

# IE8531 Project

Akshit Goyal and Jake Roth

May 11, 2022 @ 11:59pm

## Contents

<b>1. Problem</b>	<b>1</b>
1.1. Properties of model	2
1.2. Valid inequalities	3
1.2.1. Approach $F_0(\emptyset, S)$	4
1.2.2. Approach $F_0(N - S, \emptyset)$	7
1.3. Combined strength of lifted subadditive inequalities (13) & (18)	7
1.4. Constraint generation	8
<b>2. Numerical Experiments</b>	<b>8</b>
2.1. Capital Budgeting	8
2.2. Computational setup	9
2.3. Computational Results	9
<b>3. Future directions: risk-averse uncapacitated facility location</b>	<b>11</b>
<b>A. Algorithms</b>	<b>13</b>
<b>B. Proofs</b>	<b>14</b>

## 1. Problem

Consider the problem of maximizing a function  $h : 2^N \rightarrow \mathbb{R}$  over a finite set  $N$  such that  $h = f \circ a$  where  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a differentiable, strictly concave, and increasing function and  $a : 2^N \rightarrow \mathbb{R}$  is a function defined by  $a(T) := \sum_{i \in T} a_i$  where  $\{a_i\}_{i=1}^N$  are weights. Letting  $x$  denote the binary representation of the set  $T$ , with abuse of notation, we can rewrite  $a(T) = a^\top x$  and express the generic problem as

$$\begin{aligned} & \underset{x}{\text{maximize}} && f(a^\top x) \\ & \text{subject to} && x \in X \subseteq \{0, 1\}^N. \end{aligned} \tag{1}$$

It is clear that when  $a_i \geq 0$  for all  $i \in N$ , then the objective function

$$h(T) := f(a(T)) = f(a^\top x), \quad T \subseteq N, \tag{2}$$

is submodular.<sup>1</sup> An important property of submodularity is that we can express the problem (1) in the standard form of maximizing a linear function over a submodular set as

$$\begin{aligned} & \underset{w, x}{\text{maximize}} && w \\ & \text{subject to} && (w, x) \in F \end{aligned} \tag{3}$$

---

<sup>1</sup>When  $a \not\geq 0$ , if  $X = \{0, 1\}^N$ , then we may complement variables so that we can assume  $a \geq 0$  wlog.

where the mixed-integer set  $F$  is given by

$$F := \{x \in \{0, 1\}^N, w \in \mathbb{R} : w \leq f(a^\top x), x \in X\} \quad (4)$$

with  $a \in \mathbb{R}^N$ . When  $X = \{0, 1\}^N$ , denote  $F$  by  $F_0$  as in [Shi et al., 2020]. Given an explicit form of  $X$  such as  $X = \{x \in \{0, 1\}^N : Dx \leq d\}$ , we then have  $F = F_0 \cap X$ . In the remainder of our report, we focus on generating inequalities to describe the convex hull of  $F_0$  and treat the region  $X$  separately.

### 1.1. Properties of model

The majority of this section follows the presentation in [Ahmed and Atamtürk, 2011], with some additional details.

First note that  $F_0$  is the finite union of polyhedra with a common recession cone. That is, for each  $x \in \{0, 1\}^N$ , we have  $w \leq \alpha_k$  for righthand side data  $\alpha_k$  where

$$F = \bigcup_{t=1}^T \{(w, x^t) : w \leq f(a^\top x^t)\}, \quad T = 2^N,$$

and the recession cone for each  $t \in [T]$  is  $\{(0, w) : w \leq 0\}$ . This means that the convex hull of  $F_0$  is a polyhedron and can be represented with linear inequalities. However,  $F_0$  requires exponentially many constraints, and is difficult to fully characterize. In the remainder of this section, we will review an approach for lifting facet-defining inequalities of restrictions of  $F_0$ . The approach will focus on the continuous relaxation of the lifting function and use the KKT conditions of the resulting problem to construct a sequence-independent (subadditive) approximation. The subadditive approximation to the lifting function  $\gamma$  is obtained by applying the continuous relaxation and then using the KKT conditions to solve the convex program [Ahmed and Atamtürk, 2011].

Next recall a few properties of submodularity.

**Proposition 1** (Proposition 2.1 in [Nemhauser et al., 1978]). *Let  $h : 2^N \rightarrow \mathbb{R}$  be submodular function over finite set  $N$  and define  $\rho_i(S) = h(S \cup \{i\}) - h(S)$ . Then the following are equivalent.*

1.  $h(A) + h(B) \geq h(A \cup B) + h(A \cap B)$  for all  $A, B \subseteq N$ .
2.  $h(S \cup \{i\}) - h(S) \geq h(T \cup \{i\}) - h(T)$  for all  $T, S \subseteq N$  and for all  $i \in N - T$ .
3.  $h(T) \leq h(S) + \sum_{j \in T-S} \rho_j(S) - \sum_{j \in S-T} \rho_j(S \cup T - \{j\})$  for all  $S, T \subseteq N$ .

Observe that for any  $S, T \subseteq N$ , we have (i)  $\rho_i(S \cup T) \geq \rho_i(N)$  since  $S \cup T \subseteq N$  and (ii)  $\rho_i(S) \leq \rho_i(\emptyset)$  since  $\emptyset \subseteq S$ . Substituting (i) and (ii) into (3), we obtain:

$$h(T) \leq h(S) + \sum_{j \in T-S} \rho_j(S) - \sum_{j \in S-T} \rho_j(N - \{j\}), \quad \forall S, T \subseteq N \quad (5a)$$

$$h(T) \leq h(S) + \sum_{j \in T-S} \rho_j(\emptyset) - \sum_{j \in S-T} \rho_j(S - \{j\}), \quad \forall S, T \subseteq N. \quad (5b)$$

By writing the index sets  $\{i : i \in S - T\} = \{i : i \in S, i \notin T\}$  and  $\{i : i \in T - S\} = \{i : i \in T, i \in N - S\}$  in (5) and letting  $x \in \{0, 1\}^N$  be the binary representation of the set  $T$ , i.e.,

$$x_i = \begin{cases} 1, & \text{if } i \in T \\ 0, & \text{o.w.} \end{cases}$$

we obtain the *submodular inequalities* (associated with a subset  $T \subseteq N$  represented in binary by the variable  $x$ )

$$w \leq f(a^\top x) = h(T) \leq h(S) - \sum_{i \in S} \rho_i(N - \{i\})(1 - x_i) + \sum_{i \in N-S} \rho_i(S)x_i, \forall S \subseteq N \quad (6a)$$

$$w \leq f(a^\top x) = h(T) \leq h(S) - \sum_{i \in S} \rho_i(S - \{i\})(1 - x_i) + \sum_{i \in N-S} \rho_i(\emptyset)x_i, \forall S \subseteq N. \quad (6b)$$

