

# FDT<sup>★</sup>

Robert Carr<sup>1</sup>, Arash Haddadan<sup>2</sup>, and Cynthia Phillips<sup>3</sup>

<sup>1</sup> University of New Mexico, Albuquerque NM 87131, USA  
bobcarr@unm.

<sup>2</sup> Carnegie Mellon University, Pittsburgh, PA 15213, USA  
ahaddada@cmu.edu

<sup>3</sup> Sandia National Laboratories, Albuquerque, NM 87185, USA  
caphill@sandia.gov

**Abstract.** The abstract should briefly summarize the contents of the paper in 150–250 words.

**Keywords:** First keyword · Second keyword · Another keyword.

## 1 Introduction

Mixed-integer linear programming (MILP), the optimization of a linear objective function subject to linear and integrality constraints, is a classic NP-hard problem [6]. In fact it is NP-hard in general to determine if there is a feasible solution to an MILP problem [6]. However, intelligent branch-and-bound strategies allow commercial and open-source MIP solvers to give exact solutions (or near-optimal with provable bound) to many specific instances of NP-hard combinatorial optimization problems. Modeling combinatorial optimization problems with MILPs also started the very popular area of LP-based approximation algorithms.

Consider the set of point described by a set

$$S(A, G, b) = \{(x, y) \in \mathbb{Z}^n \times \mathbb{R}^p : Ax + Gy \geq b\} \quad (1)$$

and the linear relaxation of set  $S(A, G, b)$ ,

$$P(A, G, b) = \{(x, y) \in \mathbb{R}^{n+p} : Ax + Gy \geq b\}. \quad (2)$$

To make the notation simpler we use  $I = (A, G, b)$ , i.e. an instance of the formulation. Then we use  $S(I)$  and  $P(I)$  to denote  $S(A, G, b)$  and  $P(A, G, b)$ , respectively. Notice that  $\min \{cx : (x, y) \in S(I)\}$  is in fact an MILP problem and  $\min \{cx : (x, y) \in P(I)\}$  is lower bound provided from the LP relaxation. Let  $z_{IP}(I, c)$  and  $z_{LP}(I, c)$  be the optimal solution to these problems, respectively.

Many researchers (see [12,13]) have developed polynomial time algorithms that find solutions for special classes of MILPs whose cost are provably smaller than  $C \cdot z_{LP}(I, c)$  for some  $C$  which can be a constant number or dependent

---

<sup>★</sup> Supported by organization x.

on some parameter of the MILP, e.g.  $O(\log(n))$ . However, for many combinatorial optimization problems there is a limit to such techniques. Let  $g_I = \max_{c \geq 0} \frac{z_{IP}(I, c)}{z_{LP}(I, c)}$ . Parameter  $g_I$  depends on set  $S(I)$  and the linear constraints in (1) and is known as the *integrality gap* of the formulation for instance  $I$ . It is easy to see, that we cannot hope to find solutions for the MILP with objective values better than  $g_I \cdot z_{LP}(I, c)$ . More generally we can define the integrality gap for a class of instances  $\mathcal{I}$ . In this case, the integrality gap for problem  $\mathcal{I}$  is

$$g_{\mathcal{I}} = \max_{c \geq 0, I \in \mathcal{I}} \frac{z_{IP}(I, c)}{z_{LP}(I, c)} \quad (3)$$

For example, in the problem of finding a minimum weight 2-edge-connected multigraph, the gap of the the natural formulation, constraining every cut to be crossed at least twice, is at most  $\frac{3}{2}$  [14] and at least  $\frac{6}{5}$  [1]. Therefore, we cannot hope to obtain an LP-based  $(\frac{6}{5} - \epsilon)$ -approximation algorithm for this problem using this LP relaxation. If we have  $g_I = 1$ ,  $P(I) = \text{conv}(S(I))$  and  $P(I)$  is an *integral* polyhedron. There are many interesting examples of such formulations such as spanning tree polytope and the perfect matching polytope [11]. For such problems we know how to write a vector  $x \in P(I)$  as convex combination of points in  $S(I)$  in polynomial time [8].

**Proposition 1.** *If  $g_I = 1$ , then for  $(x, y) \in P(I)$ , there exists  $\theta \in [0, 1]^k$ , where  $\sum_{i=1}^k \theta_i = 1$  and  $(\tilde{x}^i, \tilde{y}^i) \in S(I)$  for  $i = 1, \dots, k$  such that  $\sum_{i=1}^k \theta_i \tilde{x}^i \leq x$ . Moreover, we can find such a convex combination in polynomial time.*

Now assume  $1 < g < \infty$ . Denote by  $\mathcal{D}(P(I))$  be the set of points  $(x', y')$  such that there exists a point  $(x, y) \in P$  with  $x' \geq x$ .

A analogous result to Proposition 1 is the following theorem due to Carr and Vempala [2], which is a generalization of Goemans proof for blocking polyhedra [7].

**Theorem 1 (Carr, Vempala [2]).** *Let  $(x, y) \in P(I)$ , there exists  $\theta \in [0, 1]^k$ , where  $\sum_{i=1}^k \theta_i = 1$  and  $(\tilde{x}^i, \tilde{y}^i) \in \mathcal{D}(S(I))$  for  $i = 1, \dots, k$  such that  $\sum_{i=1}^k \theta_i \tilde{x}^i \leq Cx$  if and only if  $g_I \leq C$ .*

In contrast to Proposition 1 which implies exact algorithms for problems with a gap of 1, Theorem 1 does not imply an approximation algorithm, since it does not suggest how to find such a convex combination in polynomial time. This points to an interesting open question.

*Question 1.* Assume reasonable complexity assumptions (such as UGC or  $P \neq NP$ ). Given instance  $I$  with  $1 < g_I < \infty$  and  $(x, y) \in P(I)$ , can we find  $\theta \in [0, 1]^k$ , where  $\sum_{i=1}^k \theta_i = 1$  and  $(\tilde{x}^i, \tilde{y}^i) \in \mathcal{D}(\text{conv}(S(I)))$  for  $i = 1, \dots, k$  such that  $\sum_{i=1}^k \theta_i \tilde{x}^i \leq g_I x$  in polynomial time?

This seems to be a very hard question. A more specific question is of more interest.

