A Note on the Implications of Approximate Submodularity in Discrete Optimization

Temitayo Ajayi[1], Taewoo Lee [*2] and Andrew J. Schaefer[1]

[1]Nature Source Improved Plants

[2]Department of Industrial Engineering, University of Houston

[3]Rice University, Department of Computational and Applied Mathematics

Abstract. Submodularity is a key property in discrete optimization. Submodularity has been widely used for analyzing the greedy algorithm to give performance bounds and providing insight into the construction of valid inequalities for mixed-integer programs. In recent years, researchers started to study approximate submodularity, with a primary focus on providing performance bounds for iterative approaches. In this paper, we study approximate submodularity from a different perspective in order to broaden its use cases in discrete optimization. We define metrics that quantify approximate submodularity, which we then use to derive new properties about both approximate submodularity preservation and the well-known Lovász extension for set functions. We also show that previous analyses of mixed-integer sets, such as the submodular knapsack polytope, can be extended to the approximate submodularity setting. Our work demonstrates that one may generalize many of the analytical tools used in submodular optimization into the approximate submodularity context.

*tlee6@uh.edu

1 Introduction

Exploiting structural properties in discrete optimization problems can lead to successful algorithms and heuristics. A classical property that is frequently used in discrete optimization is submodularity. Let Ω be a finite set of elements, and let 2^{Ω} denote the power set of Ω. A set function $f : 2^{\Omega} \to \mathbb{R}$ is *submodular* if for any $\mathcal{A} \subseteq \mathcal{B} \subseteq \Omega$ and $s \in \Omega \backslash \mathcal{B}$, $f(\mathcal{B} \cup \{s\}) - f(\mathcal{B}) \leq f(\mathcal{A} \cup \{s\}) - f(\mathcal{A})$. For some problems, submodularity provides guarantees for solution approaches such as the greedy algorithm. Recently, researchers have expanded algorithm analysis to *approximately submodular* functions (e.g., [11, 18, 29]). However, much of the initial focus on approximate submodularity has remained within performance guarantees for algorithms. In this paper, we propose approximate submodularity metrics to study multiple implications of approximate submodularity in discrete optimization, including the derivation of valid inequalities and properties of extensions on the unit hypercube. Our work applies to any nonnegative and monotonic set function, and our analyses often follow arguments similar to those of analogous results in the submodular context.

Continuous relaxations of problems are often used as direct approximation techniques for discrete optimization problems (e.g., solving the linear programming relaxation of a mixed-integer program) because they are easier to solve; these relaxations often have a polynomial-time algorithm. Extensions of set functions can transform discrete optimization problems into continuous optimization problems, for which efficient algorithms or approximation schemes may exist. An *extension* of a set function $f : 2^{\Omega} \to \mathbb{R}$ is a function $F : D \to \mathbb{R}$ such that $D \supset \mathbb{B}^{|\Omega|}$ and $F(x(\mathcal{S})) = f(\mathcal{S})$, for all $\mathcal{S} \subseteq \Omega$, where $x(\mathcal{S})$ denotes the characteristic vector of the set $\mathcal{S}$. We focus on extensions defined on the unit hypercube $[0,1]^{|\Omega|}$. Notably, the *Lovász extension* [23] for set functions is convex if and only if the set function is submodular; in this case, the Lovász extension is equal to the convex closure. The convex closure is difficult to compute in general; in contrast, computing the Lovász extension is comparatively simple, which makes submodularity a valuable property when considering solution methods that use the convex closure. We provide a new characterization that relates the approximate submodularity of a function with the approximate convexity of its Lovász extension.

Valid inequalities are crucial for solving mixed-integer programs as they can cut off solutions to relaxations so that the new problem's feasible region more closely approximates the convex hull [8]. The knapsack problem is one of the foundational problems in discrete optimization where researchers have studied its facial structure and valid inequalities (e.g., [3, 6]). In particular, Atamtürk and Narayanan (2009) [3] study valid inequalities for the submodular knapsack polytope, in which the constraint function is submodular. Submodular functions feature in the constraints of other optimization problems as well. Researchers have also studied mixed-integer programs with conic-quadratic constraints and objective functions where valid inequalities are derived by leveraging the submodularity of the objective and constraint functions [1, 2, 13]. Valid inequalities and outer approximations of the epigraphs of submodular and general set functions have

2

also been studied [4]. Our study is the first to use *approximate* submodularity to derive valid inequalities for mixed-integer sets defined by approximately submodular functions.

Previous studies on approximate submodularity focus on performance bounds for greedy algorithms and other iterative selection approaches. Performance bounds have been produced using different notions of approximate submodularity where metrics with different properties can produce different bounds; trade-offs between additive and multiplicative bounds for the greedy algorithm performance on non-submodular functions are studied in [18], and [11] and [29] also define metrics that they use to propose greedy algorithm performance bounds for approximately submodular functions. However, the notion of approximate submodularity can also be used for generalizing results in other areas of discrete optimization, such as analyzing properties of continuous extensions and deriving valid inequalities, which can provide new insights for efficient solution methods. New methodological applications are still emerging, even outside of the greedy algorithm, in which approximate submodularity can extend the existing results that depend on submodularity. Our contributions are as follows:

- In Section 2, we study fundamental properties about our approximate submodularity metrics, which we use to show which operations preserve approximate submodularity.

- In Section 3, we derive results on the approximate convexity of the Lovász extension of approximately submodular functions.

- We study mixed-integer sets defined by approximately submodular functions in Section 4. We use the proposed metrics to adapt analogous analyses for the submodular setting, thus deriving new valid inequalities for cases when the set function is approximately submodular.

We note that there are several cases in which our proofs are similar to those of analogous results in the submodular setting. Our primary message is that in a broad set of areas in discrete optimization, one can use approximate submodularity to generalize both classical and more recent results.

2   Approximate Submodularity Metrics

In this section, we discuss various approximate submodularity metrics; the term "metrics" is used loosely, as some are not subadditive and none are positive definite, both of which are part of the formal definition of a metric. However, the approximate submodularity metrics we discuss indicate a notion of distance to submodularity. In this work, if ζ is an approximate submodularity metric, $\zeta[f]$ is the metric value for f.

2.1   Proposed Notions of Approximate Submodularity

We begin with the most general (global) metric, which is inspired directly from an equivalent definition of submodularity: $f(\mathcal{A} \cup \mathcal{B}) + f(\mathcal{A} \cap \mathcal{B}) \leq f(\mathcal{A}) + f(\mathcal{B})$, for all $\mathcal{A}, \mathcal{B} \subseteq \Omega$.

Definition 1. *Let* $f : 2^\Omega \to \mathbb{R}$. *Define the* global submodularity distance $\mathcal{E}$ *by* $\mathcal{E}[f] := \max_{\mathcal{A},\mathcal{B}\subseteq\Omega} f(\mathcal{A}\cup\mathcal{B}) + f(\mathcal{A}\cap\mathcal{B}) - f(\mathcal{A}) - f(\mathcal{B})$.

The global submodularity distance is a general purpose metric; we demonstrate its value in identifying operations that preserve approximate submodularity and proving general results about set functions (Section 2.2). The remaining metrics are inspired by a characterization of increasing, submodular functions.

Lemma 1. (Edmonds 1970 [12]) *Let* $f : 2^\Omega \to \mathbb{R}$. *Then* f *is increasing and submodular if and only if for any* $\mathcal{A},\mathcal{B}\subseteq\Omega, s\in\Omega$, $f(\mathcal{A}\cup\mathcal{B}\cup\{s\}) - f(\mathcal{A}\cup\mathcal{B}) \le f(\mathcal{A}\cup\{s\}) - f(\mathcal{A})$.

Using Lemma 1, we present two metrics for approximate submodularity.

Definition 2. *Let* $f : 2^\Omega \to \mathbb{R}$. *Consider* $\ell \in \{0,...,|\Omega|-1\}, k \in \{0,...,|\Omega|\}$. *The* (ℓ,k)-pairwise violation *of* f *is defined as* $d^{\ell,k}[f] := \max_{\substack{\mathcal{A},\mathcal{B}\subseteq\Omega,s\in\Omega \\ |\mathcal{A}|=\ell,|\mathcal{B}|=k}} f(\mathcal{A}\cup\mathcal{B}\cup\{s\}) - f(\mathcal{A}\cup\mathcal{B}) - f(\mathcal{A}\cup\{s\}) + f(\mathcal{A})$.

Thus, the pairwise violation represents the worst-case violation of the condition in Lemma 1 given $\mathcal{A}$ and $\mathcal{B}$ with fixed cardinalities. In the context of a sensor placement problem, the (ℓ,k)-pairwise violation captures the case in which a single sensor added to a sparse sensor network (given by $\mathcal{A}$) creates a smaller marginal increase in information than when the same sensor is added to a denser network ($\mathcal{A}\cup\mathcal{B}$). Note that if f is submodular, $d^{\ell,k}[f] \le 0$ for all ℓ and k, and the reverse implication holds with the added condition of f being monotonic increasing. Also, for any f, $d^{\ell,k}[f] \le \mathcal{E}[f]$.