This provides a valid formulation of  $F_0$ , which is given by

$$F_0 = \{x \in \{0, 1\}^N, w \in \mathbb{R} : (6a) \text{ or } (6b)\}.$$

Note that since  $f$  is strictly increasing, it has an inverse  $g := f^{-1}$ . Since  $f$  is strictly concave and differentiable, its inverse  $g$  is strictly convex, increasing, and differentiable. Applying  $g$ , we can rewrite  $F_0$  as

$$F_0 = \{x \in \{0, 1\}^N, w \in \mathbb{R} : a^\top x \geq g(w)\}.$$

**Remark 1.** *If we are interested in an affine function  $\tilde{a}(S) = a(S) + b$ , then we can absorb  $b$  in the definition of the functions  $f, g$ . That is, if  $\tilde{f}, \tilde{g}$  are the original functions, then we can find  $f, g$  that are inverses and maintain the properties of  $\tilde{f}, \tilde{g}$  while satisfying  $f(a(S)) = \tilde{f}(a(S) + b)$ . Thus we will assume  $b = 0$ .*

Next we will show an alternative representation of  $F_0$  that introduces the value  $a(S)$  for a particular subset  $S \subseteq N$  of interest. In particular, for any  $x \in \{0, 1\}^N$ , let  $\bar{x} := \mathbf{1} - x$  so that we have

$$\begin{aligned} a^\top x &= \sum_{i \in N} a_i x_i = \sum_{i \in N-S} a_i x_i + \sum_{i \in S} a_i x_i \\ &= \sum_{i \in N-S} a_i x_i + \sum_{i \in S} a_i (1 - \bar{x}_i) \\ &= \sum_{i \in N-S} a_i x_i + \sum_{i \in S} a_i + \sum_{i \in S} -a_i \bar{x}_i \\ &= \sum_{i \in N-S} a_i x_i + a(S) + \sum_{i \in S} -a_i \bar{x}_i \\ &= \sum_{i \in N-S} a_i x_i + a(S) + \sum_{i \in S} -a_i (1 - x_i). \end{aligned}$$

Moving  $a(S)$  to the RHS, we have

$$F_0 = \left\{x \in \{0, 1\}^N, w \in \mathbb{R} : \sum_{i \in N-S} a_i x_i + \sum_{i \in S} -a_i (1 - x_i) \geq g(w) - a(S)\right\}.$$

## 1.2. Valid inequalities

Next we introduce an approach to construct valid inequalities for  $F_0$  by considering restrictions of  $F_0$  with a particular subset  $S \subseteq N$  in mind. Following [Ahmed and Atamtürk, 2011], the approach has two forms:

- (a) For any  $S \subseteq N$ , fix  $x_i = 1$  for  $i \in S$ ; lift the resulting inequalities.
- (b) For any  $S \subseteq N$ , fix  $x_i = 0$  for  $i \in N - S$ ; lift the resulting inequalities.

Either set of inequalities (a) or (b) alone is sufficient to provide a formulation of  $F_0$ . Lifting both is stronger. Refer to Section 1.3 for explanation.

### 1.2.1. Approach $F_0(\emptyset, S)$

For any  $S \subseteq N$ , consider the restriction of  $F_0$  defined by fixing  $x_i = 1$  for  $i \in S$ . Define this restriction to be

$$F_0(\emptyset, S) := \left\{ x \in \{0, 1\}^{N-S}, w \in \mathbb{R} : \sum_{i \in N-S} a_i x_i \geq g(w) - a(S) \right\}. \quad (7)$$

Applying  $f(\cdot)$  to both sides of  $F_0(\emptyset, S)$

$$w \leq f\left(\sum_{i \in N-S} a_i x_i + a(S)\right) = f\left(\sum_{i \in N-S} a_i x_i + \sum_{i \in S} a_i x_i\right) = f(a^\top x) = h(T)$$

and applying the submodular inequality in (6a) gives that

$$w \leq h(T) \leq h(S) + \sum_{i \in N-S} \rho_i(S) x_i, \quad \forall S \subseteq N \quad (8)$$

is a valid constraint for  $F_0(\emptyset, S)$ . Further, it is facet-defining for  $\text{conv}[F_0(\emptyset, S)]$ . To see this consider that there exist  $(x^t, w)$  that satisfy the desired constraint for  $t \in [2^{N-S}]$  by setting  $w$  appropriately. Given a facet-defining inequality for a restriction of  $F_0$ , it is natural to seek a lifted version for the full region  $F_0$  by lifting for dimensions  $x_i$  with  $i \in S$ .

Next, consider lifting variable  $x_j$  associated with constraint (8) where  $S$  is the set of interest,<sup>2</sup> and for all  $(x, w) \in F_0(\emptyset, S)$ ,  $x_j = 0$  we seek  $\zeta$  satisfying

$$w \leq h(S) - \zeta + \sum_{i \in N-S} \rho_i(S) x_i \iff \zeta \leq -w + \sum_{i \in N-S} \rho_i(S) x_i.$$

Translating the condition  $(x, w) \in F_0(\emptyset, S)$ ,  $x_j = 0$  gives the constraint set as

$$-a_j + \sum_{i \in N-S} a_i x_i + a(S) \geq g(w), \quad x \in \{0, 1\}^{N-S}, w \in \mathbb{R}.$$

The corresponding lifting problem is

$$\zeta := \max_{w, x} \quad w - \sum_{i \in N-S} \rho_i(S) x_i - h(S) \quad (9a)$$

$$\text{s.t.} \quad \sum_{i \in N-S} a_i x_i \geq g(w) - a(S) + a_j, \quad (9b)$$

$$x \in \{0, 1\}^{N-S}, x_j = 0, w \in \mathbb{R}. \quad (9c)$$

---

<sup>2</sup>Consider lifting the original constraint in “ $g$ ” space via

$$-\alpha_j(1 - x_j) + \sum_{i \in N-S} a_i x_i \geq g(w) - a(S).$$

When  $x_j = 1$ , we recover the original constraint, and when  $x_j = 0$ , we seek

$$\begin{aligned} \alpha_j &\leq \sum_{i \in N-S} a_i x_i + a(S) - g(w), \quad \forall (x, w) \in F_0(\emptyset, S), x_j = 0 \\ &= \min_{(x, w) \in F_0(\emptyset, S), x_j = 0} \left\{ \sum_{i \in N-S} a_i x_i + a(S) - g(w) \right\}, \end{aligned}$$

but by setting  $x_i = 0$  for  $i \in N - S$  and taking  $w = f(a(S))$  we can recover the coefficient  $\alpha_j = 0$ , which is worse than the original submodular constraint.