*Question 2.* Assume reasonable complexity assumptions, a specific problem  $\mathcal{I}$  with  $1 < g_{\mathcal{I}} < \infty$ , and  $(x, y) \in P(I)$  for some  $I \in \mathcal{I}$ , can we find  $\theta \in [0, 1]^k$ , where  $\sum_{i=1}^k \theta_i = 1$  and  $(\tilde{x}^i, \tilde{y}^i) \in S(I)$  for  $i = 1, \dots, k$  such that  $\sum_{i=1}^k \theta_i \tilde{x}^i \leq g_{\mathcal{I}} x$  in polynomial time?

Although Question 2 is wide open, for some problems  $\mathcal{I}$  there are polynomial time algorithms closing the gap. For example, for generalized Steiner forest problem [9] gave an LP-based 2-approximation algorithm. The gap for this problem is also lower bounded by 2. Same holds for the set covering problem [10]. In fact, for set cover the approximation algorithm achieving the same factor as the integrality gap lower bound, is a *randomized rounding* algorithm. Raghavan and Thompson [10] showed that this technique achieves provably good approximation for many combinatorial optimization problems.

One key step, in answering Question 1 and 2 is being able to find a solution in  $S(I)$ . In fact, for general  $I$  it is NP-hard to even decide if  $S(I)$  is empty or not. There are a number of heuristics for this purpose, such as the feasibility pump heuristic [4,5]. These heuristics are often very effective and fast in practice, however, they can sometimes fail in finding a feasible solution. In addition these algorithms do not provide any bounds on the quality of the solution they find.

### 1.1 Our results

In this paper we study Theorem 1 and the natural questions that follow from it (Questions 1 and 2).

We give a general approximation framework for solving  $\{0, 1\}$ -MILPs. Consider the set of point described by a set

$$S(I) = \{(x, y) \in \mathbb{R}^{n \times p} : Ax + Gy \geq b, x \in \{0, 1\}\}, \quad (4)$$

and a linear relaxation of set  $S$

$$P(I) = \{(x, y) \in \mathbb{R}^{n \times p} : Ax + Gy \geq b, 0 \leq x \leq 1\}. \quad (5)$$

For a vector  $x \in \mathbb{R}^n$  such that  $(x, y) \in P(I)$  for some  $y \in \mathbb{R}^p$ , let  $\text{supp}(x) = \{i \in \{1, \dots, n\} : x_i \neq 0\}$ . We prove the following algorithm.

**Theorem 2.** Assume  $1 \leq g_I < \infty$ . The FDT algorithm, given  $(x^*, y^*) \in P(I)$ , produces in polynomial time  $\lambda \in [0, 1]^k$  and  $(z^1, w^1), \dots, (z^k, w^k) \in S(I)$  such that  $k \leq |\text{supp}(x^*)|$ ,  $\sum_{i=1}^k \lambda_i z^i \leq Cx^*$ , and  $\sum_{i=1}^k \lambda_i = 1$ . Moreover,  $C \leq g_I^{|\text{supp}(x^*)|}$ .

Notice, that the FDT algorithm needs to find feasible solutions to any MILP with finite gap. This can be of independent interest.

**Theorem 3.** Assume  $1 \leq g_I < \infty$ . The DomToIP algorithm finds  $(\hat{x}, \hat{y}) \in S(I)$  in polynomial time.

In attempt to extend our results into  $\{0, 1, 2\}$ -MILPs we give an FDT algorithm for the 2-edge-connected multigraph problem. In this problem we want to find the minimum weight 2-edge-connected subgraph (which can multiple copies of each edge) in a graph  $G = (V, E)$  with respect to weights  $c \in \mathbb{R}_{\geq 0}^E$ . For this problem the natural linear programming relaxation is

$$2EC(G) = \{x \in \mathbb{R}_{\geq 0}^E : x(\delta(S)) \geq 2 \text{ for } \emptyset \subset S \subset V\}. \quad (6)$$

We prove the following theorem.

**Theorem 4.** *Let  $G = (V, E)$  and  $x$  be an extreme point of  $2EC(G)$ . The FDT algorithm for 2EC produces  $\lambda \in [0, 1]^k$  and 2-edge-connected multigraphs  $F_1, \dots, F_k$  such that  $k \leq O(|V|)$ ,  $\sum_{i=1}^k \lambda_i \chi^{F_i} \leq Cx^*$ , and  $\sum_{i=1}^k \lambda_i = 1$ . Moreover,  $C \leq g_{2EC}^{O(|V|)}$ , where  $g_{2EC}$  is the integrality gap of the 2-edge-connected multigraph problem with respect to formulation in (6)*

Although the bound guaranteed in both Theorems 2 and 4 are very large for large problems, we show that in practice, the algorithm works very well providing approximation ratios that are far better than the theoretical bound in the theorem statements. We examine FDT for binary MILPs for problems such as the tree augmentation problem (TAP) and we apply FDT for 2EC on some interesting and “hard to decompose” points in the linear relaxation.

Finally we give a stronger characterization of integrality gap than that in Theorem 1 for bounded covering problems. For this purpose assume  $P = \{x \in \mathbb{R}_{\geq 0}^n : Ax \geq b\mathbf{1}, x \leq b\mathbf{1}\}$ ,  $S = P \cap \mathbb{Z}^n$ , and  $g = \max_{c \geq 0} \frac{\min_{x \in S} cx}{\min_{x \in P} cx}$ . Examples of such problems include the 2-edge-connected multigraph problem, the tree augmentation problem, and many others.

**Theorem 5.** *Let  $x \in P$ , there exists  $\theta \in [0, 1]^k$ , where  $\sum_{i=1}^k \theta_i = 1$  and  $\tilde{x}^i \in S$  for  $i = 1, \dots, k$  such that*

- for  $\ell \in \{1, \dots, n\}$ , if  $x_\ell = 0$ , then  $\tilde{x}_\ell^i = 0$  for  $i = 1, \dots, k$ , i.e.  $\tilde{x}^i$  is in the support of  $x$ ,
- we have  $Cb = C \cdot A_j x \geq A_j(\sum_{i=1}^k \theta_i \tilde{x}^i)$ , for  $j$  such that  $A_j x = b$ ,

*if and only if  $C \geq g$ .*

This means in order to prove an upper bound on the integrality gap of a covering problem, we need to show there is a convex combination of integer feasible points that is “cheap” on all tight cuts. Notice that Theorem 1 requires the certificate convex combination to be “cheap” on every single variable.