Definition 3. *Let* $f : 2^\Omega \to \mathbb{R}$. *The* marginal violation *of* f *is defined as* $D[f] := \max\{d^{\ell,k}[f] \mid \ell \in \{0,\dots,|\Omega|-1\}, k \in \{0,\dots,|\Omega|\}\}$.

Note that $D[f]$, also used in [21], does not depend on set sizes. Although $\mathcal{E}[f]$ and $D[f]$ may be difficult to compute exactly in general, Section 2.2 details some operations that preserve approximate submodularity, with respect to $\mathcal{E}[f]$, which enables one to bound $\mathcal{E}[f]$ and $D[f]$. In addition, in Appendix B, we present a generalized version of the uncapacitated facility location problem in which the objective function is approximately submodular and $D[f]$ can be bounded analytically.

2.2 Preserving Approximate Submodularity

We prove some properties of our proposed approximate submodularity metrics (Theorem 1), as well as operations from which bounds or exact values of approximate submodularity metrics can be inferred immediately (Proposition 1). The former compares properties of our approximate submodularity metrics to true metrics. The latter concept can be thought of as "approximate submodularity preservation." Some of these results have analogs for submodular functions (for reference, see [5], [26], and [27]), but others are specific to approximate submodularity. We let $\mathcal{F}$ (resp., $\mathcal{F}_+$) be the set functions (resp., that are nonnegative and increasing) over ground set Ω.

4

Theorem 1. *Consider a nonnegative, increasing set function $f : 2^\Omega \to \mathbb{R}$ and a metric of approximate submodularity $\zeta : \mathcal{F}_+ \to \mathbb{R}$ where ζ is defined by any of the following: (I) $\zeta[f] = \mathcal{E}[f]$, (II) $\zeta[f] = D[f]$, or (III) $\zeta[f] = d^{\ell,k}[f]$, for some $\ell \in \{0, \dots, |\Omega| - 1\}, k \in \{0, \dots, |\Omega|\}$. Then we have:*

  (i) *The function ζ is sublinear. That is, ζ is subadditive (i.e., $\zeta[f_1] + \zeta[f_2] \geq \zeta[f_1 + f_2]$) and positively homogeneous with degree 1 (i.e., $\alpha\zeta[f] = \zeta[\alpha f]$, for $\alpha \in \mathbb{R}_+$).*

  (ii) *If f is not submodular, then for any $\epsilon \in [0, \zeta[f])$, there does not exist a nonnegative, increasing, submodular function $g : 2^\Omega \to \mathbb{R}$ such that $\|g - f\|_\infty < \frac{\epsilon}{4}$.*

The contrapositive of Claim (ii) of Theorem 1 can be read as a necessary condition, which can, in some cases, remove the need for testing whether any function near f is submodular (e.g., [28]). Although our notions of approximate submodularity are not "metrics" in the analytical sense, Theorem 1 proves that they are sublinear. Sublinear functions are well studied in the literature and are the "next simplest convex functions" after affine functions [17]. All metrics are sublinear. Subadditivity and positive homogeneity independently have multiple implications. They can be used to verify that a function $f_1 + f_2$ (or αf, for $\alpha \in \mathbb{R}_+$) satisfies conditions in hypotheses of results in Sections 3–4. We remark that subadditivity (and hence, sublinearity) is not a trivial property of approximate submodularity metrics in the literature; e.g., submodularity ratio proposed in Das and Kempe (2011) [11] is not subadditive.

We can relate our metrics to asymmetric seminorms, which share more properties with analytical metrics.

Definition 4. (Cobzaş 2013 [7]) *A function $\zeta : \mathbb{R}^n \to \mathbb{R}$ is an asymmetric seminorm if it is nonnegative, positively homogeneous, and subadditive.*

Corollary 1. *Define $\mathcal{E}_+ : \mathcal{F} \to \mathbb{R}$ by $\mathcal{E}_+[f] := \max\{0, \mathcal{E}[f]\}$. Then $\mathcal{E}_+$ is an asymmetric seminorm on $\mathcal{F}$.*

We provide some examples in which functions induced by an approximately submodular function inherit approximate submodularity. Denote the complement of $\mathcal{S} \subseteq \Omega$ by $\mathcal{S}^c$. Given a normalized set function f, define $f_1, f_2 : 2^\Omega \to \mathbb{R}$ by $f_1(\mathcal{S}) = f(\mathcal{S}^c), f_2(\mathcal{S}) = f(\mathcal{S}) + f(\mathcal{S}^c) - f(\Omega)$; thus, f_2 is a symmetric, nonnegative function. Given $\mathcal{A} \subseteq \Omega$, define $f_\mathcal{A} : 2^{\Omega \backslash \mathcal{A}} \to \mathbb{R}$ by $f_\mathcal{A}(\mathcal{S}) = f(\mathcal{A} \cup \mathcal{S})$. Given a factor q of $|\Omega|$, let $\Omega(q) = \{1, \dots, \frac{|\Omega|}{q}\}$, $\mathcal{S}(i) = \{(i-1)q+1, \dots, iq\}$, for all $i \in \Omega(q)$, and $f_q : 2^{\Omega(q)} \to \mathbb{R}$ be defined by $f_q(\mathcal{S}) = f(\bigcup_{i \in \mathcal{S}} \mathcal{S}(i))$. Finally, let $g : 2^\Omega \to \mathbb{R}$ be a *modular* function (g and $-g$ are submodular), and define the convolution of f and g as $f \circledast g(\mathcal{S}) = \min_{\mathcal{Z} \subseteq \mathcal{S}} f(\mathcal{Z}) + g(\mathcal{S} \backslash \mathcal{Z})$. Note $f \circledast g(\mathcal{S}) = g \circledast f(\mathcal{S})$.

Proposition 1. *Given $f : 2^\Omega \to \mathbb{R}$ and the corresponding functions $f_1, f_2, f_\mathcal{A}$, and f_q, we have: (i) $\mathcal{E}[f] = \mathcal{E}[f_1]$. (ii) $2\mathcal{E}[f] \geq \mathcal{E}[f_2]$. (iii) $\mathcal{E}[f] \geq \mathcal{E}[f_\mathcal{A}]$ (iv) $\mathcal{E}[f] \geq \mathcal{E}[f_q]$. (v) $\mathcal{E}[f] \geq \mathcal{E}[f \circledast g]$.*

Proposition 1 can be used in a fashion similar to Theorem 1. Also, Proposition 1(iv) can provide guarantees on a greedy algorithm that selects among prescribed subsets of elements. Our proof for Proposition 1(v)

follows similar arguments to that of the submodular case [26]. Note that we slightly abuse notation in Proposition 1(iii)–(iv) as the domains of $f_{\mathcal{A}}$ and f_q are not 2^{Ω}.

3 Extensions of Approximately Submodular Functions

Next, we study extensions of approximately submodular functions. For the remainder of this paper, we consider functions that are monotone increasing and normalized ($f(\emptyset) = 0$) and focus on analysis based on the marginal violation D because we consider problems of this form, namely the approximately submodular knapsack and packing problems and the generalized uncapacitated facility location problem, in Section 4 and Appendix B, respectively. Other examples for monotone submodular optimization can be found in [22], which may have relevant approximately submodular analogs. Given a set function $f : 2^{\Omega} \to \mathbb{R}$, an *extension* of f over $[0,1]^{|\Omega|}$ is a function $F : [0,1]^{|\Omega|} \to \mathbb{R}$ such that $F(x(\mathcal{S})) = f(\mathcal{S})$, for all $\mathcal{S} \subseteq \Omega$, where $x(\mathcal{S})$ is the characteristic vector of $\mathcal{S}$. Our main result in this section is that the Lovász extension is approximately convex (Theorem 2) when the underlying set function is approximately submodular. A main component of multiple key results in this section is the marginal violation D (Definition 3). Other works on extensions of set functions include [19], [23], and [25].

