $(\hat{x}, \hat{f}) \in \mathcal{S}$ such that $a^T \hat{x} - e^T \hat{f} \leq 0$ be a solution to DC − Node. Since $a_j \hat{x}_j - \hat{f}_j \geq 0$, we obtain that $\hat{f}_j = a \hat{x}_j$ for $j \in [1 : n]$. Now, set $\hat{J} = \{j : \hat{x}_j = 1\}$. Since $\sum_{j \in \hat{j}} a_j = \sum_{j=1}^n \hat{f}_j = b$, the set $\hat{J}$ is a feasible solution to Subset − Sum.

($\Leftarrow$): Let $\hat{J} \subseteq [1 : n]$ be a solution to Subset − Sum. Then, one can construct a feasible solution to DC − Node as $(\hat{x}_j, \hat{f}_j) = (1, a_j)$ if $j \in \hat{J}$ and $(\hat{x}_j, \hat{f}_j) = (0, 0)$ otherwise. □

We now present an important sub-family of instances for which the optimization can be performed efficiently, since for this sub-family we will show the existence of a polynomial size extended formulation of the integer hull. The proof of Theorem 1 is presented in the next section.

Theorem 1. *Let $\bar{\kappa}$ be a fixed positive integer. Consider the set $\mathcal{S}$ defined in (2) with $\underline{f}_0 = \overline{f}_0 = 0$, $\bar{f}_j = \kappa_j \bar{f}$ where $\kappa_j \in [1 : \bar{\kappa}]$ for $j \in [1 : n]$, $\bar{f} \in \mathbb{R}_{++}$ and $d \in [0, \bar{f})$. Then there exists an algorithm for optimizing a linear function over the set $\mathcal{S}$ whose running time is polynomial in the size of the input.*

Although the requirements of Theorem 1 might seem restrictive at first glance, we note that it is quite common in a power network to have transmission line thermal limits which are multiples of each other. Therefore, the assumptions of Theorem 1 are satisfied for the majority of the substructures that appear in instances from the literature. For example, in 118-bus instances from [8], the fraction of substructures with $\bar{\kappa} = 3$ is about 93% while 86% of the substructures in 300-bus instances can be covered by $\bar{\kappa} = 7$.

4 Proof of Theorem 1 via extended formulation of convex hull of $\mathcal{S}$

In this section, our aim is to obtain the convex hull of $\mathcal{S}$ as an extended formulation using disjunctive arguments. We start our analysis by identifying the extreme points of a certain restricted polyhedral set defined as

$$\mathcal{S}(\hat{x}) = \left\{ f \in \mathbb{R}^{n+1} : \sum_{j \in J(\hat{x})} f_j = d, \ \underline{f}_j \leq f_j \leq \bar{f}_j \ j \in J(\hat{x}), \ f_j = 0 \ j \notin J(\hat{x}) \right\}, \tag{3}$$

where $\hat{x} \in \{0, 1\}^n$ and $J(\hat{x}) := \{0\} \cup \{j : \hat{x}_j = 1\}$.

Proposition 2. *Consider the set $\mathcal{S}(\hat{x})$ defined in (3). Then, we have*

$$\text{extr}\left(\mathcal{S}(\hat{x})\right) = \bigcup_{k \in J(\hat{x})} \left\{ f \in \mathbb{R}^{n+1} : f_j \in \{\, \underline{f}_j, \bar{f}_j \}, \ j \in J(\hat{x}) \setminus \{k\}, \right.$$

$$f_k = d - \sum_{j \in J(\hat{x}) \setminus \{k\}} f_j, \underline{f}_k \leq f_k \leq \bar{f}_k,$$

$$\left. f_j = 0, \ j \notin J(\hat{x}) \right\}.$$

Proof. Observe that $\mathcal{S}(\hat{x})$ is a polyhedral set defined in $\mathbb{R}^{n+1}$. Therefore, at an extreme point, there must exist at least $n+1$ linearly independent inequalities satisfied at equality [14]. Since the set $\mathcal{S}(\hat{x})$ has already one equality in its definition, at least n of the bound constraints must be active. Therefore, enumerating over each coordinate gives the desired result. □

Now, we are ready to give the extreme points of $\mathrm{conv}(\mathcal{S})$.

Corollary 1. *Consider the set $\mathcal{S}$ defined in (2). Then, we have*

$$\mathrm{extr}\,(\mathrm{conv}(\mathcal{S})) = \{(x, f) \in \{0, 1\}^n \times \mathbb{R}^{n+1} : f \in \mathrm{extr}\,(\mathcal{S}(x))\}.$$

Proof. The result follows due to the fact that at any extreme point $(\hat{x}, \hat{f})$ of $\mathrm{conv}(\mathcal{S})$, we must have $\hat{x} \in \{0, 1\}^n$ and using Proposition 2. □

4.1 An extended formulation for $\mathrm{conv}(\mathcal{S})$

Below, we describe a way to obtain $\mathrm{conv}(\mathcal{S})$ through an extended formulation via a special layered network. To start with, let us define a set V_k, which contains the possible values the variable f_k can take between its bounds in one of the extreme points of $\mathrm{conv}(\mathcal{S})$, for $k \in [0 : n]$. In particular, we have

$$V_k = \left\{ f_k = d - \sum_{j \in [0:n] \setminus \{k\}} f_j : f_j \in E_j \ j \in [0 : n] \setminus \{k\}, \ \underline{f}_k \le f_k \le \overline{f}_k \right\},$$

for $k \in [0 : n]$, where

$$E_j := \begin{cases} \{\underline{f}_0, \overline{f}_0\} & \text{if } j = 0 \\ \{-\bar{f}_j, 0, \bar{f}_j\} & \text{if } j \in [1 : n]. \end{cases}$$

We note that each set V_k can be obtained via enumeration.

We now describe a way to obtain $\mathrm{conv}(\mathcal{S})$ through an extended formulation via a special network with $n + 2$ layers, denoted as $N = (\cup_{i=-1}^n K_i, A)$. Here, K_i is the set of nodes in layer i and A is the set of arcs. Layer -1 contains the source node and layer n contains the sink nodes. Each node in the network is designated by a triplet (i, δ, S), where i is the layer index, δ is the amount of demand satisfied until layer i and S is the set of coordinates selected so far that take value between their respective bounds. The critical element in this construction is the fact that $|S| \le 1$ due to Proposition 2 and Corollary 1. We use Algorithm 1 to construct this network, which we use to represent the extreme points of $\mathrm{conv}(\mathcal{S})$ using source to sink paths. See Figure 2 for an illustration.

6

Algorithm 1 Network construction for the extended formulation of conv($\mathcal{S}$).

Let $K_{-1} = \{(-1, 0, \emptyset)\}$, $K_i = \emptyset$, $i = 0, \ldots, n$, and $A = \emptyset$.
for $i = 0, \ldots, n$ **do**
  Set $\underline{B}_i = \max\{\sum_{j=0}^{i} \underline{f}_j, d - \sum_{j=i+1}^{n} \bar{f}_j\}$ and $\overline{B}_i = \min\{\sum_{j=0}^{i} \bar{f}_j, d - \sum_{j=i+1}^{n} \underline{f}_j\}$.
  for each $(i-1, \delta, S) \in K_{i-1}$ **do**
    for each $\vartheta \in E_i$ **do**
      if $\underline{B}_i \leq \delta + \vartheta \leq \overline{B}_i$ **then**
        $K_i = K_i \cup \{(i, \delta + \vartheta, S)\}$ and $A = A \cup \{(i-1, \delta, S) \to (i, \delta + \vartheta, S)\}$.
    if $S = \emptyset$ **then**
      for each $\vartheta \in V_i$ **do**
        if $\underline{B}_i \leq \delta + \vartheta \leq \overline{B}_i$ **then**
          $K_i = K_i \cup \{(i, \delta + \vartheta, \{i\})\}$ and $A = A \cup \{(i-1, \delta, \emptyset) \to (i, \delta + \vartheta, \{i\})\}$.