## 1.2 Notations

For vectors  $x, y \in \mathbb{R}_n$  we say  $x$  dominates  $y$  if  $x_i \geq y_i$  for  $i = 1, \dots, n$ . For  $m \times n$  matrix  $A$ , let  $A_j$  be the  $j$ -th row of  $A$  and  $A^j$  be the  $j$ -th column of  $A$ . Let  $\mathbf{1}$  be the vector of all ones of a suitable dimension. For a set  $S$  of vectors in  $\mathbb{R}_n$ ,  $\text{conv}(S)$  is the convex hull of all the points in  $S$ .

## 2 FDT on binary MIPs

In this section we present the fractional decomposition tree (FDT) algorithm and prove its correctness.

Consider a formulation instance  $I = (A, G, b)$ , set

$$S(I) = \{(x, y) \in \mathbb{R}^{n+p} : Ax + Gy \geq b, x \in \{0, 1\}\}, \quad (7)$$

and its linear relaxation

$$P(I) = \{(x, y) \in \mathbb{R}^{n+p} : Ax + Gy \geq b, 0 \leq x \leq 1\}. \quad (8)$$

Also assume we are given a point  $(x^*, y^*) \in P(I)$ . For instance,  $(x^*, y^*)$  can be the optimal solution of minimizing a cost function  $cx$  over set  $P$ , which provides a lower bound on  $\min_{(x,y) \in S(I)} cx$ . In this section, we prove Theorem 2. One key ingredient in proving Theorem 2 is Theorem 3 that we will also prove in this section.

Although the theoretical worst-case upper bound on  $C$  in the statement of Theorem 2 can be very large, we will show that in practice  $C$  can be really small and hence FDT can provide good LP-based approximation algorithms in many cases. We also remark that if  $g_I = 1$ , then the algorithm will give an exact decomposition of any feasible solution. In the remainder of this section we explain the FDT algorithm and prove Theorem 2. For simplicity in the notation we denote  $P(I)$ ,  $S(I)$ , and  $g_I$  with  $P$ ,  $S$ , and  $g$ , respectively. Let  $t = |\text{supp}(x^*)|$ . We assume without loss of generality that  $\text{supp}(x^*) = \{1, \dots, t\}$ .

We begin with the following lemmas that show how the FDT algorithm branches on a variable.

**Lemma 1.** *Given  $(x', y') \in \mathcal{D}(P)$  and  $\ell \in \{1, \dots, n\}$ , we can find in polynomial time vectors  $(\hat{x}^0, \hat{y}^0)$ ,  $(\hat{x}^1, \hat{y}^1)$  and scalars  $\gamma_0, \gamma_1 \in [0, 1]$  such that: (i)  $\gamma_0 + \gamma_1 \geq \frac{1}{g}$ , (ii)  $(\hat{x}^0, \hat{y}^0)$  and  $(\hat{x}^1, \hat{y}^1)$  are in  $P$ , (iii)  $\hat{x}_\ell^0 = 0$  and  $\hat{x}_\ell^1 = 1$ , (iv)  $\gamma_0 \hat{x}^0 + \gamma_1 \hat{x}^1 \leq x'$ .*

*Proof.* Consider the following linear program which we denote by  $\text{Branching}(\ell, x', y')$ . The variables of  $\text{Branching}(\ell, x', y')$  are  $\gamma_0, \gamma_1$  and  $(x^0, y^0)$  and  $(x^1, y^1)$ .

$$\text{Branching}(\ell, x', y') \quad \max \quad \lambda_0 + \lambda_1 \quad (9)$$

$$\text{s.t.} \quad Ax^j + Gy^j \geq b\lambda_j \quad \text{for } j = 0, 1 \quad (10)$$

$$0 \leq x^j \leq \lambda_j \quad \text{for } j = 0, 1 \quad (11)$$

$$x_\ell^0 = 0, x_\ell^1 = \lambda_1 \quad (12)$$

$$x^0 + x^1 \leq x' \quad (13)$$

$$\lambda_0, \lambda_1 \geq 0 \quad (14)$$

Let  $(x^0, y^0)$ ,  $(x^1, y^1)$ , and  $\gamma_0, \gamma_1$  be an optimal solution to the LP above. Let  $(\hat{x}^0, \hat{y}^0) = (\frac{x^0}{\gamma_0}, \frac{y^0}{\gamma_0})$ ,  $(\hat{x}^1, \hat{y}^1) = (\frac{x^1}{\gamma_1}, \frac{y^1}{\gamma_1})$ . Observe that (ii), (iii), (iv) are satisfied with this choice. In order to show that (i) is also satisfied we prove the following claim.

*Claim.* We have  $\gamma_0 + \gamma_1 \geq \frac{1}{g}$ .

*Proof.* We show that there is a feasible solution that achieves the objective value of  $\frac{1}{g}$ . By Theorem 1 there exists  $\theta \in [0, 1]^k$ , with  $\sum_{i=1}^k \theta_i = 1$  and  $(\tilde{x}^i, \tilde{y}^i) \in S$  for  $i = 1, \dots, k$  such that  $\sum_{i=1}^k \theta_i \tilde{x}^i \leq gx'$ .

$$x' \geq \sum_{i=1}^k \frac{\theta_i}{g} \tilde{x}^i = \sum_{i \in [k]: \tilde{x}_\ell^i = 0} \frac{\theta_i}{g} \tilde{x}^i + \sum_{i \in [k]: \tilde{x}_\ell^i = 1} \frac{\theta_i}{g} \tilde{x}^i \quad (15)$$

For  $j = 0, 1$ , let  $(x^j, y^j) = \sum_{i \in [k]: \tilde{x}_\ell^i = j} \frac{\theta_i}{g} (\tilde{x}^i, \tilde{y}^i)$ . Also let  $\lambda_0 = \sum_{i \in [k]: \tilde{x}_\ell^i = 0} \frac{\theta_i}{g}$  and  $\lambda_1 = \sum_{i \in [k]: \tilde{x}_\ell^i = 1} \frac{\theta_i}{g}$ . Note that  $\lambda_0 + \lambda_1 = \frac{1}{g}$ . Constraint (13) is satisfied by Inequality (15). Also, for  $j = 0, 1$  we have

$$Ax^j + Gy^j = \sum_{i \in [k], \tilde{x}_\ell^i = j} \frac{\theta_i}{g} (A\tilde{x}^i + G\tilde{y}^i) \geq b \sum_{i \in [k], \tilde{x}_\ell^i = j} \frac{\theta_i}{g} = b\lambda_j. \quad (16)$$

Hence, Constraints (10) holds. Constraint (12) also holds since  $x_\ell^0$  is obviously 0 and  $x_\ell^1 = \sum_{i \in [k]: \tilde{x}_\ell^i = 1} \frac{\theta_i}{g} = \lambda_1$ . The rest of the constraints trivially hold.  $\square$

This concludes the proof.  $\square$

We now show if  $x'$  in the statement of Lemma 1 is partially integral, we can find solutions with more integral components.