The Lovász extension of a set function $f : 2^{\Omega} \to \mathbb{R}$ is defined by $F^L : [0,1]^{|\Omega|} \to \mathbb{R}$ such that $F^L(x) :=$ $\sum_{k=0}^{|\Omega|} \lambda_k f(\mathcal{C}_k)$, where $\emptyset = \mathcal{C}_0 \subset \mathcal{C}_1 \subset \cdots \subset \mathcal{C}_{|\Omega|} = \Omega$ is a chain such that $\sum_{k=0}^{|\Omega|} \lambda_k x(\mathcal{C}_k) = x$, with $\sum_{k=0}^{|\Omega|} \lambda_k = 1$, and $\lambda \geq 0$. It is well known that by defining a permutation $(\pi_1, \ldots, \pi_{|\Omega|})$ such that $x_{\pi_1} \geq x_{\pi_2} \cdots \geq x_{\pi_{|\Omega|}}$, $\mathcal{C}_0 = \emptyset, \mathcal{C}_k = \mathcal{C}_{k-1} \cup \{\pi_k\}$, for $k \in \{1, \ldots, |\Omega|\}$, the Lovász extension is equivalently defined as $F^L(x) = \sum_{k=1}^{|\Omega|} x_{\pi_k}(f(x(\mathcal{C}_k)) - f(x(\mathcal{C}_{k-1})))$ (see [5]). The convex closure of f is the unique convex function $F^C : [0,1]^{|\Omega|} \to \mathbb{R}$ such that $F^C(x(\mathcal{S})) \leq f(\mathcal{S})$ for all $\mathcal{S} \subseteq \Omega$ and $F^C(x) \geq G(x)$ for any other convex understimator $G : [0,1]^{|\Omega|} \to \mathbb{R}$ of f. Lovász (1983) [23] shows that f is submodular if and only if F^L is convex; in fact, in this special case, the convex closure and the Lovász extension are equal ($F^C = F^L$). This property is useful in that the convex closure is generally difficult to compute in comparison to the Lovász extension. Although the Lovász extension does not equal the convex closure when f is not submodular, we prove a generalized result when f is approximately submodular. We remark that Halabi and Jegelka (2019) [15] also study the Lovász extension of non-submodular functions, including its subgradients, in the context of convex optimization solution approaches.

Consider the following linear program parametrized by $x \in [0,1]^{|\Omega|}$:

$$V(x) = \min_{y} \left\{ \sum_{\mathcal{S} \subseteq \Omega} f(\mathcal{S})y(\mathcal{S}) \ \middle|\ \sum_{\mathcal{S} \ni s} y(\mathcal{S}) = x_s, \forall\, s \in \Omega, \sum_{\mathcal{S} \subseteq \Omega} y(\mathcal{S}) = 1, y \geq 0 \right\}. \tag{1}$$

Proposition 2. (Bach 2013 [5]) *For $f : 2^{\Omega} \to \mathbb{R}$ with $f(\emptyset) = 0$, we have $V(x) = F^C(x)$, for all $x \in [0,1]^{|\Omega|}$.*

Given a permutation $(\pi_1, \cdots, \pi_{|\Omega|})$, define $\mathcal{S}_0^\pi := \emptyset \subset \mathcal{S}_1^\pi := \{\pi_1\} \cdots \mathcal{S}_k^\pi := \{\pi_1, \ldots, \pi_k\} \cdots \subset \mathcal{S}_{|\Omega|}^\pi := \Omega$.
Define the set $\Gamma(f) := \{\gamma \in \mathbb{R}^{|\Omega|} \mid \exists \text{ permutation } \pi \text{ such that } \gamma_{\pi_i} = f(\mathcal{S}_i^\pi) - f(\mathcal{S}_{i-1}^\pi), \forall \, i \in \Omega\}$.

Definition 5. *A function $F : [0,1]^{|\Omega|} \to \mathbb{R}$ is ϵ-approximately convex, if $F(\lambda x + (1-\lambda)y) \leq \epsilon + \lambda F(x) + (1-\lambda)F(y)$, for any $\lambda \in [0,1]$.*

Theorem 2. *For any increasing set function $f : 2^\Omega \to \mathbb{R}$ such that $f(\emptyset) = 0$,*

$$F^L(x) \leq \max_{\gamma \in \Gamma(f)} \sum_{s \in \Omega} \gamma_s x_s \leq F^C(x) + |\Omega|D[f] \leq F^L(x) + |\Omega|D[f] \leq \max_{\gamma \in \Gamma(f)} \sum_{s \in \Omega} \gamma_s x_s + |\Omega|D[f].$$

Hence, $F^L(x) \geq F^C(x) \geq F^L(x) - |\Omega|D[f]$, and $||F^L - F^C||_\infty \leq |\Omega|D[f]$. Moreover, F^L is $|\Omega|D[f]$-approximately convex. In addition, if for some $\epsilon > 0$, F^L is ϵ-approximately convex, then $D[f] \leq \epsilon$.

Theorem 2 states that the approximate submodularity of f implies the approximate convexity of F^L and vice-versa. The proof of Theorem 2 uses the well-known linear program (1), but a key difference is that we construct feasible primal-dual solutions with a duality gap due to the generalization to approximate submodularity.

Next, we consider the case in which there exists a submodular function g close to f. In this case, we show that the Lovász extension of g approximates the Lovász extension of f.

Proposition 3. *Given set functions $f, g : 2^\Omega \to \mathbb{R}$, where $f(\emptyset) = g(\emptyset) = 0$, and their respective Lovász extensions $F^L, G^L : [0,1]^{|\Omega|} \to \mathbb{R}$, $||F^L - G^L||_\infty = ||f - g||_\infty$.*

Thus, the approximating function g (which may be submodular) can lead to approximation methods in the discrete domain or over the hypercube using convex optimization methods.

4   Valid Inequalities of Polyhedra Associated With Approximately Submodular Functions

We use approximate submodularity metrics from Section 2.1 to derive valid inequalities for some mixed-integer sets. Our analyses are similar to analogs in submodular analysis [2, 3, 4] with additional details to generalize to approximate submodularity.

4.1   Epigraph Inequalities

First, we study the epigraphs of set functions. These mixed-integer sets can be useful when minimizing a submodular function [1]. We consider the case when the function is approximately submodular. Let $\phi : \mathbb{R} \to \mathbb{R}_+$ be increasing, and for any $\tau \in \mathbb{R}_+$, let $F_\tau : [0,1]^{|\Omega|} \to \mathbb{R}_+$ be defined by $F_\tau(x) =$

$\phi(\tau + \sum_{i\in\Omega} c_i x_i)$, where $c \in \mathbb{R}_+^{|\Omega|}$. Thus, F_τ is increasing. Consider the mixed-integer feasible region $H_\mathbb{B} = \{(x,z) \in \mathbb{B}^{|\Omega|} \times \mathbb{R}_+ \mid F_\sigma(x) \leq z\}$, where $\sigma \geq 0$, and $c \in \mathbb{R}_+^{|\Omega|}$. Define the set function $f_\tau : 2^\Omega \to \mathbb{R}$ by $f_\tau(\mathcal{S}) = F_\tau(x(\mathcal{S}))$. Notice that f_τ is increasing and $f_\tau(\emptyset) = 0$ if and only if $\phi(\tau) = 0$. Therefore, define $g_\tau : 2^\Omega \to \mathbb{R}$ by $g_\tau(\mathcal{S}) = f_\tau(\mathcal{S}) - \phi(\tau)$, which is normalized, $g_\tau(\emptyset) = 0$, and is increasing; hence, it is also nonnegative. Note that $D[f_\tau] = D[g_\tau]$. We denote the Lovász extension of g_τ by G_τ^L.

For any $\gamma \in \Gamma(g_\tau)$—i.e., $\gamma_{\pi_k} = f_\tau(\mathcal{S}_k^\pi) - f_\tau(\mathcal{S}_{k-1}^\pi) = \phi(\tau + \sum_{i=1}^k c_{\pi_i}) - \phi(\tau + \sum_{i=1}^k c_{\pi_{i-1}})$ for some permutation π of $(1,\ldots,|\Omega|)$—consider the following inequality:

$$\sum_{s\in\Omega} \gamma_s x_s \leq z - \phi(\tau). \tag{2}$$

When ϕ is the square root function, then f_τ is a submodular set function, and Atamtürk and Narayanan (2008) [2] show that inequality (2) is valid for the convex hull $\mathrm{conv}(H_\mathbb{B})$ for $\tau = \sigma$. In fact, along with the variable bounds, such inequalities describe $\mathrm{conv}(H_\mathbb{B})$. In the more general case, where ϕ is such that f_τ is approximately submodular, we show that similar inequalities are still valid for $\mathrm{conv}(H_\mathbb{B})$.

Lemma 2. *For any $\gamma \in \Gamma(g_\tau), \mathcal{S} \subseteq \Omega$, we have $-|\Omega|D[f_\tau] + \sum_{s\in\Omega} \gamma_s (x(\mathcal{S}))_s \leq f_\tau(\mathcal{S}) - \phi(\tau)$.*

Proposition 4. *For any $\gamma \in \Gamma(g_\sigma)$, the following inequality is valid for $\mathrm{conv}(H_\mathbb{B})$:*

$$-|\Omega|D[f_\sigma] + \sum_{s\in\Omega} \gamma_s x_s \leq z - \phi(\sigma). \tag{3}$$

Proposition 4 illustrates what is lost between submodularity and approximate submodularity in deriving valid inequalities in this setting. When f_σ and g_σ are approximately submodular, $D[f_\tau] > 0$ may lead to looser valid inequalities. Our proof of Proposition 4 follows arguments similar to those of Atamtürk and Narayanan (2008) [2], who establish the result when ϕ is the square root function.