Replacing  $a_j$  with a parameter  $\delta$  and removing the  $x_j = 0$  constraint leads to the  $\delta$ -parametrized lifting problem. For constraint (8) associated with the set  $S$ , the problem is given by

$$\zeta(\delta) := \max_{w, x} w - \sum_{i \in N-S} \rho_i(S) x_i - h(S) \quad (10a)$$

$$(L : S) \quad \text{s.t.} \quad \sum_{i \in N-S} a_i x_i \geq g(w) - a(S) - \delta, \quad (10b)$$

$$x \in \{0, 1\}^{N-S}, w \in \mathbb{R}. \quad (10c)$$

The scalar  $\delta$  appears in the RHS of the constraint since applying  $g(\cdot)$  to the constraint in (7) gives

$$-a_j(1-0) + \sum_{i \in N-S} a_i x_i \geq g(w) - a(S) \iff \sum_{i \in N-S} a_i x_i \geq g(w) - a(S) + a_j.$$

Identifying  $\delta = -a_j$ , we recover the standard lifting problem. For any  $\delta$  such that  $-\delta \leq a_j$  ( $-\delta > a_j$ ), we have a relaxation (restriction) of the standard problem. That is, letting  $\zeta$  denote the value of the standard lifting problem, we have that  $\zeta(\delta) \geq \zeta$  for  $-\delta \leq a_j$  and  $\zeta(\delta) \leq \zeta$  for  $-\delta > a_j$ .

In general, it is necessary to consider all orderings of liftings for all subsets  $S \subseteq N$  to find the convex hull of  $F_0$ . Since each subproblem is a submodular maximization problem, which is NP-hard in general, this is difficult. However, a greedy approach (Algorithm 1) can solve  $(L : S)$ . The idea is that taking elements from  $N - S$  with the largest weights is optimal so long as the contribution of the  $k$ th element to  $f(\sum_{i=1}^k a_i x_i + a(S))$  is at least as large as the reduction due to  $\rho_k(S)$ . Solving the subproblem provides a value  $f(a(S) + A_k + \delta)$  that is a valid upper bound  $w$  for all  $\delta \in [-A_k, -A_{k-1}]$ . In fact, it is shown in [Shi et al., 2020] that  $\zeta(\delta)$  provides a *sequence-independent* lifting function, meaning that the convex hull of  $F_0$  can be obtained by lifting in *any* order.

**Proposition 2** (Proposition 5 in [Ahmed and Atamtürk, 2011]). *Algorithm 1 solves lifting problem  $(L : S)$ .*

*Proof.* See Appendix. □

Based on Proposition 2, the optimal value of  $(L : S)$  is given by the following expression

$$\zeta(\delta) = f(a(S) + A_k + \delta) - \sum_{i=1}^k \rho_i(S) - f(a(S)), \quad \delta \in [-A_k, -A_{k-1}], \quad k \in [m+1]. \quad (11)$$

where  $m = |N - S|$  and we may take  $-A_{m+1} = -\infty$ . As an implementation note, if we introduce  $-A_{m+1} = -\infty$ , then we can extend the domain of  $\zeta$  for  $\delta \leq -A_m$ . This is important in the case when there exists an element  $i \in S$  such that  $a_i \geq \sum_{j \in N-S} a_j = A_m$ . It can be easily handled by calling  $\zeta(\delta)$  for  $\delta \leq -A_m$ .

An important property of  $\zeta$  is that it is continuous on  $\mathbb{R}_-$  and concave on each interval  $[-A_k, A_{k-1}]$  for  $k = 1, 2, \dots, m+1$ . It is stated that  $\zeta$  is not subadditive in [Ahmed and Atamtürk, 2011], but it was later shown in [Shi et al., 2020] that  $\zeta$  is in fact subadditive, and hence sequence-independent. This motivated the authors in [Ahmed and Atamtürk, 2011] to study the concave upper envelope of  $\zeta$ , which is a subadditive function. It is defined as

$$\gamma(\delta) := \begin{cases} \zeta(\mu_k - A_{k-1}) - \rho_k(S) \frac{b_k(\delta)}{a_k}, & \text{if } \mu_k - A_k \leq \delta \leq \mu_k - A_{k-1}; \\ \zeta(\delta), & \text{otherwise} \end{cases} \quad (12)$$

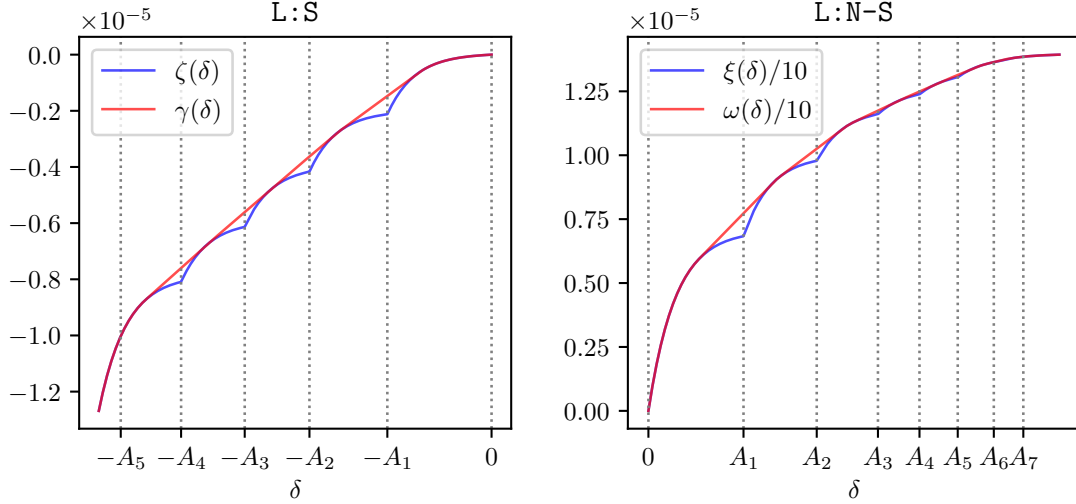


Figure 1: Lifting functions for  $f(x) = -e^{-x/\lambda}$  and  $g(y) = \lambda \log(-1/y)$  with  $\lambda = 0.1$ .

*Left:* Piecewise concave  $\zeta$  and its upper concave envelope  $\gamma$ .