**Lemma 2.** *Given  $(x', y') \in \mathcal{D}(P)$ , such that  $x'_1, \dots, x'_{\ell-1} \in \{0, 1\}$  for some  $\ell \geq 1$ , we can find in polynomial time vectors  $(\hat{x}^0, \hat{y}^0), (\hat{x}^1, \hat{y}^1)$  and scalars  $\gamma_0, \gamma_1 \in [0, 1]$  such that: (i)  $\frac{1}{g} \leq \gamma_0 + \gamma_1 \leq 1$ , (ii)  $(\hat{x}^0, \hat{y}^0)$  and  $(\hat{x}^1, \hat{y}^1)$  are in  $\mathcal{D}(P)$ , (iii)  $\hat{x}_\ell^0 = 0$  and  $\hat{x}_\ell^1 = 1$ , (iv)  $\gamma_0 \hat{x}^0 + \gamma_1 \hat{x}^1 \leq x'$ , (v)  $\hat{x}_j^i \in \{0, 1\}$  for  $i = 0, 1$  and  $j = 1, \dots, \ell - 1$ .*

*Proof.* By Lemma 1 we can find  $(\bar{x}^0, \bar{y}^0), (\bar{x}^1, \bar{y}^1), \gamma_0$  and  $\gamma_1$  that satisfy (i), (ii), (iii), and (iv). We define  $\hat{x}^0$  and  $\hat{x}^1$  as follows. For  $i = 0, 1$ , for  $j = 1, \dots, \ell - 1$ , let  $\hat{x}_j^i = \lceil \bar{x}_j^i \rceil$ , for  $j = \ell, \dots, t$  let  $\hat{x}_j^i = \bar{x}_j^i$ . We now show that  $(\hat{x}^0, \bar{y}^0), (\hat{x}^1, \bar{y}^1), \gamma_0$ , and  $\gamma_1$  satisfy all the conditions. Note that conditions (i), (ii), (iii), and (v) are trivially satisfied. Thus we only need to show (iv) holds. We need to show that  $\gamma_0 \hat{x}_j^0 + \gamma_1 \hat{x}_j^1 \leq x'_j$ . If  $j = \ell, \dots, t$ , then this clearly holds. Hence, assume  $j \leq \ell - 1$ . By the property of  $x'$  we have  $x'_j \in \{0, 1\}$ . If  $x'_j = 0$ , then by Constraint (13) we have  $\bar{x}_j^0 = \bar{x}_j^1 = 0$ . Therefore,  $\hat{x}_j^i = 0$  for  $i = 0, 1$ , so (iv) holds. Otherwise if  $x'_j = 1$ , then we have  $\gamma_0 \hat{x}_j^0 + \gamma_1 \hat{x}_j^1 \leq \gamma_0 + \gamma_1 \leq 1 \leq x'_j$ . Therefore (v) holds.  $\square$

The FDT algorithm grows a tree similar to the classic branch-and-bound search tree for integer programs. Each node represents a partially integral vector  $(\bar{x}, \bar{y})$  in  $\mathcal{D}(P)$  together with a multiplier  $\bar{\lambda}$ . The solutions contained in the nodes of the tree become progressively more integral at each level. Leaves of the FDT tree contain solutions with integer values for all the  $x$  variables that dominate a point in  $P$ . We will later see how we can turn these into points in  $S$ .

Let us define the FDT algorithm more formally. The algorithm maintains nodes  $L_i$  in iteration  $i$  of the algorithm. The nodes in  $L_i$  correspond to the nodes in level  $L_i$  of the FDT tree. The points in the leaves of the FDT tree,  $L_t$ , are points in  $\mathcal{D}(P)$  and are integral for all integer variables.

**Lemma 3.** *There is a polynomial time algorithm that produces sets  $L_0, \dots, L_t$  of pairs of  $(x, y) \in \mathcal{D}(P)$  together with multipliers  $\lambda$  with the following properties for  $i = 0, \dots, t$ : (a) If  $(x, y) \in L_i$ , then  $x_j \in \{0, 1\}$  for  $j = 1, \dots, i$ , i.e. the first  $i$  coordinates of a solution in level  $i$  are integral, (b)  $\sum_{[(x, y), \lambda] \in L_i} \lambda \geq \frac{1}{g^i}$ , (c)  $\sum_{[(x, y), \lambda] \in L_i} \lambda x \leq x^*$ , (d)  $|L_i| \leq t$ .*

*Proof.* We prove this lemma using induction but one can clearly see how to turn this proof into a polynomial time algorithm. Let  $L_0$  be the set that contains a single node (root of the FDT tree) with  $(x^*, y^*)$  and multiplier 1. It is easy to check all the requirements in the lemma are satisfied for this choice.