Next, we consider the epigraph of a general, increasing, nonnegative, approximately submodular function f, $H_f := \mathrm{conv}(\{(x,z) \in \mathbb{R}^{|\Omega|} \times \mathbb{R} \mid f(x(\mathcal{S})) \leq z\})$. Consider the associated polyhedron $P_f := \{\gamma \in \mathbb{R}^{|\Omega|} \mid \sum_{s\in\Omega} \gamma_s \leq f(\mathcal{S}), \forall\, \mathcal{S} \subseteq \Omega\}$. We refer to the variable bounds as trivial inequalities of H_f.

Proposition 5. (Atamtürk and Narayanan 2020 [4])

1. *Any nontrivial facet-defining inequality $\sum_{s\in\Omega} \gamma_s x_s \leq \alpha z + \gamma_0$ for H_f satisfies $\gamma_0 \geq 0$ and $\alpha = 1$ (up to scaling).*

2. *The inequality $\sum_{s\in\Omega} \gamma_s x_s \leq z$ is valid for H_f if and only if $\gamma \in P_f$.*

3. *The inequality $\sum_{s\in\Omega} \gamma_s x_s \leq z$ is facet-defining for H_f if and only if γ is an extreme point of P_f.*

Atamtürk and Narayanan (2020) [4] prove that nontrivial facets of H_f are homogeneous. We establish a similar result for approximately submodular functions.

Proposition 6. *Let $f : 2^\Omega \to \mathbb{R}$ be increasing with $f(\emptyset) = 0$. Suppose $\gamma \in \mathbb{R}^{|\Omega|}$ and*

$$\sum_{s \in \Omega} \gamma_s x_s \leq z + |\Omega| D[f] + \gamma_0 \tag{4}$$

defines a nontrivial facet of H_f. Let $\bar{f} : 2^\Omega \to \mathbb{R}$ be defined by $\bar{f}(\emptyset) = 0, \bar{f}(\mathcal{S}) = f(\mathcal{S}) + |\Omega| D[f] + \gamma_0$, for all nonempty $\mathcal{S} \subseteq \Omega$, and suppose $\gamma \in \Gamma(\bar{f})$. Then, $\gamma_0 \leq 0$.

The proof of Proposition 6 proceeds similarly to that of the submodular case in [4], with some additional steps to account for the approximate submodularity generalization. This includes bounding $\max_{\gamma \in \Gamma[f]} \sum_{s \in \mathcal{S}} \gamma_s$ using the marginal violation D. Given the conditions in the hypothesis of Proposition 6, the constant term $|\Omega| D[f] + \gamma_0$ is bounded below by 0 and above by $|\Omega| D[f]$; when f is submodular, the condition $\gamma \in \Gamma(\bar{f})$ is implied, $D[\bar{f}] = 0$, and the nontrivial facets are homogeneous. We also remark that Atamtürk and Narayanan (2020) [4] provide valid inequalties for general set functions, but these rely on a submodular-supermodular decomposition of f.

4.2 Knapsack Inequalities

Consider the polytope $X = \text{conv}\{x \in \mathbb{B}^n \mid f(\mathcal{S}(x)) \leq b\}$, where $\mathcal{S}(x)$ is the subset of Ω characterized by the binary vector x and f is a set function. When f is submodular, nonnegative, and increasing, X is known as the submodular knapsack polytope [3]; a special case is the well-known linear knapsack set, and optimizing over it is NP-hard [20]. We consider the case where f is approximately submodular, nonnegative and increasing. Thus, we call the set X an *approximately submodular knapsack set*. Our focus in this subsection is on deriving valid inequalities for this set. Some facets for 0-1 polytopes established by Atamtürk and Narayanan (2009) [3] apply in our setting; we list them in Proposition $A.2$ in the appendix.

Definition 6.

1. *The subset $\mathcal{S} \subseteq \Omega$ is a* cover *for X if $f(\mathcal{S}) > b$ and is* minimal *if $f(\mathcal{S}\backslash\{s\}) \leq b$ for all $s \in \mathcal{S}$.*

2. *Let $\pi = (\pi_1, \ldots, \pi_{|\Omega\backslash\mathcal{S}|})$ be a permutation of $\Omega\backslash\mathcal{S}$. Let $U_\pi(\mathcal{S}) = \{\pi_j \in \Omega\backslash\mathcal{S} \mid f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_j\}) - f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_{j-1}\}) \geq f(\{s\}), \forall s \in \mathcal{S}\}$. The* set-extension *of $\mathcal{S} \subseteq \Omega$ with respect to π is denoted by $E_\pi(\mathcal{S}) = \mathcal{S} \cup U_\pi(\mathcal{S})$.*

Proposition 7 extends the result of Proposition 5 in Atamtürk and Narayanan (2009) [3] for submodular knapsack problems into the approximately submodular context.

Proposition 7. *If $\mathcal{S} \subseteq \Omega$ is a cover for X, the extended cover inequality $\sum\limits_{s \in E_\pi(\mathcal{S})} x_s \leq |\mathcal{S}| - 1$ is valid for X if $f(\mathcal{S}) > (|\mathcal{S}| + |U_\pi(\mathcal{S})|)D[f] + b$. In addition, the inequality defines a facet of $\{x \in X \mid x_s = 0, \forall\, s \notin E_\pi(\mathcal{S})\}$ if $\mathcal{S}$ is also a minimal cover and for each $s \in U_\pi(\mathcal{S})$, there exist $t_s, u_s \in \mathcal{S}$ such that $t_s \neq u_s$, and $f(\mathcal{S} \cup \{s\}\backslash\{t_s, u_s\}) \leq b$.*

We observe from Proposition 7 that in adapting the result for approximate submodularity, we add a condition for the extended cover inequality to be valid. The proof follows similar steps as those in Atamtürk and Narayanan (2009) [3], except it accounts for violated submodularity inequalities.

4.3 Illustrative Example

In this section we provide an example in which we derive valid inequalities for the approximately submodular knapsack polytope. We apply the derived valid inequalities (Proposition 7) and show that they can be used as a tool in the process of finding integer solutions when optimizing a linear function over the polytope.

Let Ω be a set of elements, $u, w \in \mathbb{R}_+^{|\Omega|}\backslash\{0\}$, $p > 1$. Let $G : [0,1]^{|\Omega|} \to \mathbb{R}$ and $H : \mathbb{R} \to \mathbb{R}$, where $G(x) = w^\top x$ and $H(z) = z^p$ if $z \geq 0$ and 0 otherwise. Define $F : [0,1]^{|\Omega|} \to \mathbb{R}$ and $f : 2^\Omega \to \mathbb{R}$ by $F(x) = u^\top x + H(G(x))$ and $f(\mathcal{S}) = F(x(\mathcal{S}))$. Because $p > 1$, H is convex and increasing. Also, G is a linear function, so F is convex and f is supermodular. Moreover, f is increasing and nonnegative but not submodular.

Proposition 8. *We have $D[f] \leq p\|w\|_1^{p-1}\|w\|_\infty$.*

Proposition 8 gives a bound on $D[f]$ that can remove the need to compute $D[f]$ directly, which helps verify whether Proposition 7 applies to inequalities of the form $f(\mathcal{S}) > (|\mathcal{S}| + |U_\pi(\mathcal{S})|)D[f] + b$, for some $\mathcal{S} \subseteq \Omega$.