*Right:* Piecewise concave  $\xi$  and its upper concave envelope  $\omega$ .

with

$$\mu_k := g\left((g')^{-1}\left(\frac{a_k}{\rho_k(S)}\right)\right)$$

$$b_k(\delta) := \mu_k - A_{k-1} - \delta$$

for  $k \in [m+1]$ . To evaluate  $\gamma$  at a point  $\delta \in \mathbb{R}_-$ , we compute  $l_k = \mu_k - A_k$  and  $u_k = \mu_k - A_{k-1}$  for each  $k$  in  $1, 2, \dots, m+1$ . If there exists a  $k^* \in [m+1]$  such that  $\delta \in [l_{k^*}, u_{k^*}]$ , then  $\gamma(\delta)$  is given by  $\zeta(\mu_{k^*} - A_{k^*-1})$ ; otherwise  $\gamma(\delta)$  is given by  $\zeta(\delta)$ . For some intuition on where this comes from, see Appendix B.

The left panel of Figure 1 plots  $\zeta$  and its upper concave envelope  $\gamma$  on a test instance for the utility function  $f(x) = -e^{-x/\lambda}$  with  $\lambda = 0.1$ ,  $|N| = 12$ ,  $|S| = 5$ , and  $a_i \in [0, 1]$  for  $i \in [12]$ . The particular instance is

$$a = [0.95, 0.87, 0.71, 0.67, 0.59, 0.58, 0.56, 0.55, 0.38, 0.35, 0.33, 0.1]$$

with  $S = [2, 4, 7, 9, 10, 11, 12]$ .

**Proposition 3** (Proposition 6 in [Ahmed and Atamtürk, 2011]). *The function  $\gamma : \mathbb{R}_- \rightarrow \mathbb{R}$  defined in (12) is the concave upper envelope of the function  $\zeta$  defined in (11).*

The reason for this is that the version of the lifting problem with continuous  $x$  is a parametric concave maximization problem where the parameter  $\delta$  appears in the RHS of the constraints. Thus the value is piecewise linear.

Since  $\gamma \geq \zeta$ , and by the subadditivity of  $\gamma$ , it holds that

$$\zeta\left(\sum_{i \in S} -a_i \bar{x}_i\right) \leq \gamma\left(\sum_{i \in S} -a_i \bar{x}_i\right) \leq \sum_{i \in S} \gamma(-a_i) \bar{x}_i.$$

This provides the subadditive lifting inequality

$$w \leq h(S) + \sum_{i \in S} \gamma(-a_i)(1 - x_i) + \sum_{i \in N-S} \rho_i(S)x_i, \quad (13)$$

which is the main purpose of this section.

### 1.2.2. Approach $F_0(N - S, \emptyset)$

The approach for lifting (6b) follows a similar procedure as outlined in  $F_0(\emptyset, S)$ .

1. Consider the restriction of  $F_0$  defined by fixing  $x_i = 0$  for  $i \in N - S$ ; call it  $F_0(N - S, \emptyset)$ .
2. Apply (6b) to the restriction  $(N - S, \emptyset)$  to obtain

$$w \leq h(T) \leq h(S) - \sum_{i \in N-S} \rho_i(S - \{i\})(1 - x_i), \quad \forall S \subseteq N \quad (14)$$

3. Lift (14) by defining the  $d$ -parametrized lifting problem  $(L : N - S)$

$$\xi(\delta) := \max_{w, x} w + \sum_{i \in S} \rho_i(S - \{i\})(1 - x_i) - h(S) \quad (15a)$$

$$(L : N - S) \quad \text{s.t.} \quad \sum_{i \in S} -a_i(1 - x_i) \geq g(w) - a(S) - \delta, \quad (15b)$$

$$x \in \{0, 1\}^S, \quad w \in \mathbb{R}. \quad (15c)$$

4. Solve (15) with the greedy Algorithm 2 to compute  $\xi$  in closed form as

$$\xi(\delta) = f(a(S) - A_k + \delta) + \sum_{i=1}^k \rho_i(S - \{i\}) - f(a(S)) \quad (16)$$

after finding  $k$  such that  $A_{k-1} \leq \delta \leq A_k$  with  $k = 1, 2, \dots, r$ ; if  $\delta \geq A_r$ , then return  $\xi(\delta)$  with  $k = r$ .

5. Construct the upper concave envelope of  $\xi$ , again expressible in closed form as

$$\omega(\delta) := \begin{cases} \xi(A_k - \nu_k) - \rho_k(S - \{k\})b_k(\delta)/a_k, & \text{if } A_{k-1} - \nu_k \leq \delta \leq A_k - \nu_k; \\ \xi(\delta), & \text{otherwise} \end{cases} \quad (17)$$

with  $\nu_k = a(S) - g((g')^{-1}(a_k/\rho_k(S - \{k\})))$  and  $b_k(\delta) = A_k - \nu_k - \delta$  for  $k \in [r]$ ; if  $\delta \geq A_r$ , then return  $\xi(\delta)$ .

6. Lifted subadditive inequalities are given by

$$w \leq h(S) - \sum_{i \in S} \rho_i(S - \{i\})(1 - x_i) + \sum_{i \in N-S} \omega(a_i)x_i. \quad (18)$$

### 1.3. Combined strength of lifted subadditive inequalities (13) & (18)

First note that by Corollary 1 from [Ahmed and Atamtürk, 2011], any extreme point of the continuous relaxation of  $F_0$  has at most one fractional component. Based on this, consider a particular node of a branch-and-bound tree with a binary optimal solution to the LP relaxation,  $x^*$ . Define the sets  $T = \{i : x_i^* = 1\}$  and  $N - T = \{i : x_i^* = 0\}$ . Consider another point  $\hat{x}$  with

1.  $\hat{x}_i = 1$  for  $i \in T$ ;
2.  $0 < \hat{x}_k < 1$  for  $k \in N - T$ ;
3.  $\hat{x}_i = 0$  for  $i \in N - (T \cup k)$ .

The corresponding extreme point  $(\hat{w}, \hat{x})$  is such that

$$g(\hat{w}) = a(T) + a_k \hat{x}_k.$$

Evaluating (13) at  $(\hat{w}, \hat{x})$  gives

$$\begin{aligned} f(a(T) + a_k \hat{x}_k) &\leq f(a(T)) + (f(a(T) + a_k) - f(a(T))) \hat{x}_k \\ &= (1 - \hat{x}_k) f(a(T)) + \hat{x}_k f(a(T) + a_k) \end{aligned}$$

which is cut off because of strict concavity of  $f$ . Further, this inequality will cut off all extreme points with fractional component  $k$  in  $N - T$  whose components of  $T$  are 1.

Next consider another point  $\tilde{x}$  with

1.  $0 < \tilde{x}_k < 1$  for  $k \in T$ ;
2.  $\tilde{x}_i = 1$  for  $i \in T - k$ ;
3.  $\tilde{x}_i = 0$  for  $i \in N - T$ .