Figure 2: Illustration of the nodes of the layered network.

Let us define a set of binary variables $Y_{(i-1,\delta,S),(i,\delta',S')}$ which takes value 1 if the arc $(i-1, \delta, S) \to (i, \delta', S')$ is traversed, and 0 otherwise. Consider the following system:

$$\sum_{\eta \in K_0} Y_{(-1,0,\emptyset),\eta} = 1 \tag{4a}$$

$$\sum_{\eta \in K_{i-1}} Y_{\eta,\eta'} = \sum_{\eta'' \in K_{i+1}} Y_{\eta',\eta''} \qquad i \in [1 : n-1], \eta' \in K_i \tag{4b}$$

$$x_i = 1 - \sum_{(i-1,\delta,S) \in K_{i-1}, (i,\delta,S) \in K_i} Y_{(i-1,\delta,S),(i,\delta,S)} \qquad i \in [1 : n] \tag{4c}$$

$$f_i = \sum_{(i-1,\delta,S) \in K_{i-1}, (i,\delta',S') \in K_i} (\delta' - \delta) Y_{(i-1,\delta,S),(i,\delta',S')} \qquad i \in [0 : n]. \tag{4d}$$

Proposition 3. *Consider the set $\mathcal{S}$ defined in (2) and the layered network $N = (\cup_{i=-1}^{n} K_i, A)$ constructed using Algorithm 1. Then, we have that*

$$\text{conv}(\mathcal{S}) = \{(x, f) \in [0, 1]^n \times \mathbb{R}^{n+1} : \exists Y \geq 0 : (4)\}.$$

Proof. First, we note that there is a one-to-one correspondence between extreme points of conv($\mathcal{S}$)

and source-to-sink paths in N. Also, the constraint matrix of (4) is totally unimodular in Y. This is due to the fact that constraints (4a)-(4a) are defined by a network matrix, which is known to be totally unimodular, and constraints (4c)-(4d) are simply the definitions of the x and f variables in terms of the Y variable. Hence, the result follows. $\square$

4.2 Proof of Theorem 1: A special case when the extended formulation is compact

In the previous section, we constructed an extended formulation for $\text{conv}(\mathcal{S})$. However, this formulation can be of exponential-size in the worst case as optimizing a linear function over the set $\mathcal{S}$ is NP-Hard due to Proposition 1. Now, we will consider an important sub-family of instances for which the size of the extended formulation is polynomial in n with fixed parameter $\bar{\kappa}$.

Proposition 4. *Under the assumptions of Theorem 1, there exists a compact extended formulation for $\text{conv}(\mathcal{S})$ whose size is polynomial in n and $\bar{\kappa}$.*

Proof. First of all, we can assume without loss of generality that $\bar{f} = 1$ (otherwise, we can pass to new variables $f'_j = f_j/\bar{f}$). Consider the extended formulation given in Proposition 3 constructed via Algorithm 1. It suffices to show that the number of nodes in the layered network is upper bounded by a polynomial in n and $\bar{\kappa}$.

Since $\underline{f}_j, \overline{f}_j \in \mathbb{Z}$, we have that δ is either an integer or a number of the form $p + d$ where $p \in \mathbb{Z}$. Notice also that $n\bar{\kappa} \leq \underline{B}_i := \max\{\sum_{j=0}^{i} \underline{f}_j, d - \sum_{j=i+1}^{n} \overline{f}_j\}$, since $d \geq 0$ and $-\underline{f}_j = \overline{f}_j = \kappa_j \leq \bar{\kappa}$. Similarly, $n\bar{\kappa} + 1 \geq \overline{B}_i := \min\{\sum_{j=0}^{i} \overline{f}_j, d - \sum_{j=i+1}^{n} \underline{f}_j\}$. Thus $|K_i| \leq 4n\bar{\kappa}$. Therefore, $\sum_{j=-1}^{n} |K_j| \leq 4n^2\bar{\kappa}$, which is a polynomial in n, and $\bar{\kappa}$ is a fixed integer. $\square$

We are now ready to prove Theorem 1.

Proof of Theorem 1. The statement of the proposition follows from the facts that i) optimizing a linear function over $\mathcal{S}$ is equivalent to optimizing the same function over $\text{conv}(\mathcal{S})$, ii) there is a compact polyhedral extended formulation for $\text{conv}(\mathcal{S})$ due to Proposition 4, and iii) linear programming is polynomially solvable. $\square$

5 An outer-approximation for convex hull of $\mathcal{S}$

Since the size of the extended formulation for $\text{conv}(\mathcal{S})$ can be of exponential-size, we explore the possibility of obtaining a compact outer-approximation in this section.

5.1 A compact extended formulation

We start with the following proposition.

Proposition 5. *Consider the set $\mathcal{S}$ defined in* (2), *and define*

$$\mathcal{O} = \bigcup_{k=0}^{n} \left\{ (x, f) \in [0,1]^n \times \mathbb{R}^{n+1} : \sum_{j=0}^{n} f_j = d, \; \underline{f}_j x_j \leq f_j \leq \bar{f}_j x_j \; j \in [0:n], \; f_k \in \text{conv}(V_k), \; x_k = 1 \right\}.$$

(5)

Then, we have $\text{conv}(\mathcal{S}) \subseteq \text{conv}(\mathcal{O})$.

Proof. The result follows since the extreme points of $\text{conv}(\mathcal{S})$ are contained in the set $\mathcal{O}$ by construction. $\square$

We note that the size of the polyhedral representation of $\text{conv}(\mathcal{O})$ is on the order of $O(n^2)$.

5.2 A special case when the outer-approximation is exact

We will now discuss a special case for which i) $\text{conv}(\mathcal{O}) = \text{conv}(\mathcal{S})$ and ii) $\text{conv}(\mathcal{S})$ can be described in the original space.

Lemma 1. *Let* $(\bar{f}, d') \in \mathbb{R}_+ \times \mathbb{R}$ *such that* $d' = \kappa \bar{f}$ *for some* $\kappa \in [-m+1:m]$. *Then, the following polytope is integral in x:*

$$\left\{ (x, f) \in [0,1]^m \times \mathbb{R}^m : d' - \bar{f} \leq \sum_{j=1}^{m} f_j \leq d', \; -\bar{f} x_j \leq f_j \leq \bar{f} x_j \; j \in [1:m] \right\}.$$

Proof. First of all, we can assume without loss of generality that $\bar{f}_j = 1$, $j = 1, \ldots, m$ (otherwise, we can pass to new variables $f'_j = f_j / \bar{f}$). Suppose that there is a fractional extreme point of the polytope, denoted by $(\hat{x}, \hat{f})$, with $\hat{x}_i \in (0, 1)$ for some i. We will look at two cases:

<u>Case 1</u>. Suppose $\sum_{j=1}^{m} \hat{f}_j = d'$ holds (the case with $\sum_{j=1}^{m} \hat{f}_j = d' - 1$ can be handled similarly). First note that the point $(\hat{x}, \hat{f})$ with $-\hat{x}_i < \hat{f}_i < \hat{x}_i$ cannot be extreme, since it can be written as a convex combination of points with $\hat{x}_i \pm \epsilon$. Therefore, we must have that $\hat{f}_i = \pm \hat{x}_i$. Let us assume that $\hat{f}_i = \hat{x}_i$ (the case with $\hat{f}_i = -\hat{x}_i$ is similar). Since d' is integer, there must exist another index k such that $\hat{f}_k \in (-1, 1)$. Now, we construct two new points $(x^{\pm}, f^{\pm}) = (\hat{x} \pm \epsilon e_i \mp \epsilon e_k, \hat{f} \pm \epsilon e_i \mp \epsilon e_k)$ with $\epsilon > 0$ small enough such that $(x^{\pm}, f^{\pm})$ belong to the polytope (here, e_i is the i−th unit vector). However, this contradicts to the fact that $(\hat{x}, \hat{f})$ is an extreme point.