We provide an example instance in which we optimize a linear function over the integer hull of the approximately submodular knapsack polytope. In particular, we show that by adding our valid inequalities, it is possible to obtain an integral solution from the continuous relaxation. Let $\Omega = \{1, 2, \ldots, 6\}$, $u = [9, 9, 9, 9, 8.85, 0]^\top$, $w = [0, 0, 0, 0, 1, 1]^\top$, $p = 1.1$, and define f, F, G, and H as stated above. Also let $c = [3, 3, 3, 3, 2, 2]^\top, b = 28.3$. Define the following instance of an approximately submodular knapsack problem (ASK) written as a binary program:

$$z_{\text{ASK}} = \max\left\{\sum_{s \in \Omega} c_s x_s \mid F(x) \leq b, x \in \mathbb{B}^{|\Omega|}\right\}. \tag{ASK}$$

The continuous relaxation of (ASK) was solved using Gurobi 9.1.1 [14] through the Gurobipy python interface (Python version 3.6.8) using a piecewise approximation of the nonlinear function with maximum

absolute error of .001, with an optimal solution of $[0.0\overline{33}, 1, 1, 1, 0, 1]^\top$ and objective value 11.1. Observe that $\mathcal{S} = \{1, 2, 3, 4\}$ is a (minimal) cover. Consider the permutation of $(5, 6)$ ($\pi_1 = 5, \pi_2 = 6$). Then $f(\mathcal{S} \cup \{\pi_1\}) - f(\mathcal{S}) = 9.85 > 9 = f(\{s\})$, and $f(\mathcal{S} \cup \{\pi_1, \pi_2\}) - f(\mathcal{S} \cup \{\pi_1\}) \approx 1.14 < f(\{s\})$, for all $s \in \mathcal{S}$. Hence, $U_\pi(\mathcal{S}) = \{5\}$ and $E_\pi(\mathcal{S}) = \{1, 2, \ldots, 5\}$. By Proposition 8, $D[f] \leq p||w||_1^{p-1}||w||_\infty \approx 1.18$, which implies $f(\mathcal{S}) = 36 > 1.18|\Omega| + 28.3 = 7.08 + 28.3 \geq |\Omega|D[f] + b \geq (|\mathcal{S}| + |U_\pi(\mathcal{S})|)D[f] + b$; thus, Proposition 7 implies $\sum_{s=1}^{5} x_s \leq 3$ is a valid inequality for (ASK). Solving the relaxation of (ASK) with this valid inequality yields an optimal solution of $[1, 1, 1, 0, 0, 1]^\top$ with an objective value of 11. Thus, this solution is optimal for (ASK).

In general, $f(\mathcal{S}) > (|\mathcal{S}| + |U_\pi(\mathcal{S})|)D[f] + b$ does not hold, so not every extended cover inequality is valid.

4.4   Randomly Generated Instances

We further explore the utility of the valid inequalities presented in Proposition 7 by solving 40 randomly generated instances with and without a selection of these valid inequalities. The instances are larger versions of the example instance illustrated in Section 4.3; each instance contains $q = 10$ approximately submodular knapsack constraints, and we refer to this problem as an *approximately submodular packing problem*. Table 1 shows the lower and upper bounds of the uniform distributions of the parameters of the problem instances.

Because there were $n = 120$ decision variables, it was not practical to examine all possible valid inequalities for each instance. Instead, for each knapsack constraint, we generated 90 sets of $\mathcal{S}$ where $|\mathcal{S}| \in \{5, 6, 7, 8\}$ and searched for valid inequalities based on these sets. In particular, for five repetitions, we randomly selected a subset $\mathcal{T} \subset \mathcal{S}^c$ where $|\mathcal{T}| = 8$ and determined if an extended cover valid inequality could be generated. To aid the search for valid inequalities near the best feasible solutions, the probability that variable i was added to set $\mathcal{S}$ was given by $\frac{c_i}{\sum_j c_j}$. All generated inequalities were added to the formulation before the start of the solve. We note that optimizing the inequality generation process is outside the scope of this study; thus, we do not include the generation time in our results and leave this subject to future research. All problems were solved using Gurobi 9.1.1 [14] through the Gurobipy Python interface (Python version 3.6.8) on a Linux machine with a 20-core 2.4 GHz processor and 256 GB memory.

Overall, we observe that the extended cover valid inequalities improve the solution time for most of the instances (see Figure 1). The cumulative time to solve all of the instances with the extended cover inequalities was less than half that without the additional inequalities. Table C.1 in the appendix provides the solution times and the number of added inequalities for each instance.

Parameter	LB	UB	Set Value
w_i	.2	.6	·
u_i	1	11	·
c_i	8.5	11.5	·
b	42	48	·
p	1.05	1.2	·
q	·	·	10
n	·	·	120

Table 1: Approximately submodular packing problem instance parameter values for w, u, c, b, and p were randomly generated from uniform distributions with the above lower and upper bounds. All problem instances had exactly 120 decision variables.
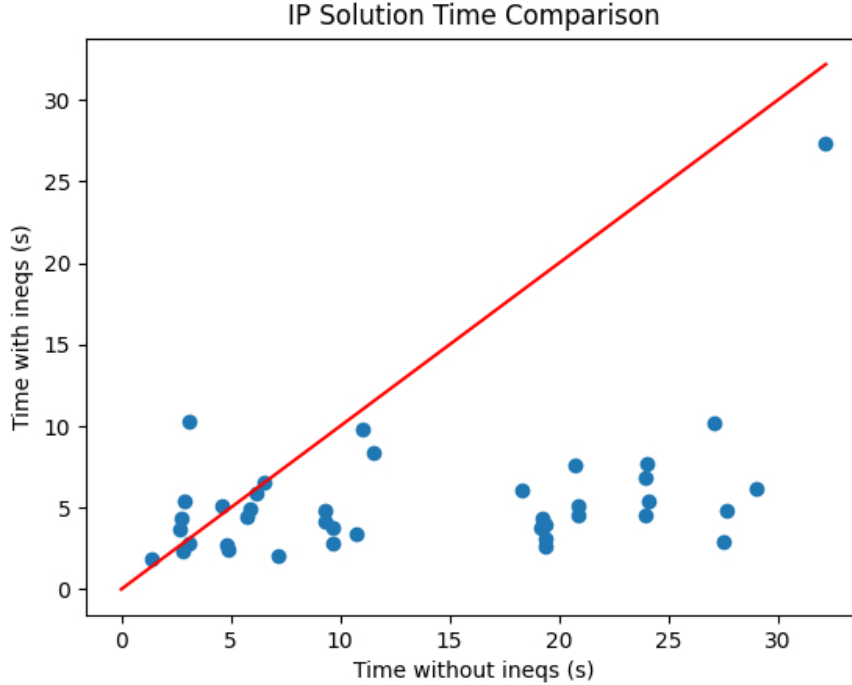

Figure 1: Solution times (with and without inequalities) in seconds for 40 instances of approximately submodular packing problems. Points below the diagonal line (where the time ratio equals 1) indicate reduced solution times by adding extended cover inequalities.

5 Conclusion

The value of submodularity in discrete optimization has long been established. Recently, notions of approximate submodularity have been applied to the greedy algorithm and similar approaches. We introduce new approximate submodularity metrics that have broad applicability in discrete optimization. We derive fundamental properties about our metrics, including which set function operations preserve approximate submodularity. We establish connections between our notions of approximate submodularity and the approximate convexity of the Lovász extension. Our approximate submodularity metrics can directly extend analyses in areas such as valid inequality derivations. Our numerical results show that valid inequalities derived based on the proposed metrics reduce solution times for approximately submodular packing problems in general. Optimizing the generation of valid inequalities for these problems and comparing their performances to prior works remains our future work.

Data availability statement

The implementable instances used in this study are available at:

`https://bitbucket.org/tayoajayi/approxsubmodinstances2021/src/main/.`

References

[1] A. Atamtürk and A. Gómez. Submodularity in conic quadratic mixed 0-1 optimization. *Operations Research*, 68(2):609–630, 2020.

[2] A. Atamtürk and V. Narayanan. Polymatroids and mean-risk minimization in discrete optimization. *Operations Research Letters*, 36(5):618–622, 2008.

[3] A. Atamtürk and V. Narayanan. The submodular knapsack polytope. *Discrete Optimization*, 6(4):333–344, 2009.

[4] A. Atamtürk and V. Narayanan. Submodular function minimization and polarity. *Mathematical Programming Series B*, 2021.

[5] F. Bach. *Learning with Submodular Functions: A Convex Optimization Perspective*. Now Publishers Inc., Hanover, MA, USA, 2013.

[6] E. Balas. Facets of the knapsack polytope. *Mathematical Programming*, 8:146–164, 1975.

[7] Ş. Cobzaş. *Functional Analysis in Asymmetric Normed Spaces*. Birkhäuser, Basel, Basel, 2013.

[8] G. Cornuéjols. Valid inequalities for mixed integer linear programs. *Mathematical Programming*, 112:3–44, 2008.

[9] G. Cornuéjols, M. L. Fisher, and G. L. Nemhauser. Location of bank accounts to optimize float: An analytic study of exact and approximate algorithms. *Management Science*, 23(8):789–810, 1977.

[10] G. Cornuéjols, G. Nemhauser, and L. Wolsey. The uncapacitated facility location problem. In P. E. Mirchandani and R. L. Francis, editors, *Discrete Location Theory*, pages 119–171. John Wiley and Sons Inc., New York, NY, 1983.

[11] A. Das and D. Kempe. Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, pages 1057–1064, USA, 2011. Omnipress.

[12] J. Edmonds. Submodular functions, matroids, and certain polyhedra. In *Combinatorial Structures and Their Applications (Proceedings of Calgary International Conference on Combinatorial Structures and Their Applications)*, pages 69–87, New York, NY, USA, 1970. Gordon and Breach.

[13] A. Gómez. Strong formulations for conic quadratic optimization with indicator variables. *Mathematical Programming*, 188:193–226, 2021.