The corresponding extreme point  $(\tilde{w}, \tilde{x})$  is such that

$$g(\tilde{w}) = a(T - k) + a_k \tilde{x}_k = a(T) - (1 - \tilde{x}_k) a_k$$

Evaluating (18) at  $(\tilde{w}, \tilde{x})$  gives

$$\begin{aligned} f(a(T) - (1 - \tilde{x}_k) a_k) &\leq f(a(T)) - (f(a(T)) - f(a(T) - a_k))(1 - \tilde{x}_k) \\ &= \tilde{x}_k f(a(T)) + (1 - \tilde{x}_k) f(a(T) - a_k) \end{aligned}$$

which is cut off because of strict concavity of  $f$ . Further, this inequality will cut off all extreme points with fractional component  $k$  in  $T$  whose components of  $N - T$  are 0. Based on this argument, we see that adding both forms of the lifted subadditive inequalities results in a strengthened LP relaxation at the remaining unexplored nodes relative to adding either one alone.

## 1.4. Constraint generation

A heuristic cut procedure for removing fractional solutions after solving the LP relaxation is introduced in [Ahmed and Atamtürk, 2011]; we do not consider this approach further.

## 2. Numerical Experiments

We conduct numerical experiments on randomly generated instances of a capital budgeting problem with varying— problem size  $n$ , # of scenarios  $m$ , risk tolerance  $\lambda$ . The aim is to demonstrate the effectiveness of lifted inequalities (13) & (18) to close the optimality gap on large-sized instances as compared to the usual submodular inequalities (6a) & (6b).

### 2.1. Capital Budgeting

The problem is to commit to an investment decision under uncertainty subject to a budget constraint. Let  $q_j$ ,  $j \in N = \{1, \dots, n\}$  denote the capital requirements of a set  $N$  of investment options. Let  $a_i \in \mathbb{R}^N$  denote the vector of value of investment at some future time under scenario  $i = 1, \dots, m$ . Using an exponential utility function, the problem is stated as

$$\begin{aligned} \max \quad & \sum_{i=1}^m \pi_i \left( 1 - \exp \left( -\frac{a_i^\top x}{\lambda} \right) \right) \\ \text{s.t.} \quad & q^\top x \leq 1, \quad x \in \{0, 1\}^N \end{aligned}$$

We can rewrite the problem as given below by introducing a continuous variable

$$\begin{aligned} \max \quad & 1 + \sum_{i=1}^m \pi_i w_i \\ \text{s.t.} \quad & w_i \leq -\exp\left(-\frac{a_i^\top x}{\lambda}\right) \quad \forall i = 1, \dots, m \\ & q^\top x \leq 1, \quad x \in \{0, 1\}^N \end{aligned}$$

For each  $i = 1, \dots, m$ , we generate usual submodular inequalities and lifted ones for the following polyhedron,

$$F_0^i = \left\{ x \in \{0, 1\}^N, w \in \mathbb{R} : w \leq -\exp\left(-\frac{a_i^\top x}{\lambda}\right) \right\}.$$

## 2.2. Computational setup

Following [Ahmed and Atamtürk, 2011], the parameters for the capital budgeting problem are generated according to the following:

1. The capital requirement  $q_j$  is sampled independently for each  $j \in N$  from a continuous uniform distribution over  $[0, 0.2]$ , i.e.  $\mathcal{U}(0, 0.2)$ .
2. The return on investment  $a_i = (a_{ij})_{j \in N}$  under scenario  $i = 1, \dots, m$  with probability  $\pi_i = 1/m$  is given by

$$a_{ij} = r_{ij} q_j \quad \text{where} \quad r_{ij} = \exp(\alpha_j + \beta_j \ln f_i + \epsilon_{ij})$$

3. For each  $j \in N$ ,  $\alpha_j \sim \mathcal{U}(0.05, 0.10)$  and  $\beta_j \in \mathcal{U}(0, 1)$  which are fixed across scenarios.
4. For each scenario  $i$ ,

$$\ln f_i \sim \mathcal{N}(0.05, 0.0025) \quad \text{and for each } j \in N : \epsilon_{ij} \sim \mathcal{N}(0, 0.0025)$$

where  $\mathcal{N}(\mu, \sigma^2)$  is the normal distribution with mean  $\mu$  and variance  $\sigma^2$ .

## 2.3. Computational Results

To compare the performance of the two sets of inequalities-(6a)- (6b) and (13) & (18), we consider values of  $n \in \{15, 25, 50\}$ ,  $m \in \{1, 25, 100\}$ ,  $\lambda \in \{1, 2\}$  which gives the total 18 possible instance settings. For each instance setting, we generate 5 random instances of parameters of the capital budgeting problem. All the results are reported after averaging across these 5 instances.

We run our computations using Gurobi 9.1.2 on 64-bit Windows having Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz and 16 GB of memory. The inequalities are added in a branch & cut fashion using Gurobi's `callback` function and setting `LazyConstraints` paramter to 1. Each inequality is added as a lazy constraint whenever an integer solution is encountered. The `TimeLimit` paramter is set to 600 seconds. For more precise comparison of the individual strength of the two sets of inequalities, we further shut off the cuts generated by Gurobi by setting `Cuts` paramter to 0. The results of the computations are presented in Table 1.

In addition, we performed a single set of experiments at  $n = 25$ ,  $m = 100$ , and  $\lambda = 1, 2$  for the same problem to compare the performance of the exact lifted inequalities (coefficients  $\zeta$  and  $\xi$  rather than  $\gamma$  and  $\omega$ ). Results of the computation are presented in Table 2.

$n$	$m$	$\lambda$	Submodular Inequalities (6a) & (6b)						Lifting Inequalities (13) & (18)					
			# cuts	# nodes	ObjVal	ObjBnd	%Gap	Runtime	# cuts	# nodes	ObjVal	ObjBnd	%Gap	Runtime
15	1	1	434	1778	0.69339	0.69339	0.00	0.42	20	135	0.69339	0.69340	0.00	0.07
		2	222	1255	0.44631	0.44631	0.00	0.24	11	40	0.44631	0.44631	0.00	0.05
	25	1	15945	2333	0.68460	0.68460	0.00	22.80	1295	928	0.68460	0.68460	0.00	1.79
		2	10105	2028	0.43852	0.43852	0.00	12.94	695	624	0.43852	0.43852	0.00	1.10
	100	1	66980	2140	0.68317	0.68317	0.00	114.53	5700	1082	0.68317	0.68318	0.00	10.32
		2	42900	1717	0.43724	0.43724	0.00	66.85	2300	423	0.43724	0.43724	0.00	3.69
25	1	1	336	152455	0.70603	0.75027	6.27	600.29	245	6112	0.70762	0.70762	0.00	1.04
		2	249	160668	0.45838	0.48189	5.14	600.20	52	799	0.45938	0.45938	0.00	0.22
	25	1	21225	12062	0.68816	0.78541	14.13	600.56	4485	30230	0.69111	0.69254	0.21	423.74
		2	14055	15291	0.44303	0.50141	13.17	600.65	3305	7427	0.44444	0.44445	0.00	18.08
	100	1	77100	3709	0.68702	0.81744	18.99	600.62	15780	11350	0.69116	0.69382	0.38	596.80
		2	59780	3815	0.44199	0.51606	16.75	600.48	15980	5105	0.44446	0.44446	0.00	97.49
50	1	1	95	149703	0.70248	0.88498	25.98	600.18	98	196744	0.70961	0.70967	0.01	311.55
		2	91	164873	0.46045	0.56340	22.35	600.19	141	6914	0.46114	0.46115	0.00	2.70
	25	1	4415	10564	0.68833	0.92382	34.21	600.68	1875	26280	0.69319	0.70089	1.11	600.64
		2	3865	14394	0.44505	0.56188	26.25	600.99	1345	36215	0.44703	0.44887	0.41	600.86
	100	1	32020	2850	0.68842	0.94189	36.82	600.70	8580	8529	0.69214	0.70196	1.42	601.48
		2	12180	4782	0.44485	0.56282	26.51	601.69	5220	11115	0.44670	0.44947	0.62	601.80