Suppose by induction that we have constructed sets  $L_0, \dots, L_i$ . Let the solutions in  $L_i$  be  $(x^j, y^j)$  for  $j = 1, \dots, k$  and  $\lambda_j$  be their multipliers, respectively. For each  $j = 1, \dots, k$  by Lemma 2 (setting  $(x', y') = (x^j, y^j)$  and  $\ell = i + 1$ ) we can find  $(x^{j0}, y^{j0}), (x^{j1}, y^{j1})$  and  $\lambda_j^0, \lambda_j^1$  with the properties (i) to (v) in Lemma 2. Define  $L'$  to be the set of nodes with solutions  $(x^{j0}, y^{j0}), (x^{j1}, y^{j1})$  and multipliers  $\lambda_j \lambda_j^0, \lambda_j \lambda_j^1$ , respectively, for  $j = 1, \dots, k$ . It is easy to check that set  $L'$  is a suitable candidate for  $L_{i+1}$ , i.e. set  $L'$  satisfies (a), (b) and (c). However we can only ensure that  $|L'| \leq 2k \leq 2t$ , and might have  $|L'| > t$ . We call the following linear program  $\text{Pruning}(L')$ . Let  $L' = \{[(x^1, y^1), \gamma_1], \dots, [(x^{2k}, y^{2k}), \gamma_{2k}]\}$ . The variables of  $\text{Pruning}(L')$  is a scalar variable  $\theta_j$  for each node  $j$  in  $L'$ .

$$\text{Pruning}(L') \quad \max \quad \sum_{j=1}^{2k} \theta_j \quad (17)$$

$$\text{s.t.} \quad \sum_{j=1}^{2k} \theta_j x_i^j \leq x_i^* \quad \text{for } i = 1, \dots, t \quad (18)$$

$$\theta \geq 0 \quad (19)$$

Notice that  $\theta = \gamma$  is in fact a feasible solution to  $\text{Pruning}(L')$ . Let  $\theta^*$  be the optimal vertex point solution to this LP. Since the problem is in  $\mathbb{R}^{2k}$ ,  $\theta^*$  has to satisfy  $2k$  linearly independent constraints at equality. However, there are only  $t$  constraints of type (18). Therefore, there are at most  $t$  coordinates of  $\theta_j^*$  that are non-zero. We claim that  $L_{i+1}$  which consists of  $(x^j, y^j)$  for  $j = 1, \dots, 2k$  and their corresponding multipliers  $\theta_j^*$  satisfy the properties in the statement of the lemma. Notice that, we can discard the nodes in  $L_{i+1}$  that have  $\theta_j^* = 0$ , so  $|L_{i+1}| \leq t$ . Also, since  $\theta^*$  is optimal and  $\gamma$  is feasible for  $\text{Pruning}(L')$ , we have  $\sum_{j=1}^k \theta_j^* \geq \sum_{j=1}^{2k} \gamma_j \geq \frac{1}{g^{i+1}}$ .  $\square$

The leaves of the FDT tree,  $L_t$ , have the property that every solution  $(x, y)$  in  $L_t$  has  $x \in \{0, 1\}^n$  and  $(x, y) \in \mathcal{D}(P)$ . Our goal to now replace these leaves

solutions in  $\mathcal{D}(P)$  with solutions in  $S$ . To this end, we need to reduce some variables with value of 1 into variables of value 0 and obtain feasibility in the meantime. We introduce a procedure that given a point  $(x, y) \in \mathcal{D}(P)$  with  $x \in \{0, 1\}^n$  outputs a point  $z \leq x$  with  $z \in \{0, 1\}^n$  such that  $(z, w)$  is in  $S$  for some vector  $w \in \mathbb{R}^p$ .

The algorithm “fixes” the variables iteratively, from  $x_1$  to  $x_t$ . Suppose we run the algorithm for  $\ell \in \{0, \dots, t-1\}$  iterations and we have  $(x^{(\ell)}, y^{(\ell)}) \in \mathcal{D}(P)$  such that  $x_i^{(\ell)} \in \{0, 1\}$  for  $i = 1, \dots, \ell$ . Now consider the following linear program. The variables of this LP are the  $z \in \mathbb{R}^n$  variables and  $w \in \mathbb{R}^p$ .

$$\text{DominantToFeasible}(x^{(\ell)}) \quad \min \quad z_{\ell+1} \quad (20)$$

$$\text{s.t.} \quad Az + Gw \geq b \quad (21)$$

$$z_j = x_j^{(\ell)} \quad j = 1, \dots, \ell \quad (22)$$

$$z_j \leq x_j^{(\ell)} \quad j = \ell + 1, \dots, n \quad (23)$$

$$z \geq 0 \quad (24)$$

If the optimal value to  $\text{DominantToFeasible}(x^{(\ell)})$  is 0, then let  $x_{\ell+1}^{(\ell+1)} = 0$ . Otherwise if the optimal value is strictly positive let  $x_{\ell+1}^{(\ell+1)} = 1$ . Let  $x_j^{(\ell+1)} = x_j^{(\ell)}$  for  $j \in \{1, \dots, n\} \setminus \{\ell + 1\}$ . This procedure is the  $\text{DomToIP}$  algorithm. The following lemma proves Theorem 3.

**Lemma 4.** *Given  $(x, y) \in \mathcal{D}(P)$ ,  $\text{DomToIP}$  algorithm correctly finds  $(x^{(t)}, y^{(t)}) \in S$  in polynomial time.*

The proof is in Appendix A.

### 3 FDT for 2EC

In Section 2 our focus was on binary MIPs. In this section, in an attempt to extend FDT to 0,1,2 problems we introduce an FDT algorithm for a 2-edge-connected multigraph problem. Given a graph  $G = (V, E)$  a multi-subset of edges  $F$  of  $G$  is a 2-edge-connected multigraph of  $G$  if for each set  $\emptyset \subset U \subset V$ , the number of edge in  $F$  that have one endpoint in  $U$  and one not in  $U$  is at least 2. In the 2EC problem, we are given non-negative costs on the edge of  $G$  and the goal is to find the minimum cost 2-edge-connected multigraph of  $G$ . Notice that, no optimal solution ever takes 3 copies of an edge in 2EC, hence we assume that we can take an edge at most 2 times. The natural linear programming relaxation is  $2\text{EC}(G) = \{x \in [0, 2]^E : x(\delta(U)) \geq 2 \text{ for } \emptyset \subset U \subset V\}$ . Notice that  $\mathcal{D}(2\text{EC}(G)) \cap [0, 2]^E = 2\text{EC}(G)$ , since 2EC is a covering problem. We want to prove the following theorem.