[14] Gurobi Optimization LLC. Gurobi optimizer reference manual, 2020.

[15] Marwa El Halabi and Stefanie Jegelka. Optimal approximation for unconstrained non-submodular minimization. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 3961–3972. PMLR, 13–18 Jul 2020.

[16] P. L. Hammer, E. L. Johnson, and U. N. Peled. Facet of regular 0-1 polytopes. *Mathematical Programming*, 8:179–206, 1975.

[17] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis*. Springer-Verlag, Berlin, Heidelberg, 2001.

[18] T. Horel and Y. Singer. Maximization of approximately submodular functions. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 3045–3053. Curran Associates, Inc., 2016.

[19] R. Iyer and J. Bilmes. Polyhedral aspects of submodularity, convexity and concavity, 2015. arXiv 1506.07329.

[20] R. M. Karp. Reducibility among combinatorial problems. In R.E. Miller, J.W. Thatcher, and J.D. Bohlinger, editors, *Complexity of Computer Computations*, pages 85–103. Springer US, 1972.

[21] A. Krause and V. Cevher. Submodular dictionary selection for sparse representation. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. Omnipress, 2010.

[22] A. Krause and D. Golovin. *Submodular Function Maximization*, pages 71–104. Cambridge University Press, 2014.

[23] L. Lovász. Submodular functions and convexity. In *Mathematical Programming: The State of the Art*, pages 235–257. Springer, 1983.

[24] P. B. Mirchandani and R. L. Francis, editors. *Discrete Location Theory*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, 1st edition, 1990.

[25] K. Murota. Discrete convex analysis. *Mathematical Programming*, 83:313–371, 1998.

[26] H. Narayanan. *Submodular Functions and Electrical Networks*, volume 54 of *Annals of Discrete Mathematics*. North Holland, 1997.

[27] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization*. Wiley-Interscience. Wiley, 1988.

[28] C. Seshadri and J. Vondrák. Is submodularity testable. *Algorithmica*, 69:1–25, 2014.

[29] Y. Zhou and C. J. Spanos. Causal meets submodular: Subset selection with directed information. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 2649–2657. Curran Associates, Inc., 2016.

A Omitted Proofs

THEOREM 1. *Consider a nonnegative, increasing set function $f : 2^\Omega \to \mathbb{R}$ and a metric of approximate submodularity $\zeta : \mathcal{F}_+ \to \mathbb{R}$ where ζ is defined by any of the following: (I) $\zeta[f] = \mathcal{E}[f]$, (II) $\zeta[f] = D[f]$, or (III) $\zeta[f] = d^{\ell,k}[f]$, for some $\ell \in \{0, \ldots, |\Omega| - 1\}, k \in \{0, \ldots, |\Omega|\}$. Then we have:*

(i) *The function ζ is sublinear. That is, ζ is subadditive (i.e., $\zeta[f_1] + \zeta[f_2] \geq \zeta[f_1 + f_2]$) and positively homogeneous with degree 1 (i.e., $\alpha\zeta[f] = \zeta[\alpha f]$, for $\alpha \in \mathbb{R}_+$).*

(ii) *If f is not submodular, then for any $\epsilon \in [0, \zeta[f])$, there does not exist a nonnegative, increasing, submodular function $g : 2^\Omega \to \mathbb{R}$ such that $\|g - f\|_\infty < \frac{\epsilon}{4}$.*

Proof: For both claims, we prove case (III); proofs for other cases are similar. Let $f_j : 2^\Omega \to \mathbb{R}_+$ be increasing set functions for $j \in \{1, 2\}$.

Consider $\zeta(\cdot) = d^{\ell,k}(\cdot), \ell \in \{0, \ldots, |\Omega| - 1\}, k \in \{0, \ldots, |\Omega|\}$. Observe that

$$
\begin{aligned}
\zeta[f_1 + f_2] &= \max_{\substack{\mathcal{A},\mathcal{B} \subseteq \Omega, s \in \Omega \\ |\mathcal{A}|=\ell, |\mathcal{B}|=k}} \sum_{j=1}^{2} \Big(f_j(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f_j(\mathcal{A} \cup \mathcal{B}) - f_j(\mathcal{A} \cup \{s\}) + f_j(\mathcal{A}) \Big) \\
&\leq \sum_{j=1}^{2} \left(\max_{\substack{\mathcal{A},\mathcal{B} \subseteq \Omega, s \in \Omega \\ |\mathcal{A}|=\ell, |\mathcal{B}|=k}} \Big(f_j(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f_j(\mathcal{A} \cup \mathcal{B}) - f_j(\mathcal{A} \cup \{s\}) + f_j(\mathcal{A}) \Big) \right) \\
&= \zeta[f_1] + \zeta[f_2],
\end{aligned}
$$

which proves subadditivity.

Let $(\mathcal{A}^*, \mathcal{B}^*, s^*) \in \arg\max_{\substack{\mathcal{A},\mathcal{B} \subseteq \Omega, s \in \Omega \\ |\mathcal{A}|=\ell, |\mathcal{B}|=k}} \Big(f(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f(\mathcal{A} \cup \mathcal{B}) - f(\mathcal{A} \cup \{s\}) + f(\mathcal{A}) \Big)$. Then, for any $\alpha \in \mathbb{R}_+$, we have $(\mathcal{A}^*, \mathcal{B}^*, s^*) \in \arg\max_{\substack{\mathcal{A},\mathcal{B} \subseteq \Omega, s \in \Omega \\ |\mathcal{A}|=\ell, |\mathcal{B}|=k}} \Big(\alpha f(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - \alpha f(\mathcal{A} \cup \mathcal{B}) - \alpha f(\mathcal{A} \cup \{s\}) + \alpha f(\mathcal{A}) \Big)$, which implies $\zeta(\alpha f) = d^{\ell,k}(\alpha f) = \alpha\zeta[f]$, which proves positive homogeneity.

Now suppose $\zeta[f] > \epsilon$, for some $\epsilon \geq 0$. Let $g : 2^\Omega \to \mathbb{R}$ be any nonnegative, increasing set function such that $\|g - f\|_\infty < \epsilon/4$. Consider $(\mathcal{A}^*, \mathcal{B}^*, s^*) \in \arg\max_{\substack{\mathcal{A},\mathcal{B} \subseteq \Omega, s \in \Omega \\ |\mathcal{A}|=\ell, |\mathcal{B}|=k}} \Big(f(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f(\mathcal{A} \cup \mathcal{B}) - f(\mathcal{A} \cup \{s\}) + f(\mathcal{A}) \Big)$, we have

$$
\begin{aligned}
\zeta(g) &\geq g(\mathcal{A}^* \cup \mathcal{B}^* \cup \{s^*\}) - g(\mathcal{A}^* \cup \mathcal{B}^*) - g(\mathcal{A}^* \cup \{s^*\}) + g(\mathcal{A}^*) \\
&\geq f(\mathcal{A}^* \cup \mathcal{B}^* \cup \{s^*\}) - f(\mathcal{A}^* \cup \mathcal{B}^*) - f(\mathcal{A}^* \cup \{s^*\}) + f(\mathcal{A}^*) - \epsilon = \zeta[f] - \epsilon > 0,
\end{aligned}
$$

which implies g is not submodular. $\square$

COROLLARY 1. *Define $\mathcal{E}_+ : \mathcal{F} \to \mathbb{R}$ by $\mathcal{E}_+[f] := \max\{0, \mathcal{E}[f]\}$. Then $\mathcal{E}_+$ is an asymmetric seminorm on $\mathcal{F}$.*

Proof: Let $f_1, f_2 \in \mathcal{F}$. Observe that $\mathcal{E}_+[f_1] + \mathcal{E}_+[f_2] \geq \mathcal{E}[f_1] + \mathcal{E}[f_2] \geq \mathcal{E}[f_1 + f_2]$, by Theorem 1. If $\mathcal{E}[f_1 + f_2] \geq 0$, then $\mathcal{E}_+[f_1] + \mathcal{E}_+[f_2] \geq \mathcal{E}[f_1 + f_2] = \mathcal{E}_+[f_1 + f_2]$. Otherwise, $\mathcal{E}_+[f_1 + f_2] = 0$. We have $\mathcal{E}_+[f_1] \geq 0, \mathcal{E}_+[f_2] \geq 0$, which imply $\mathcal{E}_+[f_1] + \mathcal{E}_+[f_2] \geq 0 = \mathcal{E}_+[f_1 + f_2]$, thus proving subadditivity.