Table 1: Comparison of computational performance for different values of  $n, m, \lambda$ . Results in each row are averaged across 5 instances. (**ObjVal**, **ObjBnd**: Best known lower bound, upper bound respectively within the time limit. % **Gap** is the percentage gap between best known upper and lower bound after time limit reached.)

### Observations:

1. For  $n = 15$ , all instances solved to optimality within the time limit for all values of  $m$  and  $\lambda$ . The runtimes for lifted inequalities is considerably smaller than weak submodular inequalities. Moreover, for each of the inequalities, the runtime decreases with  $\lambda$ . This is because the utility function is more nonlinear for small  $\lambda$  which corresponds to higher risk-averseness.
2. From all instances, one can easily observe that the lifted inequalities provides more strength to the LP relaxation since very few of them added as compared to the usual submodular inequalities.
3. For  $n = 25$  and  $50$ , the lifted inequalities closes the gap in time limit for most instances whereas the usual submodular inequalities provide large gap. For instances not solved in time limit, lifted inequalities terminate with gap approximately around  $1 - 1.5\%$ . Similar to  $n = 15$ , the risk-tolerance parameter  $\lambda = 2$  provide either smaller runtime or gap as compared to  $\lambda = 1$  for both sets of inequalities.
4. We observed minimal improvement in runtime and optimality gap using exact lifting as compared to the lifting using upper concave envelope, as shown in Table 2.

$n$	$m$	$\lambda$	instance	Lifting Inequalities ( $\gamma$ and $\omega$ ) (13) & (18)				Exact Lifting Inequalities ( $\zeta$ and $\xi$ )			
				# cuts	# nodes	%Gap	Runtime	# cuts	# nodes	%Gap	Runtime
25	100	1	1	20900	8344	0.43	600.68	12300	10297	0.31	601.11
			2	22900	13041	0.00	578.00	17500	6380	0.00	114.13
			3	13300	10064	0.48	600.46	6500	14662	0.49	600.68
			4	14500	9392	0.43	601.22	20900	12978	0.28	600.45
			5	7300	15866	0.58	600.11	15700	12771	0.56	600.09
		2	1	15700	3650	0.00	49.81	15700	3208	0.00	51.73
			2	4500	973	0.00	9.47	3900	865	0.00	7.42
			3	26500	9325	0.00	217.78	22100	7625	0.00	138.93
			4	8900	2037	0.00	25.30	7100	1260	0.00	15.77
			5	24300	9538	0.00	168.31	16900	10483	0.00	238.34

Table 2: Comparing lifting inequalities using concave upper envelope vs exact lifting.

### 3. Future directions: risk-averse uncapacitated facility location

In this section, we present some possible future directions. Recall the *uncapacitated facility location* (UFL) problem. Given a set  $[m]$  of customers and a set  $[n]$  of facilities, the UFL can be stated as

$$\text{minimize} \quad \sum_{i \in [m]} \sum_{j \in [n]} c_{ij} y_{ij} + \sum_{j=1}^n f_j x_j \quad (19a)$$

$$\text{subject to} \quad \sum_{j \in [n]} y_{ij} = 1, \quad i \in [m], \quad (19b)$$

$$y_{ij} \leq x_j, \quad i \in [m], j \in [n], \quad (19c)$$

$$x_j \in \{0, 1\}, j \in [n], \quad (19d)$$

$$y_{ij} \in \mathbb{R}_+, \quad i \in [m], j \in [n]. \quad (19e)$$

Suppose that  $f_j \geq 0$  for  $j \in [n]$ . Then the objective function for UFL can be expressed equivalently as

$$\max \quad \sum_{i \in [m]} \sum_{j \in [n]} -c_{ij} y_{ij} + \sum_{j \in [n]} -f_j x_j \iff \max \quad \sum_{i \in [m]} \sum_{j \in [n]} -c_{ij} (1 - \bar{y}_{ij}) + \sum_{j \in [n]} -f_j (1 - \bar{x}_j)$$

$$\iff \max \sum_{i \in [m]} \sum_{j \in [n]} c_{ij} \bar{y}_{ij} + \sum_{j \in [n]} f_j \bar{x}_j$$

for  $\bar{x}_j = 1 - x_j$  and  $\bar{y}_{ij} = 1 - y_{ij}$ . Thus a “risk-averse” UFL (RUFL) can be expressed by the objective function

$$f_1\left(\sum_{i \in [m]} \sum_{j \in [n]} c_{ij} \bar{y}_{ij}\right) + f_2\left(\sum_{j \in [n]} f_j \bar{x}_j\right)$$

for concave functions  $f_1, f_2$ . The RUFL problem can be expressed as

$$\begin{array}{ll} \text{maximize} & w_1 + w_2 \\ w_1, w_2, \bar{x}, \bar{y} & \end{array} \quad (20a)$$

$$\text{subject to} \quad w_1 \leq f_1\left(\sum_{i \in [m]} \sum_{j \in [n]} c_{ij} \bar{y}_{ij}\right), \quad (20b)$$

$$w_2 \leq f_2\left(\sum_{j \in [n]} f_j \bar{x}_j\right), \quad (20c)$$

$$\sum_{j \in [n]} \bar{y}_{ij} = n - 1, \quad i \in [m], \quad (20d)$$

$$\bar{y}_{ij} \geq \bar{x}_j, \quad i \in [m], j \in [n], \quad (20e)$$

$$\bar{x}_j \in \{0, 1\}, \quad j \in [n], \quad (20f)$$

$$\bar{y}_{ij} \leq 1, \quad i \in [m], j \in [n]. \quad (20g)$$

We could also consider an alternative mixed-integer set corresponding to the joint objective

$$w \leq f_1\left(\sum_{i \in [m]} \sum_{j \in [n]} c_{ij} \bar{y}_{ij}\right) + f_2\left(\sum_{j \in [n]} f_j \bar{x}_j\right).$$