**Theorem 4.** *Let  $G = (V, E)$  and  $x$  be an extreme point of  $2\text{EC}(G)$ . The FDT algorithm for 2EC produces  $\lambda \in [0, 1]^k$  and 2-edge-connected multigraphs  $F_1, \dots, F_k$  such that  $k \leq O(|V|)$ ,  $\sum_{i=1}^k \lambda_i \chi^{F_i} \leq Cx^*$ , and  $\sum_{i=1}^k \lambda_i = 1$ . Moreover,  $C \leq g_{2\text{EC}}^{O(|V|)}$ , where  $g_{2\text{EC}}$  is the integrality gap of the 2-edge-connected multigraph problem with respect to formulation in (6)*



We do not know the exact value for  $g_{2EC}$ , but we know  $\frac{6}{5} \leq g_{2EC} \leq \frac{3}{2}$  [1,14]. Also, we need to remark that polyhedral version of Christofides' algorithm provides a  $\frac{3}{2}$ -approximation for 2EC, i.e. we already have an algorithm with  $C \leq \frac{3}{2}$ . However, we show later in Section ?? that in practice the constant  $C$  for the FDT algorithm for 2EC is much better than  $\frac{3}{2}$ .

The FDT algorithm for 2EC is very similar to the one for binary MIPs, but there are some differences as well. First, we need a branching lemma. Observe that the following branching lemma is essentially a translation of Lemma 1 for 0,1,2 problems expect for one additional clause.

**Lemma 5.** *Given  $x \in 2EC(G)$ , and  $e \in E$  we can find in polynomial time vectors  $x^0, x^1$  and  $x^2$  and scalars  $\gamma_0, \gamma_1$ , and  $\gamma_2$  such that: (i)  $\gamma_0 + \gamma_1 + \gamma_2 \geq \frac{1}{g_{2EC}}$ , (ii)  $x^0, x^1$ , and  $x^2$  are in  $2EC(G)$ , (iii)  $x_e^0 = 0$ ,  $x_e^1 = 1$ , and  $x_e^2 = 2$ , (iv)  $\gamma_0 x^0 + \gamma_1 x^1 + \gamma_2 x^2 \leq x$ , (v) for  $f \in E$  with  $x_f \geq 1$ , we have  $x_f^j \geq 1$  for  $j = 0, 1, 2$ .*

*Proof.* Consider the following LP with variables  $\lambda_j$  and  $x^j$  for  $j = 0, 1, 2$ .

$$\max \sum_{j=0,1,2} \lambda_j \quad (25)$$

$$\text{s.t. } x^j(\delta(U)) \geq 2\lambda_j \quad \text{for } \emptyset \subset U \subset V, \text{ and } j = 0, 1, 2 \quad (26)$$

$$0 \leq x^j \leq 2\lambda_j \quad \text{for } j = 0, 1, 2 \quad (27)$$

$$x_e^j = j \quad \text{for } j = 0, 1, 2 \quad (28)$$

$$x_f^j \geq j \quad \text{for } f \in E \text{ where } x_f \geq 1, \text{ and } j = 0, 1, 2 \quad (29)$$

$$x^0 + x^1 + x^2 \leq x \quad (30)$$

$$\lambda_0, \lambda_1, \lambda_2 \geq 0 \quad (31)$$

Let  $x^j, \gamma_j$  for  $j = 0, 1, 2$  be an optimal solution solution to the LP above. Let  $\hat{x}^j = \frac{x^j}{\gamma_j}$  for  $j = 0, 1, 2$  where  $\gamma_j > 0$ . If  $\gamma_j = 0$ , let  $\hat{x}^j = 0$ . Observe that (ii), (iii), (iv), and (v) are satisfied with this choice. We can also show that  $\gamma_0 + \gamma_1 + \gamma_2 \geq \frac{1}{g_{2EC}}$ , which means that (i) is also satisfied. The proof is similar to the proof of the claim in Lemma 1, but we need to replace each  $f \in E$  with  $x_f \geq 0$  with a suitably long path to ensure that Constraint (29) is also satisfied. We skip the details.  $\square$

In contrast to FDT for binary MIPs where we round up the fractional variables that are already branched on at each level, in FDT for 2EC we keep all coordinates as they are and perform a rounding procedure at the end. Formally, let  $L_i$  for  $i = 1, \dots, |\text{supp}(x^*)|$  be collections of pairs of feasible points in  $2EC(G)$  together with their multipliers. Let  $t = |\text{supp}(x^*)|$  and assume without loss of generality that  $\text{supp}(x^*) = \{e_1, \dots, e_t\}$ .

**Lemma 6.** *The FDT algorithm for 2EC in polynomial time produces sets  $L_0, \dots, L_t$  of pairs  $x \in 2EC(G)$  together with multipliers  $\lambda$  with the following properties. (a) If  $x \in L_i$ , then  $x_{e_j} = 0$  or  $x_{e_j} \geq 1$  for  $j = 1, \dots, i$ , (b)  $\sum_{(x,\lambda) \in L_i} \lambda \geq \frac{1}{g_{2EC}^i}$ , (c)  $\sum_{(x,\lambda) \in L_i} \lambda x \leq x^*$ , (d)  $|L_i| \leq t$ .*

The proof is similar to Lemma 3, but we need to use property (i) in Lemma 5 to prove that (a) also holds.

Consider the solutions  $x$  in  $L_t$ . For each variable  $e$  we have  $x_e = 0$  or  $x_e \geq 1$ .

**Lemma 7.** *Let  $x$  be a solution in  $L_t$ . Then  $\lfloor x \rfloor \in 2\text{EC}(G)$ .*

*Proof.* Suppose not. Then there is a set of vertices  $\emptyset \subset U \subset V$  such that  $\sum_{e \in \delta(U)} \lfloor x_e \rfloor < 2$ . Since  $x \in 2\text{EC}(G)$  we have  $\sum_{e \in \delta(U)} x_e \geq 2$ . Therefore, there is an edge  $f \in \delta(U)$  such that  $x_f$  is fractional. By property (a) in Lemma 6, we have  $1 < x_f < 2$ . Therefore, there is another edge  $h$  in  $\delta(U)$  such that  $x_h > 0$ , which implies that  $x_h \geq 1$ . But in this case  $\sum_{e \in \delta(U)} \lfloor x_e \rfloor \geq \lfloor x_f \rfloor + \lfloor x_h \rfloor \geq 2$ . This is a contradiction.  $\square$