<u>Case 2</u>. Suppose $d' - \bar{f} < \sum_{j=1}^{m} \hat{f}_j < d'$ holds. As before, we must have that $\hat{f}_i = \pm \hat{x}_i$. Let us assume that $\hat{f}_i = \hat{x}_i$ (the case with $\hat{f}_i = -\hat{x}_i$ is similar). Now, we construct two new points $(x^{\pm}, f^{\pm}) = (\hat{x} \pm \epsilon e_i, \hat{f} \pm \epsilon e_i)$ with $\epsilon > 0$ small enough such that $(x^{\pm}, f^{\pm})$ belong to the polytope. However, this contradicts to the fact that $(\hat{x}, \hat{f})$ is an extreme point. $\square$

Proposition 6. *Consider the set $\mathcal{S}$ defined in* (2) *with* $\underline{f}_0 = \bar{f}_0 = 0$, $\bar{f}_j = \bar{f}$ *for* $j \in [1:n]$, $0 \leq d < \bar{f}$, *and the set $\mathcal{O}$ defined in* (5). *Then, we have* $\text{conv}(\mathcal{O}) = \text{conv}(\mathcal{S})$.

Proof. Since $\underline{f}_0 = \bar{f}_0 = 0$, $\bar{f}_j = \bar{f}$ for $j \in [1:n]$, we have that $V_k = \{d, -\bar{f} + d\}$ for $k \in [1:n]$. Now, in order to prove that $\text{conv}(\mathcal{O}) = \text{conv}(\mathcal{S})$, it suffices to show that $\text{conv}(\mathcal{O})$ is integral in x.

We first note that the following polytope in integral in x for $\hat{f}_k \in V_k$ as a consequence of Lemma 1:

$$\left\{ (x, f) \in [0,1]^n \times \mathbb{R}^{n+1} : \sum_{j=1}^{n} f_j = d, \ -\bar{f}x_j \le f_j \le \bar{f}x_j \ j \in [1:n], \ f_k \in [-\bar{f} + d, d], \ x_k = 1 \right\}.$$

Since $\mathrm{conv}(\mathcal{O})$ is the convex hull of the disjunctive union of integral polytopes (see Equation (5)), it follows that $\mathrm{conv}(\mathcal{O})$ is integral as well, and the result follows. $\square$

Note that the conditions under which the outer approximation $\mathrm{conv}(O)$ is equal to $\mathrm{conv}(\mathcal{S})$ are a special case of conditions presented in Theorem 1 when we can show that the problem can be solved in polynomial-time.

5.3   The description of $\mathrm{conv}(\mathcal{S})$ in the original space under the special case

In the setting of the previous subsection, it is also possible to obtain the convex hull description in the original set of variables. Let us define the following inequalities:

$$\sum_{j \in J_1} (\bar{f} - d)(\bar{f}x_j + f_j) + \sum_{j \in J_2} (\bar{f} - d)dx_j + \sum_{j \in J_3} d(\bar{f}x_j - f_j) \ge (\bar{f} - d)d, \tag{6}$$

$$(J_1, J_2, J_3) \text{ is a partition of } [1:n].$$

Theorem 2. *Under the assumptions of Proposition 6, the inequalities* (6) *are valid for* $\mathrm{conv}(\mathcal{S})$.

Proof. We first prove the validity of inequalities (6) by checking them at each binary vector $\hat{x} \in \{0,1\}^n$. Let us define $J_i' := \{j : \ \hat{x}_j = 1, j \in J_i\}$, $i = 1, 2, 3$. If $J_2' \ne \emptyset$, i.e. there exists $j^* \in J_2'$ such that $x_{j^*} = 1$, then inequality (6) is trivially valid since the expressions $\bar{f}x_j + f_j$ and $\bar{f}x_j - f_j$ are nonnegative by the bound constraints. Now, suppose that $J_2' = \emptyset$. In this case, it suffices to show that

$$\sum_{j \in J_1' \cup J_3'} f_j = d, \ -\bar{f} \le f_j \le \bar{f}, j \in J_1' \cup J_3' \implies \sum_{j \in J_1'}(\bar{f} - d)(\bar{f} + f_j) + \sum_{j \in J_3'} d(\bar{f} - f_j) \ge (\bar{f} - d)d$$

$$\iff \sum_{j \in J_1'} f_j \ge d - (\bar{f} - d)|J_1'| - d|J_3'|,$$

where the equivalence follows due to the linear equality. We will proceed by showing that

$$\min_f \left\{ \sum_{j \in J_1'} f_j : \ \sum_{j \in J_1' \cup J_3'} f_j = d, \ -\bar{f} \le f_j \le \bar{f}, j \in J_1' \cup J_3' \right\} \ge d - (\bar{f} - d)|J_1'| - d|J_3'|, \tag{7}$$

by examining two cases.

<u>Case 1</u>. $|J_1'| \ge |J_3'|$: In this case, any optimal solution of the optimization problem in (7) satisfies $\sum_{j \in J_3'} f_j^* = |J_3'|\bar{f}$ and $\sum_{j \in J_1'} f_j^* = d - |J_3'|\bar{f}$. Hence, we obtain that

$$\sum_{j \in J_1'} f_j^* = d - |J_3'|\bar{f} \ge d - |J_3'|\bar{f} - (\bar{f} - d)(|J_1'| - |J_3'|) = d - (\bar{f} - d)|J_1'| - d|J_3'|.$$

10

<u>Case 2</u>. $|J_1'| < |J_3'|$: In this case, any optimal solution of the optimization problem in (7) satisfies $\sum_{j \in J_1'} f_j^* = -|J_1'|\bar{f}$ and $\sum_{j \in J_3'} f_j^* = d + |J_1'|\bar{f}$. Hence, we obtain that

$$\sum_{j \in J_1'} f_j^* = -|J_1'|\bar{f} \geq -|J_1'|\bar{f} - d(|J_3'| - |J_1'| - 1) = d - (\bar{f} - d)|J_1'| - d|J_3'|.$$

$\square$

Theorem 3. *Let us define*

$$X := \left\{ (x, f) \in [0, 1]^n \times \mathbb{R}^{n+1} : (6),\ f_0 = 0,\ \sum_{j=1}^{n} f_j = d,\ -\bar{f}_j x_j \leq f_j \leq \bar{f}_j x_j\ j \in [1 : n] \right\}.$$

Under the assumptions of Proposition 6, we have $\mathrm{conv}(\mathcal{S}) = X$.

Proof. We first claim that $\dim(X) = \dim(\mathrm{conv}(\mathcal{S})) = 2n - 1$. In fact, let us consider the set of points

$$T = \left\{ (x, f) : (f_j, x_j) = \begin{cases} (d, 1), & j = i \\ (0, 0), & j \neq i \end{cases}, j \in [1 : n] \right\} \cup \left\{ (x, f) : (f_j, x_j) = \begin{cases} (d, 1), & j = i \\ (1, 1), & j = (i+1) \mod n \\ (-1, 1), & j = (i+2) \mod n \end{cases} \right\}.$$

Since $|T| = 2n$, the points of T are affinely independent and contained in $\mathcal{S}$ and X, we have that $\dim(\mathrm{conv}(\mathcal{S})) \geq 2n - 1$ and $\dim(X) \geq 2n - 1$. As the description of $\mathcal{S}$ and X contains an equation, $\dim(\mathcal{S}) \leq 2n - 1$ and $\dim(X) \leq 2n - 1$. Hence, we conclude that $\dim(X) = \dim(\mathrm{conv}(\mathcal{S})) = 2n - 1$.