To solve the RUFL, consider the following cases

1.  $f_1$  is linear: an approach using a combination of submodular cuts for  $f_2$  and Benders’ cuts for  $f_1$  should solve the problem.
2.  $f_2$  is concave: an approach using submodular cuts for each set separately should solve the problem.
3. Question: are there strong valid inequalities for UFL that can be incorporated into the submodular lifting problem? We observe that if we want to construct the convex hull of a region  $F = F_0 \cap \{x : Dx \leq d\}$  via lifting, then it is important to be able to solve efficiently the continuous relaxation for the lifting problem associated with a valid inequality.

## References

- [Ahmed and Atamtürk, 2011] Ahmed, S. and Atamtürk, A. (2011). Maximizing a class of submodular utility functions. *Mathematical Programming*, 128:149–169. <https://doi.org/10.1007/s10107-009-0298-1>.
- [Nemhauser et al., 1978] Nemhauser, G. L., Wolsey, L., and Fisher, M. (1978). An analysis of approximations for maximizing submodular set functions. *Mathematical Programming*, 14:265–294.
- [Shi et al., 2020] Shi, X., Prokopyev, O. A., and Zeng, B. (2020). Sequence independent lifting for the set of submodular maximization problem. *Integer Programming and Combinatorial Optimization*, pages 378–390. [https://doi.org/10.1007/978-3-030-45771-6\\_29](https://doi.org/10.1007/978-3-030-45771-6_29).

## Appendices

### A. Algorithms

We include the two greedy algorithms for the lifting problem below.

---

**Algorithm 1: Greedy L:S**

---

```
Input:  $a; d; S; N$  ; /* Sorted  $a$  (desc.),  $S$ , and  $N$ . */
1  $m = \text{length}(N - S)$ ;
2  $A_k = \sum_{i=1}^k a[i]$ , for  $k = 1, \dots, m$ ;  $A_0 = 0$ ;
3  $k = 0$ ;  $x = 0$ ;
4 while  $k < m$  and  $A_k + d < 0$  do
5   |  $k \leftarrow k + 1$ ;  $x_k \leftarrow 1$ ;
6 end
7  $w = f(a(S) + A_k + d)$ ;
8  $v = w - \sum_{i \in N-S} \rho_i(S) x[i] - f(a(S))$ ;
9 return  $v, w, x$ 
```

---

---

**Algorithm 2: Greedy L:N-S**

---

```
Input:  $a; d; S; N$  ; /* Sorted  $a$  (desc.),  $S$ , and  $N$ . */
1  $r = \text{length}(S)$ ;
2  $A_k = \sum_{i=1}^k a[i]$ , for  $k = 1, \dots, r$ ;  $A_0 = 0$ ;
3  $k = 0$ ;  $x = 0$ ;
4 while  $k < r$  and  $A_k < d < 0$  do
5   |  $k \leftarrow k + 1$ ;  $x_k \leftarrow 0$ ;
6 end
7  $w = f(a(S) - A_k + d)$ ;
8  $v = w - \sum_{i \in S} \rho_i(S - \{i\}) x[i] - f(a(S))$ ;
9 return  $v, w, x$ 
```

---

## B. Proofs

*Analysis of  $\gamma$  via continuous relaxation.* To see where the expression for  $\gamma$  comes from, consider the continuous relaxation of the lifting problem (10), which we will introduce as below

$$\gamma(\delta) := \max_{w, x} \quad w - \sum_{i \in N-S} \rho_i(S) x_i - h(S) \quad (21a)$$

$$\text{s.t.} \quad \sum_{i \in N-S} a_i x_i \geq g(w) - a(S) - \delta, \quad (21b)$$

$$x \in [0, 1]^{N-S}, \quad w \in \mathbb{R}. \quad (21c)$$

Then for any  $\delta \leq 0$ ,  $\hat{\zeta}(\delta)$  is a convex program. Further, if at least one  $a_i > 0$ , then strong duality holds because there exists a Slater point by taking  $x = \mathbf{1}$  and  $g$  appropriately. Thus the KKT conditions are sufficient for optimality. Introducing the Lagrangian

$$\mathcal{L}(x, w; \lambda, \alpha, \beta) := w - \sum_{i \in N-S} \rho_i x_i + \lambda \cdot \left( a(S) + \delta - g(w) + \sum_{i \in N-S} a_i x_i \right) + \beta^\top x + \alpha^\top (\mathbf{1} - x), \quad (22)$$

then

$$\gamma(\delta) = \max_{x, w} \min_{\lambda, \alpha, \beta} \mathcal{L}(x, w; \lambda, \alpha, \beta) = \min_{\lambda, \alpha, \beta} \max_{x, w} \mathcal{L}(x, w; \lambda, \alpha, \beta) \quad (23)$$

and

$$\mathcal{L}(x, w; \lambda, \alpha, \beta) := w - \sum_{i \in N-S} \rho_i(S) x_i + \lambda \cdot \left( a(S) + \delta - g(w) + \sum_{i \in N-S} a_i x_i \right) + \beta^\top x + \alpha^\top (\mathbf{1} - x)$$

$$\nabla_{x_i} \mathcal{L} = -\rho_i(S) + \lambda a_i + \beta_i - \alpha_i = 0 \implies \lambda a_i + \beta_i - \alpha_i = \rho_i(S)$$

$$\nabla_w \mathcal{L} = 1 - \lambda g'(w) = 0 \implies \lambda = 1/g'(w)$$

$$0 \leq \lambda \perp \left( a(S) + \delta - g(w) + \sum_{i \in N-S} a_i x_i \right) \geq 0$$

$$0 \leq \beta \perp x \geq 0$$

$$0 \leq \alpha \perp (\mathbf{1} - x) \geq 0.$$

1. When  $x_i^* = 1$ , then  $\beta_i^* = 0$ . We obtain a threshold on  $\lambda^*$  since  $\alpha_i^* \geq 0$  from  $\lambda^* \geq \rho_i(S)/a_i$ .
2. When  $x_i^* = 0$ , then  $\alpha_i^* = 0$ . We obtain a similar threshold from  $\lambda^* \leq \rho_i(S)/a_i$ .
3. When  $0 < x_i^* < 1$ , and  $\alpha_i^* = \beta_i^* = 0$ . Then  $\lambda^* = \rho_i(S)/a_i = 1/g'(w^*)$ .