The FDT algorithm for 2EC iteratively applies Lemmas 5 and 6 to variables  $x_1, \dots, x_t$  to obtain leaf point solutions  $L_t$ . Then, we just need to apply Lemma 7 to obtain the 2-edge-connected multigraphs from every solution in  $L_t$ . Notice that since  $x$  is an extreme point we have  $t \leq 2|V| - 1$  [3]. By Lemma 6 we have

$$\sum_{(x, \lambda) \in L_t} \frac{\lambda}{\sum_{(x, \lambda) \in L_t} \lambda} \lfloor x \rfloor \leq \frac{1}{\sum_{(x, \lambda) \in L_t} \lambda} \sum_{(x, \lambda) \in L_t} \lambda x \leq g_{2\text{EC}}^t x^*.$$

## 4 Stronger version of Theorem 1 for covering problems

Theorem 1 characterizes the integrality gap as the smallest number such that for any point  $x \in \mathcal{D}(P)$ ,  $gx$  dominates a convex combination of points in  $S$ , i.e. there exists convex combination  $\sum_{i=1}^k \lambda_i x^i \leq gx$ . In a covering problem, we assume that matrix  $A$  in the description of  $P$  is a non-negative matrix, hence if  $y \geq x$  and  $x \in P$  we have  $y \in P$ . We also assume the right hand side vector in (2) is of the form  $b\mathbf{1}$ , i.e. it is a uniform vector. Finally we assume we have bound constraints  $x \leq b\mathbf{1}$ . This class of problems include a broad class of problems. Notice that for a covering problem  $\mathcal{D}(P) = P$ . We show that for these problems, we can make Theorem 1 stronger in the sense of Theorem 5.

In other words, the integrality gap is the smallest number such that for any  $x \in P$ , there exists a convex combination  $y = \sum_{i=1}^k \lambda_i x^i$ , such that  $A_j y \leq g(A_j x)$  for all  $j$  such that  $A_j x = b$ .

Notice that the proof of this theorem, same as Theorem 1, does not imply a polynomial time algorithm.

Observe that if  $C \geq g$ , by Theorem 1, there is a convex combination of points in  $S$  that is dominated by  $Cx$ , i.e. there exists  $\theta \in [0, 1]^k$ , where  $\sum_{i=1}^k \theta_i = 1$ , and  $x^i \in S$ , such that  $\sum_{i=1}^k \theta_i x^i \leq gx$ . The same convex combination also trivially satisfies both requirements in Theorem 5. The more interesting direction is the converse. We begin with the following theorem. Suppose  $v^1, \dots, v^r$  are the extreme points of  $P$ . Assume for each  $i = 1, \dots, r$ , there exists a convex combination of points in  $S$ , namely  $\sum_{\ell=1}^{k_i} \lambda_\ell^i w_\ell^i$  such that  $A_j \sum_{\ell=1}^{k_i} \lambda_\ell^i w_\ell^i \leq C A_j v^i$  for  $j$  such that  $A_j v^i = b$ . Let  $z^i = \sum_{\ell=1}^{k_i} \lambda_\ell^i w_\ell^i$ .

**Lemma 8.** *For any  $i \in \{1, \dots, r\}$ , there exists  $\epsilon_i > 0$  and  $x^i \in P$  where  $x^i \leq \frac{Cv^i - \epsilon_i z^i}{C(1 - \epsilon_i)}$ .*

*Proof Sketch.* □

**Lemma 9.** *There exists a convex combination of point in  $S$  that is dominated by  $Cv^i$  for  $i = 1, \dots, r$ .*

*Proof.* We prove something slightly stronger. We show that for  $i \in \{1, \dots, r\}$ , and for  $j = 1, \dots, i$ , the vector  $Cv^j$  dominates a convex combination of  $z^1, \dots, z^i$  and  $Cv^{i+1}, \dots, Cv^r$ . Note that this is enough to prove our statement since for  $i = r$ , the statement implies that  $Cv^j$  dominates a convex combination of  $z^1, \dots, z^r$  for  $j = 1, \dots, r$ .

We can prove this statement by induction on  $i$  and applying Lemma 8 to prove the induction step. □

## References

1. Robert Carr and R. Ravi. A new bound for the 2-edge connected subgraph problem. In *Proceedings of the 6th International IPCO Conference on Integer Programming and Combinatorial Optimization*, pages 112–125, London, UK, UK, 1998. Springer-Verlag.
2. Robert Carr and Santosh Vempala. On the held-karp relaxation for the asymmetric and symmetric traveling salesman problems. *Mathematical Programming*, 100(3):569–587, Jul 2004.
3. Gérard Cornuéjols, Jean Fonlupt, and Denis Naddef. The traveling salesman problem on a graph and some related integer polyhedra. *Mathematical Programming*, 33(1):1–27, Sep 1985.
4. Matteo Fischetti, Fred Glover, and Andrea Lodi. The feasibility pump. *Mathematical Programming*, 104(1):91–104, Sep 2005.
5. Matteo Fischetti and Domenico Salvagnin. Feasibility pump 2.0. *Mathematical Programming Computation*, 1(2):201–222, Oct 2009.
6. M.R. Garey and D.S. Johnson. *Computers and Intractability – A Guide to the Theory of NP-Completeness*. Freeman, San Francisco, CA, 1979.
7. Michel X. Goemans. Worst-case comparison of valid inequalities for the TSP. *Math. Program.*, 69(2):335–349, August 1995.
8. Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2. Second corrected edition edition, 1993.
9. Kamal Jain. A factor 2 approximation algorithm for the generalized steiner network problem. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, FOCS '98, pages 448–, Washington, DC, USA, 1998. IEEE Computer Society.
10. Prabhakar Raghavan and Clark D. Tompson. Randomized rounding: A technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7(4):365–374, Dec 1987.
11. A. Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency*. Springer, 2003.
12. Vijay V. Vazirani. *Approximation Algorithms*. Springer-Verlag, Berlin, Heidelberg, 2001.

13. David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, New York, NY, USA, 1st edition, 2011.
14. Laurence A. Wolsey. *Heuristic analysis, linear programming and branch and bound*, pages 121–134. Springer Berlin Heidelberg, Berlin, Heidelberg, 1980.