Our proof strategy is to show that for any $c = [c^x, c^f] \in \mathbb{R}^{2n}$, *all extreme optimal solutions* of the following problem

$$M(c) := \mathrm{argmin} \left\{ \sum_{j=1}^{n} (c_j^x x_j + c_j^f f_j) : (x, f) \in \mathrm{conv}(\mathcal{S}) \right\}, \tag{8}$$

satisfy some inequality defining X at equality. One of the key results we repeatedly use in the proof is the characterization of the extreme points of $\mathrm{conv}(S)$ established in Proposition 2 and Corollary 1, which implies the following: In an extreme solution, we have that $x_k \in \{0, 1\}$ for all $k \in [n]$, and there exists $i \in [1 : n]$ such that $f_i \in \{d, -(1 - d)\}$, $x_i = 1$ and $f_j = \pm x_j$ for all $j \in [1 : n] \setminus \{i\}$.

We start the proof with the following simple fact.

Fact 1. *Suppose that the cost vector $[c^x, c^f]$ has an index $j \in [1 : n]$ such that $c_j^x < 0$. In this case, all optimal solutions satisfy $x_j = 1$.*

Due to Fact 1, we focus on the non-trivial case in which $c_j^x \geq 0$ for all $j \in [1 : n]$ in the remainder

11

of the proof. Now, let us define

$$\mathcal{I} := \operatorname{argmin}(c_i^x + c_i^f), \ \mathcal{J} := \operatorname{argmin}(c_j^x - c_j^f), \text{ and } z := c_i^x + c_i^f + c_j^x - c_j^f, i \in \mathcal{I}, j \in \mathcal{J}.$$

We have three cases based on the sign of z.

<u>Case 1</u>. $z > 0$:

<u>Claim 1</u> Consider any extreme point $(x, f) \in \operatorname{conv}(\mathcal{S})$ satisfying that there exist $i, j \in [1 : n]$ such that $f_i = 1, f_j = -1$. We claim that such an extreme point is not optimal when $z > 0$: Indeed, let $x_k^* = x_k, k \notin \{i, j\}$ and $f_k^* = f_k, k \notin \{i, j\}$; and $f_k = x_k = 0, k \in \{i, j\}$. Then (x^*, f^*) is feasible, and observe that

$$\sum_{i=1}^{n}(c_i^x x_i^* + c_i^f f_i^*) + z \le \sum_{i=1}^{n}(c_i^x x_i + c_i^f f_i) \Rightarrow \sum_{i=1}^{n}(c_i^x x_i^* + c_i^f f_i^*) < \sum_{i=1}^{n}(c_i^x x_i + c_i^f f_i),$$

thus proving that (x, f) is not an optimal solution.

Let $A = \operatorname{argmin}(c_i^x + dc_i^f), B = \operatorname{argmin}(c_i^x + c_i^f) = \mathcal{I}, C = \operatorname{argmin}(c_i^x - (1 - d)c_i^f)$; and let $\tilde{a}, \tilde{b}, \tilde{c}$ be their respective values. Based on the above claim, it is easy to verify that all extreme points of $M(c)$ satisfy either:

(i) Exactly one arc i has $x_i = 1$, and $i \in A$.

(ii) Exactly two arcs i, j exist such that their corresponding x is positive, and $i \in B, j \in C$.

   Now, we consider three sub-cases:

- If $\tilde{a} < \tilde{b} + \tilde{c}$, then we are in situation (i) above, i.e., exactly one arc has $x_i = 1$ and i is in set A. Thus in any optimal solution $(x, f) \in M(c)$ we have $\sum_i x_i = 1$.

  Therefore, we can assign $J_2 = [1 : n]$; the associated constraint from (6) is satisfied at equality for all points in $M(c)$.

- If $\tilde{b} + \tilde{c} < \tilde{a}$. We first verify that in this case $B \cap C = \emptyset$. Assume by contradiction that there exists $i \in B \cap C$. Observe then that

$$c_i^x + dc_i^f \ge \tilde{a} > \tilde{b} + \tilde{c} = c_i^x + c_i^f + c_i^x - (1 - d)c_i^f \implies c_i^x < 0,$$

  a contradiction to our assumption that $c_i^x \ge 0$. Thus $B \cap C = \emptyset$. Now it is clear that the optimal solution is of the form (ii) above, i.e., exactly two arcs are active, one each from sets B and C.

  Assign $J_3 = B, J_2 = [1 : n] \setminus B$, and it is straightforward to verify that the corresponding constraint from (6) is satisfied at equality for all points in $M(c)$.

- Otherwise, $\tilde{a} = \tilde{b} + \tilde{c}$.

12

- If $B \cap C = \emptyset$, then we can assign $J_3 = B, J_2 = [1 : n] \setminus B$. If we have a solution of type (i) above with $x_i = 1$ and $f_i = d$ for some $i \in A$, note that whether $i \in J_2$ or $i \in J_3$, the inequality (6) is satisfied at equality. If we have a solution of type (ii) with $x_i, x_j = 1$, $f_i = 1$ and $f_j = d - 1$ with $i \in B$ and $j \in C$, then we have that $i \in J_3$ and $j \in J_2$ and the inequality (6) is satisfied at equality.

- If $B \cap C \neq \emptyset$. In this case we first verify that $|B| = 1$. Let $i \in B \cap C$. Then

$$c_i^x + c_i^f + c_i^x - (1 - d)c_i^f = 2c_i^x + dc_i^f = \tilde{b} + \tilde{c} = \tilde{a} \leq c_i^x + dc_i^f \implies c_i^x = 0. \qquad (9)$$

This implies $\tilde{b} = c_i^x + c_i^f = c_i^f$. Now assume there exists $j \in B$ where $j \neq i$. Then $0 < z \leq c_j^x + c_j^f + c_i^x - c_i^f = \tilde{b} + c_i^x - \tilde{b} = c_i^x$, which contradicts (9).

Again we can assign $J_3 = B, J_2 = [1 : n] \setminus B$. If we have a solution of type (i) above with $x_i = 1$ and $f_i = d$ for some $i \in A$, note that whether $i \in J_2$ or $i \in J_3$, the inequality (6) is satisfied at equality. If $|C| = 1$, then we cannot have a solution of type (ii), since from the above claim we have $B = C$ with $|B| = |C| = 1$ (and therefore any solution of type (ii) will have an objective function value strictly greater than $\tilde{b} + \tilde{c}$). If $|C| > 1$, then we have a solution of type (ii) with $x_i, x_j = 1$, $f_i = 1$ and $f_j = d - 1$ with $i \in B$ and therefore crucially $j \in C \setminus B$. Then we have that $i \in J_3$ and $j \in J_2$ and thus the inequality (6) is satisfied at equality.

<u>Case 2</u>. $z < 0$:

<u>Claim 2</u> We first claim that, if $z < 0$, then $\mathcal{I} \cap \mathcal{J} = \emptyset$. If not, then suppose $i \in \mathcal{I} \cap \mathcal{J}$. Then

$$0 > z = c_i^x + c_i^f + c_i^x - c_i^f = 2c_i^x,$$

which contradicts the assumption that $c_i^x \geq 0$.

We will consider two subcases: either $|\mathcal{J}| < |\mathcal{I}|$ or $|\mathcal{J}| \geq |\mathcal{I}|$.

- Suppose $|\mathcal{J}| < |\mathcal{I}|$. We will prove that $x_j = 1$ for all $j \in \mathcal{J}$ for any $(x, f) \in M(c)$. First observe that $\forall (x, f) \in M(c), \forall i \in \mathcal{I}, j \in \mathcal{J}, x_i + x_j \geq 1$ (otherwise, set both to one for a feasible solution with $f_i = 1$ and $f_j = -1$ which has a strictly smaller objective value). Now suppose there exists $(x, f) \in M(c)$ such that there exists $j \in \mathcal{J}$ with $x_j = 0$ and $x_i = 1$ for all $i \in \mathcal{I}$. Also let $\tilde{J} \subseteq \mathcal{J}$ such that $x_j = 1$ for $j \in \tilde{J}$. Then observe:

  - There does not exist $i \in \mathcal{I}$ such that $f_i = -1$: If there exists $i \in \mathcal{I}$ such that $f_i = -1$, we obtain a better solution by setting $x_i = 0, f_i = 0$ and $x_j = 1$ and $f_j = -1$. Thus, we cannot have $f_i = -1$ for any $i \in \mathcal{I}$ for $(x, y) \in M(c)$.