Suppose $f \in \mathcal{F}$ is such that $\mathcal{E}[f] \geq 0$. By Theorem 1, for any $\alpha \geq 0, \alpha\mathcal{E}_+[f] = \alpha\mathcal{E}[f] = \mathcal{E}[\alpha f] = \mathcal{E}_+[\alpha f]$. If $\mathcal{E}[f] < 0$, then $\mathcal{E}_+[f] = \mathcal{E}[f] = 0$. By similar arguments to those of Theorem 1, $\alpha\mathcal{E}_+[f] = 0 = \mathcal{E}_+[\alpha f]$; hence $\mathcal{E}_+$ satisfies positive homogeneity. Clearly, $\mathcal{E}_+$ is nonnegative, which concludes the proof. $\quad\square$

Lemma A.1. (Narayanan 1997 [26]) *For any modular set function $g' : 2^\Omega \to \mathbb{R}$ and $\mathcal{A}, \mathcal{B}, \mathcal{S}, \mathcal{T} \subseteq \Omega$ with $\mathcal{A} \cup \mathcal{B} = \mathcal{S} \cup \mathcal{T}$ and $\mathcal{A} \cap \mathcal{B} = \mathcal{S} \cap \mathcal{T}$, we have $g'(\mathcal{A}) + g'(\mathcal{B}) = g'(\mathcal{S}) + g'(\mathcal{T})$.*

PROPOSITION 1. *Given $f : 2^\Omega \to \mathbb{R}$ and the corresponding functions $f_1, f_2, f_\mathcal{A}$, and f_q, we have: (i) $\mathcal{E}[f] = \mathcal{E}[f_1]$. (ii) $2\mathcal{E}[f] \geq \mathcal{E}[f_2]$. (iii) $\mathcal{E}[f] \geq \mathcal{E}[f_\mathcal{A}]$ (iv) $\mathcal{E}[f] \geq \mathcal{E}[f_q]$. (v) $\mathcal{E}[f] \geq \mathcal{E}[f \circledast g]$.*

Proof: (i): For any $\mathcal{A}, \mathcal{B} \subseteq \Omega$, by de Morgan's laws, $(\mathcal{A} \cup \mathcal{B})^c = \mathcal{A}^c \cap \mathcal{B}^c$ and $(\mathcal{A} \cap \mathcal{B})^c = \mathcal{A}^c \cup \mathcal{B}^c$, from which the result immediately follows.

(ii): Define $\tilde{f} : 2^\Omega \to \mathbb{R}$ where $\tilde{f}(\mathcal{S}) = f(\mathcal{S}^c) - f(\Omega)$. Because $f(\Omega)$ is a constant, by (i), $\mathcal{E}[\tilde{f}] = \mathcal{E}[f]$. Thus, by Theorem 1, $\mathcal{E}[f_2] \leq \mathcal{E}[f] + \mathcal{E}[\tilde{f}] = 2\mathcal{E}[f]$.

(iii): Let $\mathcal{S}, \mathcal{T} \subseteq \mathcal{A}^c$. Then $\mathcal{S} \cup \mathcal{A}$ and $\mathcal{T} \cup \mathcal{A}$ are subsets of Ω. Hence, $f_\mathcal{A}(\mathcal{S}) + f_\mathcal{A}(\mathcal{T}) + \mathcal{E}[f] = f(\mathcal{A} \cup \mathcal{S}) + f(\mathcal{A} \cup \mathcal{T}) + \mathcal{E}[f] \geq f((\mathcal{A} \cup \mathcal{S}) \cup (\mathcal{A} \cup \mathcal{T})) + f((\mathcal{A} \cup \mathcal{S}) \cap (\mathcal{A} \cup \mathcal{T})) = f(\mathcal{A} \cup (\mathcal{S} \cup \mathcal{T})) + f(\mathcal{A} \cup (\mathcal{S} \cap \mathcal{T})) = f_\mathcal{A}(\mathcal{S} \cup \mathcal{T}) + f_\mathcal{A}(\mathcal{S} \cap \mathcal{T})$.

(iv): Consider $\mathcal{A}, \mathcal{B} \subseteq \Omega(q)$. We have $\mathcal{E}[f] + f_q(\mathcal{A}) + f_q(\mathcal{B}) = \mathcal{E}[f] + f(\bigcup_{i \in \mathcal{A}} \mathcal{S}(i)) + f(\bigcup_{i \in \mathcal{B}} \mathcal{S}(i)) \leq f(\bigcup_{i \in \mathcal{A} \cup \mathcal{B}} \mathcal{S}(i)) + f(\bigcup_{i \in \mathcal{A} \cap \mathcal{B}} \mathcal{S}(i)) = f_q(\mathcal{A} \cup \mathcal{B}) + f_q(\mathcal{A} \cap \mathcal{B})$.

(v) This proof is similar to that of [26]. Let $\mathcal{S}, \mathcal{T}, \mathcal{Z}_\mathcal{S}, \mathcal{Z}_\mathcal{T} \subseteq \Omega$, where $f \circledast g(\mathcal{S}) = f(\mathcal{Z}_\mathcal{S}) + g(\mathcal{C} \backslash \mathcal{Z}_\mathcal{S})$, and $f \circledast g(\mathcal{T}) = f(\mathcal{Z}_\mathcal{T}) + g(\mathcal{T} \backslash \mathcal{Z}_\mathcal{T})$. It is not hard to show that $(\mathcal{S} \backslash \mathcal{Z}_\mathcal{S}) \cup (\mathcal{T} \backslash \mathcal{Z}_\mathcal{T}) = ((\mathcal{S} \cup \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cup \mathcal{Z}_\mathcal{T})) \cup ((\mathcal{S} \cap \mathcal{T}) \backslash ((\mathcal{Z}_\mathcal{S} \cap \mathcal{Z}_\mathcal{T}))$ and $(\mathcal{S} \backslash \mathcal{Z}_\mathcal{S}) \cap (\mathcal{T} \backslash \mathcal{Z}_\mathcal{T}) = ((\mathcal{S} \cup \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cup \mathcal{Z}_\mathcal{T})) \cap ((\mathcal{S} \cap \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cap \mathcal{Z}_\mathcal{T}))$.

By Lemma A.1, $g(\mathcal{S} \backslash \mathcal{Z}_\mathcal{S}) + g(\mathcal{T} \backslash \mathcal{Z}_\mathcal{T}) = g((\mathcal{S} \cup \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cup \mathcal{Z}_\mathcal{T})) + g((\mathcal{S} \cap \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cap \mathcal{Z}_\mathcal{T}))$.

By the definition of $\mathcal{Z}_\mathcal{S}$ and $\mathcal{Z}_\mathcal{T}$,

$$f \circledast g(\mathcal{S}) + f \circledast g(\mathcal{T})$$
$$= f(\mathcal{Z}_\mathcal{S}) + f(\mathcal{Z}_\mathcal{T}) + g((\mathcal{S} \cup \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cup \mathcal{Z}_\mathcal{T})) + g((\mathcal{S} \cap \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cap \mathcal{Z}_\mathcal{T}))$$
$$\geq -\mathcal{E}[f] + f(\mathcal{Z}_\mathcal{S} \cup \mathcal{Z}_\mathcal{T}) + f(\mathcal{Z}_\mathcal{S} \cap \mathcal{Z}_\mathcal{T}) + g((\mathcal{S} \cup \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cup \mathcal{Z}_\mathcal{T})) + g((\mathcal{S} \cap \mathcal{T}) \backslash (\mathcal{Z}_\mathcal{S} \cap \mathcal{Z}_\mathcal{T}))$$
$$\geq -\mathcal{E}[f] + \min_{\mathcal{Z} \subseteq \mathcal{S} \cup \mathcal{T}} f(\mathcal{Z}) + g((\mathcal{S} \cup \mathcal{T}) \backslash \mathcal{Z}) + \min_{\mathcal{Z} \subseteq \mathcal{S} \cap \mathcal{T}} f(\mathcal{Z}) + g((\mathcal{S} \cap \mathcal{T}) \backslash \mathcal{Z})$$
$$= -\mathcal{E}[f] + f \circledast g(\mathcal{S} \cup \mathcal{T}) + f \circledast g(\mathcal{S} \cap \mathcal{T}).$$

Hence, $\mathcal{E}[f] \geq \mathcal{E}[f \circledast g]$. $\quad\square$

Lemma A.2. (Bach 2013 [5]) *Given a set function f, its Lovász extension is positively homogeneous of degree 1.*

Lemma A.3. *Suppose $x = w + \alpha x(\mathcal{A}) \in [0,1]^{|\Omega|}$, where $w, \alpha x(\mathcal{A}) \in [0,1]^{|\Omega|}, \mathcal{A} \subseteq \Omega$, $\alpha \in \mathbb{R}_+$, and $x_s \geq x_{s'}$ for any $s \in \mathcal{A}, s' \in \mathcal{A}^c$. If there exists a permutation π of Ω such that $x_{\pi_1} \geq \cdots \geq x_{\pi_{|\Omega|}}$ and $w_{\pi_1} \geq \cdots \geq w_{\pi_{|\Omega|}}$, then $F^L(x) = F^L(w) + \alpha f(\mathcal{A})$.*