Since the  $a_i$  elements are sorted in descending order, the  $\rho_i(S)$  components are ordered as well. Thus  $\rho_i(S)/a_i$  are non-decreasing. We have that when  $\rho_i(S)/a_i < \lambda^*$ , then  $x_i^* = 1$ , when  $\rho_i(S)/a_i > \lambda^*$ , then  $x_i^* = 0$ , and when  $\rho_i(S)/a_i = \lambda^*$ , then  $x_i^*$  is fractional. Its value can be determined from substituting the expression for  $w^*$  (from the constraint) so that once the index  $k$  is found such that  $x_{k+1}^*$  is fractional, then

$$\begin{aligned} \rho_{k+1}(S)/a_{k+1} &= \lambda^* = 1/g'(w^*) = 1/g' \left( f \left( \sum_{i \in N-S} a_i x_i^* + a(S) + \delta \right) \right) \\ &= 1/g' \left( f \left( \sum_{i=1}^k a_i + a_{k+1} x_{k+1}^* + a(S) + \delta \right) \right). \end{aligned}$$

Solving for  $x_{k+1}^*$  gives

$$x_{k+1}^* = \frac{g \left( (g')^{-1} \left( \frac{a_{k+1}}{\rho_{k+1}} \right) \right) - A_k - a(S) - \delta}{a_{k+1}}$$

$$= \frac{b_{k+1}(\delta) - a(S)}{a_{k+1}}, \quad \mu_{k+1} = g\left((g')^{-1}\left(\frac{a_{k+1}}{\rho_{k+1}(S)}\right)\right), \quad b_{k+1}(\delta) = \mu_{k+1} - A_k - \delta$$

$$w^* = (g')^{-1}\left(\frac{a_{k+1}}{\rho_{k+1}(S)}\right).$$

Substituting into the objective function gives

$$\gamma(\delta) = (g')^{-1}\left(\frac{a_{k+1}}{\rho_{k+1}(S)}\right) - \sum_{i=1}^k \rho_i(S) - \rho_{k+1}(S) \frac{b_{k+1}(\delta) - a(S)}{a_{k+1}} - f(a(S))$$

Notice that the index  $k$  will change depending on the value of  $\delta$  and that it can be found using the greedy Algorithm 1. This expression is valid when  $A_{k+1} + \delta \geq 0$  and  $A_k + \delta \leq 0$ , i.e.,  $\delta \in [-A_{k+1}, -A_k]$ .  $\square$

*Proof, Proposition 2.* We will show by contradiction.

Suppose that  $(x, w)$  is an optimal solution to  $(L : S)$ . Let  $T := \{i \in N - S : x_i = 1\}$ . The objective value of  $(x, w)$  is

$$z = w - \sum_{i \in N-S} \rho_i(S) x_i - h(S).$$

Since  $(x, w)$  are optimal, the inequality is satisfied with equality, and it holds that

$$g(w) = \sum_{i \in N-S} a_i x_i + a(S) + \delta \iff w = f\left(\sum_{i \in N-S} a_i x_i + a(S) + \delta\right).$$

Substituting the expression for  $w$  into the expression for  $z$  gives

$$\begin{aligned} z &= f\left(\sum_{i \in N-S} a_i x_i + a(S) + \delta\right) - \sum_{i \in N-S} \rho_i(S) x_i - h(S) && \text{(substitution)} \\ &= f(a(T) + a(S) + \delta) - \sum_{i \in N-S} \rho_i(S) x_i - h(S). && \text{(definition of } T) \end{aligned}$$

Next suppose that  $T$  does not satisfy the greedy ordering. Then there must exist some  $j \in T$  such that  $a_i > a_j$  for  $i \notin T$  where the element  $i$  is from  $S$  or  $(N - S) - T$ . Define a new set  $T' := (T \cup \{i\}) - \{j\}$  corresponding to swapping  $j$  for  $i$ . Let  $(x', w'(x'))$  be the point corresponding to  $T'$  so that  $x_i = 0, x'_i = 1$  and  $x_j = 1, x'_j = 0$  and where  $w'(x')$  is the optimal for the given  $x'$ . For notation, let  $w' = w'(x')$ . Then the objective value of  $(x', w')$  is

$$z' = w' - \sum_{i \in N-S} \rho_i(S) x'_i - h(S)$$

where

$$g(w') = \sum_{i \in N-S} a_i x'_i + a(S) + \delta \iff w' \leq f\left(\sum_{i \in N-S} a_i x'_i + a(S) + \delta\right)$$

and

$$\sum_{i \in N-S} a_i x'_i = \sum_{i \in N-S} a_i x_i + a_i - a_j$$

so

$$z' = f\left(\sum_{i \in N-S} a_i x_i + a_i - a_j + a(S) + \delta\right) - \sum_{i \in N-S} \rho_i(S) x'_i - h(S)$$

$$= f(a(T) + a(S) + a_i - a_j + \delta) - \sum_{i \in N-S} \rho_i(S)x_i - \rho_i(S) + \rho_j(S) - h(S).$$

Computing  $z' - z$  gives

$$\begin{aligned} z' - z &= \left[ f(a(T) + a(S) + a_i - a_j + \delta) - \sum_{i \in N-S} \rho_i(S)x_i - \rho_i(S) + \rho_j(S) - h(S) \right] \\ &\quad - \left[ f(a(T) + a(S) + \delta) - \sum_{i \in N-S} \rho_i(S)x_i - h(S) \right] \\ &= f(a(T) + a(S) + a_i - a_j + \delta) - f(a(T) + a(S) + \delta) - \rho_i(S) + \rho_j(S) \\ &\stackrel{(i)}{=} f(a(T) + a(S) + a_i - a_j + \delta) - f(a(T) + a(S) + \delta) - f(a(S) + a_i) + f(a(S) + a_j) \\ &\stackrel{(ii)}{=} \left[ f(a(S) + a_i + a(T) - a_j + \delta) - f(a(S) + a_i) \right] \\ &\quad - \left[ f(a(S) + a_j + a(T) - a_j + \delta) - f(a(S) + a_j) \right]. \end{aligned}$$

where (i) follows from the definition of  $\rho_k(\cdot)$  and (ii) is a regrouping of terms that introduces  $\pm a_j$ . Since  $a_i > a_j$  and  $f$  is concave,  $z' \geq z$  iff  $a(T - \{j\}) + \delta \leq 0$ .

*Claim 1.* The inequality  $a(T - \{i\}) + \delta \leq 0$  holds for all  $i \in T$ .

*Claim 2.* Either  $a(T) + \delta \geq 0$  or  $T = N - S$ .

For claims 1 & 2, see [Ahmed and Atamtürk, 2011]. □