  - There does not exist $i \in \mathcal{I}$ such that $f_i = d$: If there exists $i \in \mathcal{I}$ such that $f_i = d$ then (due to previous case) $f_u = 1$ for all $u \in \mathcal{I} \setminus \{i\}$. Since $|\tilde{J}| \leq |\mathcal{J}| - 1$, we have

13

$|\tilde{J}| \leq |\mathcal{I} \setminus \{i\}| - 1$. Thus, there exists $k \in [1:n] \setminus (\mathcal{I} \cup \mathcal{J})$ such that $f_k = -1, x_k = 1$, and $c_k^x - c_k^f > c_j^x - c_j^f$, contradicting optimality.

– There does not exist $i \in \mathcal{I}$ such that $f_i = d - 1$: Suppose there exists $i \in \mathcal{I}$ such that $f_i = d - 1$. Based on previous case, $f_u = 1$ for all $u \in \mathcal{I} \setminus \{i\}$. Consider a new solution where we set $x_i = 1$, $f_i = d$, $x_j = 1$ $f_j = -1$ and all other variables have the same value. Then observe that the change in the objective function value is:

$$c_i^x + dc_i^f + c_j^x - c_j^f - (c_i^x + (d-1)c_i^f) = c_j^x - c_j^f + c_i^f \leq c_j^x - c_j^f + c_i^x + c_i^f = z < 0.$$

Thus, there does not exist $i \in I$ such that $f_i = d - 1$.

From the above we have that for all $i \in \mathcal{I}$ we have $f_i = 1$. Now since $|\tilde{J}| \leq |\mathcal{J}| - 1$, we have $|\tilde{J}| \leq |\mathcal{I}| - 2$. Thus, there exists $k \in [1:n] \setminus (\mathcal{I} \cup \mathcal{J})$ such that $f_k = -1, x_k = 1$, and $c_k^x - c_k^f > c_j^x - c_j^f$, contradicting optimality. Thus $x_j = 1$ for all $j \in \mathcal{J}$.

Since $(x, f) \in M(c)$ we have $x_j = 1$ for all $j \in \mathcal{J}$. Thus select some $j \in \mathcal{J}$ and observe that $x_j \leq 1$ is satisfied at equality for all $(x, f) \in M(c)$.

• Otherwise, $|\mathcal{I}| \leq |\mathcal{J}|$. We will prove that for all $i \in \mathcal{I}, x_i = 1$, for all $(x, f) \in M(c)$. Again it is true that $\forall (x, f) \in M(c), \forall i \in \mathcal{I}, j \in \mathcal{J}, x_i + x_j \geq 1$.

Assume there exists an index $i \in \mathcal{I}$ such that $x_i = 0$. Then $x_j = 1$, for all $j \in \mathcal{J}$. Let $\tilde{I} \subseteq \mathcal{I}$ such that $x_i = 1$ for $i \in \tilde{I}$.

Then observe:

– There does not exist $j \in \mathcal{J}$ such that $f_j = 1$: Otherwise, we obtain a better solution by setting $x_j, f_j = 0$ and $f_i, x_i = 1$, contradicting optimality.

– There does not exist $j \in \mathcal{J}$ such that $f_j = d - 1$: If there exists $j \in \mathcal{J}$ such that $f_j = d$ then (due to previous case) $f_u = -1$ for all $u \in \mathcal{J} \setminus \{j\}$. Since $|\tilde{I}| \leq |\mathcal{I}| - 1$, we have $\tilde{I} \leq |\mathcal{J} \setminus \{j\}|$. Thus, there exists $k \in [1:n] \setminus (\mathcal{I} \cup \mathcal{J})$ such that $f_k = 1, x_k = 1$, and $c_k^x + c_k^f > c_i^x + c_i^f$, contradicting optimality.

– There cannot exist $j \in \mathcal{J}$ such that $f_j = d$: Suppose otherwise, for some j. Set $x_i = 1, f_i = 1, f_j = -(1 - d)$. Then the objective function change is

$$c_i^x + c_i^f + c_j^x - (1-d)c_j^f - c_j^x - dc_j^f = c_i^x + c_i^f - c_j^f = z - c_j^x < 0,$$

which contradicts optimality.

Finally, if for all $j \in J$, $f_j = -1$, then there exists $k \in [1:n] \setminus (\mathcal{I} \cup \mathcal{J})$ such that $f_k = 1, x_k = 1$, and $c_k^x + c_k^f > c_i^x + c_i^f$, contradicting optimality. Thus, for $(x, f) \in M(c)$, we have $x_i = 1$ for all $i \in \mathcal{I}$.

Since $(x, f) \in M(c)$, we have $x_i = 1$ for all $i \in \mathcal{I}$ when $|\mathcal{J}| \geq |\mathcal{I}|$. Thus select some $i \in \mathcal{I}$ and observe that $x_i \leq 1$ is satisfied at equality for all $(x, f) \in M(c)$.

Case 3. $z = 0$:

Claim 3

We claim that when $z = 0$, $i \in \mathcal{I} \cap \mathcal{J}$ if and only if $c_i^x = 0$. If there exists $i \in \mathcal{I} \cap \mathcal{J}$, then $c_i^x + c_i^f + c_i^x - c_i^f = z = 0$ implies $c_i^x = 0$. If for some $i \in \mathcal{I} \cup \mathcal{J}$, $c_i^x = 0$, then $c_i^f + c_i^x = c_i^f - c_i^x$, so $i \in \mathcal{I} \cap \mathcal{J}$.

- Suppose $\mathcal{I} \cap \mathcal{J} = \emptyset$. This implies that $c_i^x > 0$ for all $i \in \mathcal{I} \cup \mathcal{J}$ by Claim 3. Consider an extreme point (x, f) of $M(c)$. We will show that there exists a partition of $[1 : n]$ such that the inequality (6) is satisfied at equality.

  - Firstly, for all $k \in \mathcal{I}$, we show $f_k \geq 0$:
    * We will show that if there exists j such that $f_j = -1$, then $j \in \mathcal{J}$. Suppose by contradiction that there exists $j^* \in [1 : n] \setminus \mathcal{J}$ such that $f_{j^*} = -1$. Then because the set of possible flows at extreme points is $\{-1, -(1-d), 0, d, 1\}$ and $\sum_{i=1}^{n} f_i = d$, there must be some i such that $f_i = 1$. Because $\mathcal{J}$ minimizes $c_j^x - c_j^f$ and $z = 0$, $c_i^x + c_i^f + c_{j^*}^x - c_{j^*}^f > 0$, and we could obtain a better solution by setting $x_i, x_{j^*}, f_i, f_{j^*} = 0$, contradicting optimality.
    * We will show that there cannot exist $i \in \mathcal{I}$ such that $f_i = -(1-d)$. Suppose by contradiction that there exists $i^* \in \mathcal{I}$ such that $f_{i^*} = -(1-d)$. By flow conservation, there must exist some j such that $f_j = 1$; and $c_j^x + c_j^f \geq c_{i^*}^x + c_{i^*}^f$, since $\mathcal{I}$ minimizes that value. Suppose we set $x_j, f_j = 0, f_{i^*} = d$. Then the objective value change is

$$c_{i^*}^x + dc_{i^*}^f - (c_j^x + c_j^f + c_{i^*}^x - (1-d)c_{i^*}^f)$$
$$\leq c_{i^*}^x + dc_{i^*}^f - (c_{i^*}^x + c_{i^*}^f + c_{i^*}^x - (1-d)c_{i^*}^f) = -c_{i^*}^x < 0,$$

    which contradicts optimality.