## A Proof of Lemma 4

First, we need to show that in any iteration  $\ell = 0, \dots, t-1$  of Algorithm ?? the  $\text{DominantToFeasible}(x^\ell)$  is feasible. We show something stronger. For  $\ell = 0, \dots, t-1$  let

$$\begin{aligned} \text{LP}^{(\ell)} &= \{(z, w) \in P : z \leq x^{(\ell)} \text{ and } z_j = x_j^{(\ell)} \text{ for } j = 1, \dots, \ell\}, \text{ and} \\ \text{IP}^{(\ell)} &= \{(z, w) \in \text{LP}^{(\ell)} : z \in \{0, 1\}^n\}. \end{aligned}$$

Notice that if  $\text{LP}^{(\ell)}$  is a non-empty set then  $\text{DominantToFeasible}(x^{(\ell)})$  is feasible. We show by induction on  $\ell$  that  $\text{LP}^{(\ell)}$  and  $\text{IP}^{(\ell)}$  are not empty sets for  $\ell = 0, \dots, t-1$ . First notice that  $\text{LP}^{(0)}$  is clearly feasible since by definition  $(x^{(0)}, y^{(0)}) \in \mathcal{D}(P)$ , meaning there exists  $(z, w) \in P$  such that  $z \leq x^{(0)}$ . By Theorem 1, there exists  $(\tilde{z}^i, \tilde{w}^i) \in S$  and  $\theta_i \geq 0$  for  $i = 1, \dots, k$  such that  $\sum_{i=1}^k \theta_i = 1$  and  $\sum_{i=1}^k \theta_i \tilde{z}^i \leq gz$ . Hence,  $\sum_{i=1}^k \theta_i \tilde{z}^i \leq gz \leq gx^{(0)}$ . So if  $x_j^{(0)} = 0$ , then  $\sum_{i=1}^k \theta_i \tilde{z}_j^i = 0$ , which implies that  $\tilde{z}_j^i = 0$  for all  $i = 1, \dots, k$  and  $j = 1, \dots, n$  where  $x_j^{(0)} = 0$ . Hence,  $z^i \leq x^{(0)}$  for  $i = 1, \dots, k$ . Therefore  $(\tilde{z}^i, \tilde{w}^i) \in \text{IP}^{(0)}$  for  $i = 1, \dots, k$ , which implies  $\text{IP}^{(0)} \neq \emptyset$ .

Now assume  $\text{IP}^{(\ell)}$  is non-empty for some  $\ell \in \{0, \dots, t-2\}$ . Since  $\text{IP}^{(\ell)} \subseteq \text{LP}^{(\ell)}$  we have  $\text{LP}^{(\ell)} \neq \emptyset$  and hence the  $\text{DominantToFeasible}(x^{(\ell)})$  has an optimal solution  $(z^*, w^*)$ .

We consider two cases. In the first case, we have  $z_{\ell+1}^* = 0$ . In this case we have  $x_{\ell+1}^{(\ell+1)} = 0$ . Since  $z^* \leq x^{(\ell+1)}$ , we have  $(z^*, w^*) \in \text{LP}^{(\ell+1)}$ . Also,  $(z^*, w^*) \in P$ . By Theorem 1 there exists  $(\tilde{z}^i, \tilde{w}^i) \in S$  and  $\theta_i \geq 0$  for  $i = 1, \dots, k$  such that  $\sum_{i=1}^k \theta_i = 1$  and  $\sum_{i=1}^k \theta_i \tilde{z}^i \leq gz^*$ . We have

$$\sum_{i=1}^k \theta_i \tilde{z}^i \leq gz^* \leq gx^{(\ell+1)} \quad (32)$$

So for  $j \in \{1, \dots, n\}$  where  $x_j^{(\ell+1)} = 0$ , we have  $\tilde{z}_j^i = 0$  for  $i = 1, \dots, k$ . Hence,  $\tilde{z}^i \leq x^{(\ell+1)}$  for  $i = 1, \dots, k$ . Hence, there exists  $(z, w) \in S$  such that  $z \leq x^{(\ell+1)}$ . We claim that  $(z, w) \in \text{IP}^{(\ell+1)}$ . If  $(z, w) \notin \text{IP}^{(\ell+1)}$  we must have  $1 \leq j \leq \ell$  such that  $z_j < x_j^{(\ell+1)}$ , and thus  $z_j = 0$  and  $x_j^{(\ell+1)} = 1$ . Without loss of generality assume  $j$  is minimum number satisfying  $z_j < x_j^{(\ell+1)}$ . Consider iteration  $j$  of Algorithm ?. Notice that  $z \leq x^{(\ell+1)} \leq x^{(j)}$ . We have  $x_j^{(j)} = 1$  which implies when we solved  $\text{DominantToFeasible}(x^{(j-1)})$  the optimal value was strictly larger

than zero. However,  $(z, w)$  is a feasible solution to  $\text{DominantToFeasible}(x^{(j-1)})$  and gives an objective value of 0. This is a contradiction, so  $(z, w) \in \text{IP}^{(\ell+1)}$ .

Now for the second case, assume  $z_{\ell+1}^* > 0$ . We have  $x_{\ell+1}^{(\ell+1)} = 1$ . Notice that for each point  $z \in \text{LP}^{(\ell)}$  we have  $z_{\ell+1} > 0$ , so for each  $z \in \text{IP}^{(\ell)}$  we have  $z_{\ell+1} > 0$ , i.e.  $z_{\ell+1} = 1$ . This means that  $(z, w) \in \text{IP}^{(\ell+1)}$ , and  $\text{IP}^{(\ell+1)} \neq \emptyset$ .

Now consider  $(x^{(t)}, y^{(t)})$ . Let  $(z, y^{(t)})$  be the optimal solution to  $\text{LP}^{(t-1)}$ . If  $x^{(t)} = 0$ , we have  $x^{(t)} = z$ , which implies that  $(x^{(t)}, y^{(t)}) \in P$ , and since  $x^{(t)} \in \{0, 1\}^n$  we have  $(x^{(t)}, y^{(t)}) \in S$ . If  $x^{(t)} = 1$ , it must be the case that  $z_t > 0$ . By the argument above there is a point  $(z', w') \in \text{IP}^{(t-1)}$ . We show that  $x^{(t)} = z'$ . Observe that for  $j = 1, \dots, n-1$  we have  $z'_j = x_j^{(t-1)} = x_j^{(t)}$ . We just need to show that  $z'_j = 1$ . Assume  $z'_j = 0$  for contradiction, then  $(z', w') \in \text{LP}^{(t-1)}$  has objective value of 0 for  $\text{DominantToFeasible}(x^{(t-1)})$ , this is a contradiction to  $(z, w)$  being the optimal solution.