  The two cases above show that $f_k \geq 0$ for all $k \in \mathcal{I}$.

  - For all $k \in \mathcal{J}$, we show that $f_k \leq 0$:
    * We will show that if there exists i such that $f_i = 1$, then $i \in \mathcal{I}$. Suppose by contradiction that there exists $i^* \in [1 : n] \setminus \mathcal{I}$ such that $f_{i^*} = 1$. Firstly, we must have $f_i > 0$ for all $i \in \mathcal{I}$; consider otherwise:
      · If there exists $i \in \mathcal{I}$ such that $f_i = 0$, then we could set $x_i = f_i = 1, f_{i^*} = x_{i^*} = 0$ for a better solution, contradicting optimality.
      · We have already shown that for all $i \in \mathcal{I}$, $f_i \geq 0$.
    Then we have at minimum $1 + d$ units of flow accounted for between i^* and $\mathcal{I}$; so by flow conservation, there must exist $j \in [1 : n]$ such that $f_j = -1$. As before, we can set $x_j, x_{i^*}, f_j, f_{i^*} = 0$ and obtain a better solution, contradicting optimality.
    * We will show that there cannot exist $j \in \mathcal{J}$ such that $f_j = d$. Suppose by contradiction that for some $j^* \in \mathcal{J}$, we have $f_{j^*} = d$.

15

· If there exists $i \in \mathcal{I}$ such that $f_i = 0$, we can set $x_i = 1, f_i = 1, f_{j^*} = -(1-d)$. Then the objective function change is

$$c_i^x + c_i^f + c_{j^*}^x - (1-d)c_{j^*}^f - c_{j^*}^x - dc_{j^*}^f = c_i^x + c_i^f - c_{j^*}^f = z - c_{j^*}^x < 0,$$

which contradicts optimality.

· Otherwise, we have $f_i = 1$ for all $i \in \mathcal{I}$. Then by flow conservation, there must exist j such that $f_j = -1$; and $c_i^x + c_i^f + c_j^f - c_j^f \geq z = 0$. Then we can set $f_i, x_i, f_j, x_j = 0$ for a feasible solution that is not worse; and we have already shown that such a solution cannot be optimal in the previous case. Then this solution cannot be optimal, either.

The two cases above show that $f_k \leq 0$ for all $k \in \mathcal{J}$.

As we have already shown that $f_i = 1 \implies i \in \mathcal{I}$ and $f_i = -1 \implies i \in \mathcal{J}$, then for any $i \in [1:n] \setminus (\mathcal{I} \cup \mathcal{J})$, $-1 < f_i < 1$.

Assign $J_1 = \mathcal{J}, J_2 = [1:n] \setminus (\mathcal{I} \cup \mathcal{J}), J_3 = \mathcal{I}$, and consider the constraint (6). Since $\mathcal{I} \subseteq J_3, \mathcal{J} \subseteq J_1$, when there exists i such that $f_i = \pm 1$, the associated terms cancel out to zero. If there exists i such that $f_i = d$, then $i \in J_2 \cup J_3$, and equality is satisfied; otherwise, there exists i such that $f_i = -(1-d)$, and $i \in J_1 \cup J_2$, again satisfying equality.

- Otherwise there exists $i^* \in \mathcal{I} \cap \mathcal{J}$, and $c_{i^*}^x = 0$.

We claim that for all i with $c_i^x = 0$, we have that $c_i^f = c_{i^*}^f$. For contradiction, suppose otherwise. If $c_i^f < c_{i^*}^f$, then $i^* \in \mathcal{I}$ is contradicted. If $c_i^f > c_{i^*}^f$, then $i^* \in \mathcal{J}$ is contradicted.

We will consider some cases and show that in each of the cases either we can find an inequality describing a facet of X which all extreme points of $M(c)$ lie on, or else $M(c) = \mathcal{S}$. Before presenting the cases, we require the following claim.

Claim 4
When $z = 0$ and $\mathcal{I} \cap \mathcal{J} \neq \emptyset$, if $c_i^x > 0$ for some i, then $f_i \in \{-1, 0, 1\}$. Consider an i such that $c_i^x > 0$. We will show that $f_i \neq d$. Suppose by contradiction that $f_i = d$, and note that $c_i^x + c_i^f \geq c_{i^*}^f$, so $c_i^x + dc_i^f > dc_{i^*}^x$.

- If $f_{i^*} = 0$, we can assign $f_i, x_i = 0$ and $f_{i^*} = d$ for a strictly better solution, contradicting optimality.

- If $f_{i^*} = 1$, then by flow conservation there must exist j such that $f_j = -1$. We can set $f_j, x_j, f_{i^*}, x_{i^*} = 0$, for an objective change not worse than $z = 0$. Then we can set $f_i, x_i = 0$ and $f_{i^*} = d$ for a strictly better solution, contradicting the optimality of the unmodified solution.

- If $f_{i^*} = -1$, then by flow conservation there must exist k such that $f_k = 1$, and we can set $f_k, x_k, f_{i^*}, x_{i^*} = 0$, for an objective change not worse than $z = 0$. Then we can set

16

$f_i, x_i = 0$ and $f_{i^*} = d$ for a strictly better solution, contradicting the optimality of the unmodified solution.

We next show that $f_i \neq -(1-d)$. Suppose by contradiction that $f_i = -(1-d)$, and note that $c_i^x - c_i^f \geq -c_{i^*}^f$, so $c_i^x - (1-d)c_i^f > -(1-d)c_{i^*}^f$.

- If $f_{i^*} = 0$, we can assign $f_i, x_i = 0$ and $f_{i^*} = -(1-d)$ for a strictly better solution, contradicting optimality.

- If $f_{i^*} = -1$, then by flow conservation there must exist k such that $f_k = 1$, and we can set $f_k, x_k, f_{i^*}, x_{i^*} = 0$, for an objective change not worse than $z = 0$. Then we can set $f_i, x_i = 0$ and $f_{i^*} = -(1-d)$ for a strictly better solution, contradicting the optimality of the unmodified solution.

- If $f_{i^*} = 1$, we can assign $f_i, x_i = 0$ and $f_{i^*} = d$; the resulting objective change is $dc_{i^*}^f - (c_{i^*}^f + c_i^x - (1-d)c_i^f) = -(1-d)c_{i^*}^f - (c_i^x - (1-d)c_i^f) < 0$, contradicting optimality.

The above two cases prove that if $c_i^x > 0$ for some i, then $f_i \in \{-1, 0, 1\}$.

Now, we will consider some cases that are analyzed using Claim 4.

- If there exists $i \in [1:n] \setminus (\mathcal{I} \cup \mathcal{J})$, then we will show $x_i = 0$ for all $(x, f) \in M(c)$:
  Suppose there exists such an i and $f_i \neq 0$ for some extreme point of $M(c)$.
  * $f_i \in \{-1, 0, 1\}$ since $c_i^x > 0$.
  * If $f_i = -1$, then by flow conservation there must exist k such that $f_k = 1$; and $c_k^x + c_k^f + c_i^x - c_i^f > 0$ since $i \notin J$. Then we can set $f_k, x_k, f_i, x_i = 0$ to obtain a better solution, contradicting optimality.
  * If $f_i = 1$, and there exists k such that $f_k = -1$, $c_i^x + c_i^f + c_k^x - c_k^f > 0$ since $i \notin \mathcal{I}$. Then we can set $f_k, x_k, f_i, x_i = 0$ to obtain a better solution, contradicting optimality. Otherwise, by flow conservation, there exists some k such that $f_k = -(1-d)$, and all other arcs carry no flow. For all $q \notin \mathcal{I} \cap \mathcal{J}$ we have $c_q^x > 0$, and $-c_{i^*}^f \leq c_q^x - c_q^f$, then $-(1-d)c_{i^*}^f \leq (1-d)c_q^x - (1-d)c_q^f < c_q^x - (1-d)c_q^f$ and we must have $k \in \mathcal{I} \cap \mathcal{J}$; else we have $f_{i^*} = 0$ and we can swap i^* and k for a better solution, contradicting optimality. Then $c_k^x = 0$ and $c_k^f = c_{i^*}^f$. Then we can set $f_i, x_i = 0$ and $f_k = d$, and the change in objective is $dc_{i^*}^f - (c_i^x + c_i^f - (1-d)c_{i^*}^f) = c_{i^*}^f - c_i^x - c_i^f < 0$ since $i \notin \mathcal{I}$, contradicting optimality.

  Then $f_i = 0$ in all optimal solutions. Since $c_i^x > 0$, $x_i = 0$ in all optimal solutions as well. Thus, $x_i \geq 0$ is satisfied at equality for all $(x, f) \in M(c)$.

  Now suppose no such i exists for the remainder of the proof.

- If there exists $i \in \mathcal{I} \cup \mathcal{J}$ such that $c_i^x > 0$, then we will show that we have either $f_i = x_i$ or $f_i = -x_i$, for all $(x, f) \in M(c)$.
  * If $i \in \mathcal{I}$ and $c_i^x > 0$, then we will show that $f_i = x_i$.
    We will show that $f_i \in \{0, 1\}$. Since $c_i^x > 0$, we must have $f_i \in \{-1, 0, 1\}$. Suppose

17

by contradiction that $f_i = -1$. By flow conservation, there must exist k such that $f_k = 1$. Since $i \notin \mathcal{J}$ and $z = 0$, we can set $f_i, x_i, f_k, x_k = 0$ and obtain a better solution, contradicting optimality.

Then $f_i \in \{0, 1\}$ and if $f_i = 0$ then $x_i = 0$ since $c_i^x > 0$, so $f_i = x_i$.

* If $i \in \mathcal{J}$, then we will show that $f_i = -x_i$.

First, we show that $f_i \in \{-1, 0\}$. Since $c_i^x > 0$, we must have $f_i \in \{-1, 0, 1\}$. Then suppose by contradiction that $f_i = 1$.

   · If $f_{i^*} > 0$, then by flow conservation there must exist k such that $f_k = -1$. Then since $i \notin \mathcal{I}$, we can set $f_i, x_i, f_k, x_k = 0$ for a better solution, contradicting optimality.

   · If $f_{i^*} = 0$, then since $i \notin \mathcal{I}$ and $z = 0$, we can set $f_{i^*} = 1$ and $x_i, f_i = 0$ for a better solution, contradicting optimality.

   · If $f_{i^*} = -1$, then since $i \notin \mathcal{I}$ and $z = 0$, we can set $f_{i^*}, f_i, x_i = 0$ for a better solution, contradicting optimality.

   · If $f_{i^*} = -(1 - d)$, then we can set $f_{i^*} = d$ and $f_i, x_i = 0$, and the objective value change is $dc_{i^*}^f - (c_i^x + c_i^f - (1-d)c_{i^*}^f) = c_{i^*}^f - c_i^x - c_i^f < 0$ since $i \notin \mathcal{I}$, contradicting optimality.

Then $f_i \in \{0, -1\}$ and if $f_i = 0$ then $x_i = 0$ since $c_i^x > 0$, so $f_i = -x_i$. Suppose no such i exists for the remainder of the proof.

  – Then otherwise $I \cap J = [1 : n]$ and there exists some $\tilde{C}$ such that $c_i^x = 0, c_i^f = \tilde{C}, \forall i \in [1 : n]$. Then due to flow balance, $M(c) = \mathcal{S}$.

$\square$

Proposition 7. *Inequalities* (6) *can be separated in linear time.*

Proof. We now show that the separation of inequalities (6) can be performed in polynomial time. Given a point $(\hat{x}, \hat{f})$, we define $\hat{\alpha}_j^1 := (\bar{f} - d)(\bar{f}\hat{x}_j + \hat{f}_j)$, $\hat{\alpha}_j^2 := (\bar{f} - d)d\hat{x}_j$ and $\hat{\alpha}_j^3 := d(\bar{f}\hat{x}_j - \hat{f}_j)$, $j = 1, \ldots, n$, and construct the partition $(\hat{J}_1, \hat{J}_2, \hat{J}_3)$ such that

$$\hat{J}_1 := \{j : \hat{\alpha}_j^1 \leq \hat{\alpha}_j^2, \hat{\alpha}_j^1 \leq \hat{\alpha}_j^3\}, \ \ \hat{J}_2 := \{j : \hat{\alpha}_j^2 < \hat{\alpha}_j^1, \hat{\alpha}_j^2 \leq \hat{\alpha}_j^3\}, \ \text{and} \ \hat{J}_3 := \{j : \hat{\alpha}_j^3 < \hat{\alpha}_j^1, \hat{\alpha}_j^3 < \hat{\alpha}_j^2\}.$$

If $\sum_{i=1}^3 \sum_{j \in \hat{J}_i} \hat{\alpha}_j^i < (\bar{f} - d)d$, then this partition corresponds to a most violated inequality and can be obtained in linear time with respect to n. $\square$

6 Computational results

In this section, we explain our computational setting and present our computational results that compare various approaches to solve the DC-OTS problem for challenging instances from the literature.

6.1 Instances

In our computational experiments, we use the instances presented in [8]. There are five groups of 35 instances each and some statistics are reported in Table 1. In each instance, there are $|\mathcal{B}|$ many binary variables and $|\mathcal{B}| + |\mathcal{G}| + |\mathcal{L}|$ many continuous variables.

Instance	$|\mathcal{B}|$	$|\mathcal{L}|$	$|\mathcal{G}|$
118Blumsack_15	118	186	19
118Blumsack_9G	118	186	19
118Blumsack_5p	118	191	19
118Blumsack_15p	118	191	19
300New_5	300	411	61

Table 1: Instance statistics.

We refer the reader to [8] for the details about these instances.

6.2 Computational setting

In our experiments, we use the commercial MILP solver CPLEX 12.8 on a 64-bit personal computer with Intel Core i7 CPU 2.60GHz processor (16 GB RAM). A time limit of 3600 seconds is given for all settings and the relative optimality gap is set to 0.1%. We compare the following four settings in our computational experiments:

- Default: Default CPLEX.

- Feasibility Oriented: Default CPLEX except `MIPEmph=4` and `DiveType=2`. As we mentioned earlier, finding feasible solution is also challenging and we will see that changing CPLEX to this setting improves its performance.

- Feasibility Oriented setting and cuts separated from the convex hull defined in (4) in the support of x and f variables. The cuts are separated using a cut generation linear program. Five rounds of cuts are separated at the root node.

- Feasibility Oriented setting and inequalities (6) are separated. Notice that (6) can always be used by relaxing the upper bounds $\bar{f}$ in the set $\mathcal{S}$ (corresponding to a node). Only the nodes without a generator are considered in this setting due to the assumption of Theorem 3. Again five rounds of cuts are separated at the root node.

6.3 Computational results

We present our computational results in Tables 2-4. The following abbreviations are used in the tables:

- OptTime: Time in seconds used by CPLEX to prove optimality (except time for cut separation).

19

- #Nodes: Number of nodes of the branch-and-bound tree

- SepTime: Time in seconds needed to separate the cuts.

- TotalTime: Total time in seconds to prove optimality.

Table 2 shows results comparing Default CPLEX with Feasibility Oriented CPLEX. Table 3 shows the performance of Feasibility Oriented CPLEX together with cuts separated from the extended formulation. Table 4 shows the performance of Feasibility Oriented CPLEX together with (6) cuts separated.

		Default			Feasibility Oriented		
		Unsolved	OptTime	#Nodes	Unsolved	OptTime	#Nodes
118Blumsack_15	GA	0	5.97	19825	0	6.03	11236
	AA		100.67	586191		7.69	18461
118Blumsack_9G	GA	7	32.78	180040	5	26.10	76669
	AA		686.52	4943363		647.66	2825909
118Blumsack_5p	GA	0	7.33	26880	0	8.06	16441
	AA		19.96	149245		10.32	30145
118Blumsack_15p	GA	5	147.36	527989	4	83.56	254854
	AA		757.16	2484548		481.14	1169934
300New_5	GA	12	192.72	234234	6	78.99	88550
	AA		1284.98	1449385		665.87	698018

Table 2: DC-OTS Results without cuts (GA: Geometric Average, AA: Arithmetic Average).

		Unsolved	SepTime	#Cuts	OptTime	#Nodes	TotalTime
118Blumsack_15	GA	0	0.55	187	6.99	12048	7.71
	AA		0.56	187	29.12	76788	29.68
118Blumsack_9G	GA	2	0.54	185	24.76	69342	26.26
	AA		0.54	185	356.82	1443889	357.36
118Blumsack_5p	GA	0	0.65	192	8.81	16811	9.56
	AA		0.65	192	11.54	31664	12.19
118Blumsack_15p	GA	1	0.58	186	41.44	119906	42.17
	AA		0.58	186	146.05	368832	146.64
300New_5	GA	4	2.17	455	91.54	89363	95.93
	AA		2.18	455	523.17	485283	525.35

Table 3: DC-OTS Results with Feasibility Oriented setting and cuts separated from the convex hull.

		Unsolved	SepTime	#Cuts	OptTime	#Nodes	TotalTime
118Blumsack_15	GA	0	0.02	148	5.26	8250	5.28
	AA		0.02	148	6.06	12511	6.08
118Blumsack_9G	GA	2	0.01	139	28.57	76579	28.62
	AA		0.02	139	422.73	1692401	422.74
118Blumsack_15p	GA	0	0.02	141	8.13	16297	8.15
	AA		0.02	141	10.18	33747	10.20
118Blumsack_15p	GA	2	0.02	137	55.90	153063	55.94
	AA		0.02	137	282.95	597637	282.97
300New_5	GA	3	0.04	152	64.99	62704	65.05
	AA		0.04	152	379.84	402154	379.88

Table 4: DC-OTS Results with Feasibility Oriented setting and inequalities (6) are separated.

We summarize our observations as below:

- Feasibility oriented (12 unsolved) is much better than Default (24 unsolved). All other statistics are improved as well.

- The setting with cuts separated from the convex hull and the setting with inequalities separated from (6) are better than the setting with feasibility oriented (they both solve 5 more instances). Most of the statistics are improved (especially for the harder instances 118Blumsack_9G, 118Blumsack_15p, 300New_5).

- The overall performances of the setting with cuts separated from the convex hull and the setting with inequalities separated from (6) are similar.

  - The latter adds less cuts: This is more advantageous in easier instances 118Blumsack_15 and 118Blumsack_5p.

  - The former adds more cuts: This is more advantageous in harder instances 118Blumsack_9G and 118Blumsack_15p.

  - For 300New_5 instances, the former adds significantly more cuts, which seem to result in a heavier and slower model (one more unsolved).

- The setting with inequalities separated from (6) does not improve the percentage gap closed, defined as

$$100 \times \frac{UB - LP'}{UB - LP},$$

except for three instances from 118Blumsack_9G with negligible improvements (here, UB is the best upper bound found, LP is the lower bound obtained from the LP relaxation and LP' is the lower bound obtained from the LP relaxation with the added inequalities). The setting with cuts separated from the convex hull is slightly more successful with average percentage gap closed reported as 0.08%, 0.38%, 0.08%, 0.00% and 0.16% respectively for the

five instance families. The OTS Problem is known to have dual degeneracy (the cuts separate the current fractional point but there exist other optimal solutions of the LP with the same objective function value) and these results are not very surprising in this respect.

7   Conclusion

We consider a substructure in DC-OTS representing a single bus node and the lines connected to it. We show that optimizing over this substructure is NP-Hard, and we propose an extended formulation for its convex hull and show when this convex hull is of a compact size. We also propose an outer approximation for its convex hull, which is more compact than our extended formulation, and we show that in a special case of the problem, the outer approximation is exactly the convex hull. In this case, we are able to obtain the convex hull in the original space. Computationally, these cuts are able to help solve more instances in lesser time over the commercial solver CPLEX.

We note that while the substructure we studied is motivated from DC-OTS, this set may also be obtained as a substructure in the context of AC-OPF. For example, we can consider a single bus and lines connected to it, together with the real power flowing in these line. It may also be possible to use the results of this study to other network based applications like gas and water networks.

References

[1] C. Barrows, S. Blumsack, and R. Bent. Computationally efficient optimal transmission switching: Solution space reduction. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8, July 2012.

[2] E.B. Fisher, R.P. O'Neill, and M.C. Ferris. Optimal transmission switching. *IEEE Transactions on Power Systems*, pages 1346–1355, 2008.

[3] J David Fuller, Raynier Ramasra, and Amanda Cha. Fast heuristics for transmission-line switching. *IEEE Transactions on Power Systems*, 27(3):1377–1386, 2012.

[4] J.D. Fuller, R. Ramasra, and A. Cha. Fast heuristics for transmission-line switching. *IEEE Transactions on Power Systems*, 27(3):1377–1386, Aug 2012.

[5] M.R. Garey and D.S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1990.

[6] Kory W Hedman, Richard P O'Neill, Emily Bartholomew Fisher, and Shmuel S Oren. Smart flexible just-in-time transmission and flowgate bidding. *IEEE Transactions on Power Systems*, 26(1):93–102, 2010.

[7] Emma S Johnson, Shabbir Ahmed, Santanu S Dey, and Jean-Paul Watson. A k-nearest neighbor heuristic for real-time dc optimal transmission switching. *arXiv preprint arXiv:2003.10565*, 2020.

[8] B. Kocuk, H. Jeon, S.S. Dey, J. Linderoth, J. Luedtke, and X.A. Sun. A cycle-based formulation and valid inequalities for DC power transmission problems with switching. *Operations Research*, 64(4):922–938, 2016.

[9] Burak Kocuk. *Global optimization methods for optimal power flow and transmission switching problems in electric power systems*. PhD thesis, Georgia Institute of Technology, 2016.

[10] Karsten Lehmann, Alban Grastien, and Pascal Van Hentenryck. The complexity of dc-switching problems. *arXiv preprint arXiv:1411.4369*, 2014.

[11] Cong Liu, Jianhui Wang, and James Ostrowski. Heuristic prescreening switchable branches in optimal transmission switching. *IEEE Transactions on Power Systems*, 27(4):2289–2290, 2012.

[12] R. O'Neill, R. Baldick, U. Helman, M. Rothkopf, and J. Stewart. Dispatchable transmission in RTO markets. *IEEE Transactions on Power Systems*, 20(1):171–179, 2005.

[13] Pablo A Ruiz, Justin M Foster, Aleksandr Rudkevich, and Michael C Caramanis. Tractable transmission topology control using sensitivity analysis. *IEEE Transactions on Power Systems*, 27(3):1550–1559, 2012.

[14] Laurence A Wolsey and George L Nemhauser. *Integer and combinatorial optimization*, volume 55. John Wiley & Sons, 1999.

[15] J. Wu and K.W. Cheung. On selection of transmission line candidates for optimal transmission switching in large power networks. In *Power and Energy Society General Meeting (PES), 2013 IEEE*, pages 1–5, July 2013.