Proof: Let π be the ranking permutation in the hypothesis, and note that it is also a ranking permutation of $x(\mathcal{A})$; i.e., $x(\mathcal{A})_{\pi_1} \geq \cdots \geq x(\mathcal{A})_{\pi_{|\Omega|}}$. We have $F^L(x) = \sum\limits_{k=1}^{|\Omega|} x_{\pi_k}(f(x(\mathcal{S}_k^\pi)) - f(x(\mathcal{S}_{k-1}^\pi)))$, and $F^L(w) = \sum\limits_{k=1}^{|\Omega|} w_{\pi_k}(f(x(\mathcal{S}_k^\pi)) - f(x(\mathcal{S}_{k-1}^\pi)))$, implying that $F^L(x) - F^L(w) = \sum\limits_{k=1}^{|\Omega|} \alpha x(\mathcal{A})_{\pi_k}(f(x(\mathcal{S}_k^\pi)) - f(x(\mathcal{S}_{k-1}^\pi))) = F^L(\alpha x(\mathcal{A})) = \alpha F^L(x(\mathcal{A})) = \alpha f(\mathcal{A})$, where we have used the positive homogeneity of F^L (Lemma A.2). $\square$

Lemma A.4. *Let $f : 2^\Omega \to \mathbb{R}$ be increasing with $f(\emptyset) = 0$, and let $\tilde{\gamma} \in \Gamma(f)$. For any $\mathcal{S} \subseteq \Omega$, $f(\mathcal{S}) \geq -|\mathcal{S}|D[f] + \sum\limits_{s \in \mathcal{S}} \tilde{\gamma}_s$.*

Proof: Consider a permutation $(\rho_1, \ldots, \rho_{|\Omega|})$ such that $\tilde{\gamma}_{\rho_1} \geq \cdots \geq \gamma_{\rho_{|\Omega|}}$ and set $\theta^* = -|\Omega|D[f]$. We prove by induction on $|\mathcal{S}|$ that $f(\mathcal{S}) \geq -|\mathcal{S}|D[f] + \sum\limits_{s \in \mathcal{S}} \tilde{\gamma}_s$, for all $\mathcal{S} \subseteq \Omega$. The base case is confirmed as $f(\emptyset) = 0 = \sum\limits_{s \in \emptyset} \tilde{\gamma}_s$. Assume for all $\tilde{\mathcal{S}} \subseteq \Omega$ with $|\tilde{\mathcal{S}}| \leq \alpha$, $f(\tilde{\mathcal{S}}) \geq -|\tilde{\mathcal{S}}|D[f] + \sum\limits_{s \in \tilde{\mathcal{S}}} \tilde{\gamma}_s$, and let $|\mathcal{S}| = \alpha + 1$. Set $k = \max\{i \mid \rho_i \in \mathcal{S}\}$. Then $\mathcal{S} \cup \mathcal{S}_{k-1}^\rho = \mathcal{S}_k^\rho$ and $\mathcal{S} \cap \mathcal{S}_{k-1}^\rho = \mathcal{S} \backslash \{\rho_k\}$. Observe that by the definition of $D[f]$,

$$f(\mathcal{S}) \geq f(\mathcal{S} \cup \mathcal{S}_{k-1}^\rho) + f(\mathcal{S} \cap \mathcal{S}_{k-1}^\rho) - f(\mathcal{S}_{k-1}^\rho) - D[f]$$

$$= f(\mathcal{S}_k^\rho) - f(\mathcal{S}_{k-1}^\rho) - D[f] + f(\mathcal{S} \backslash \{\rho_k\})$$

$$= \tilde{\gamma}_{\rho_k} + f(\mathcal{S} \backslash \{\rho_k\}) - D[f]$$

$$\geq \sum\limits_{s \in \mathcal{S}} \tilde{\gamma}_s - |\mathcal{S}|D[f],$$

where the last line uses the induction hypothesis. $\square$

THEOREM 2. *For any increasing set function $f : 2^\Omega \to \mathbb{R}$ such that $f(\emptyset) = 0$,*

$$F^L(x) \leq \max_{\gamma \in \Gamma(f)} \sum_{s \in \Omega} \gamma_s x_s \leq F^C(x) + |\Omega|D[f]$$

$$\leq F^L(x) + |\Omega|D[f] \leq \max_{\gamma \in \Gamma(f)} \sum_{s \in \Omega} \gamma_s x_s + |\Omega|D[f].$$

Hence, $F^L(x) \geq F^C(x) \geq F^L(x) - |\Omega|D[f]$, and $\|F^L - F^C\|_\infty \leq |\Omega|D[f]$. Moreover, F^L is $|\Omega|D[f]$-approximately convex.

In addition, if for some $\epsilon > 0$, F^L is ϵ-approximately convex, then $D[f] \leq \epsilon$.

Proof: The following proof uses a version of well-known linear programming duality arguments (e.g., [23]). We first suppose f is not submodular (hence $D[f] > 0$).

Given $x \in [0,1]^{|\Omega|}$ there exists a permutation $(\pi_1, \ldots, \pi_{|\Omega|})$ such that $x_{\pi_1} \geq \cdots \geq x_{\pi_{|\Omega|}}$. Let $x_{\pi_0} = 1$. Consider the dual of (1):

$$\max_{\gamma, \theta} \left\{ \theta + \sum_{s \in \Omega} x_s \gamma_s \,\middle|\, \theta + \sum_{s \in \mathcal{S}} x_s \leq f(\mathcal{S}), \forall \, \mathcal{S} \subseteq \Omega \right\}. \tag{5}$$

Define $y^* \in \mathbb{R}^{|2^\Omega|}$ by $y^*_{\mathcal{S}}$ equals $x_{\pi_i} - x_{\pi_{i+1}}$, if $\mathcal{S} = \mathcal{S}^\pi_i, i \in \{0, \ldots, |\Omega| - 1\}$, $x_{\pi_{|\Omega|}}$, if $\mathcal{S} = \Omega$, and 0 otherwise.

We first show y^* is feasible for (1). Observe that $\sum_{s \in \Omega} y^*(\mathcal{S}) = \sum_{i=0}^{|\Omega|-1} (x_{\pi_i} - x_{\pi_{i+1}}) + x_{\pi_{|\Omega|}} = x_{\pi_0} = 1$.

In addition, for any $s \in \Omega, s = \pi_j$ for some $j \in \Omega$; hence, $\sum_{\mathcal{S} \ni s} y^*(\mathcal{S}) = \sum_{i=j}^{|\Omega|} y^*(\mathcal{S}^\pi_i) = x_{\pi_j} = x_s$. Moreover, it is easy to observe that y^* is nonnegative. Hence, y^* is feasible for (1), and $F^C(x) \leq \sum_{\mathcal{S} \subseteq \Omega} f(\mathcal{S}) y^*(\mathcal{S}) = x_{\pi_{|\Omega|}} f(\Omega) + \sum_{i=0}^{|\Omega|-1} (x_{\pi_i} - x_{\pi_{i+1}}) f(\mathcal{S}^\pi_i) = F^L(x)$.

Consider $\gamma^* \in \Gamma(f)$ such that $\gamma^*_{\pi_1} \geq \cdots \geq \gamma^*_{\pi_{|\Omega|}}$. Then $\sum_{s \in \Omega} \gamma^*_s x_s = \sum_{i=1}^{|\Omega|} (f(\mathcal{S}^\pi_i) - f(\mathcal{S}^\pi_{i-1})) x_{\pi_i} = F^L(x)$. By Lemma A.4, $(\gamma^*, -|\Omega|D[f])$ is feasible for (5).

Therefore, $F^L(x) \leq \max_{\gamma \in \Gamma(f)} \sum_{s \in \Omega} \gamma_s x_s \leq F^C(x) + |\Omega|D[f] \leq F^L(x) + |\Omega|D[f] = |\Omega|D[f] + \max_{\gamma \in \Gamma(f)} \sum_{s \in \Omega} \gamma_s x_s$. This also implies that $F^L(x) \geq F^C(x) \geq F^L(x) - |\Omega|D[f]$ and $||F^L - F^C||_\infty \leq |\Omega|D[f]$.

To show that F^L is approximately convex, consider $x, y \in [0,1]^{|\Omega|}, \lambda \in [0,1]$, then we have

$$F^L(\lambda x + (1 - \lambda)y) \leq F^C(\lambda x + (1 - \lambda)y) + |\Omega|D[f]$$
$$\leq \lambda F^C(x) + (1 - \lambda)F^C(y) + |\Omega|D[f]$$
$$\leq \lambda F^L(x) + (1 - \lambda)F^L(y) + |\Omega|D[f].$$

For the last statement, suppose F^L is ϵ-approximately convex. For any $\mathcal{A}, \mathcal{B} \in \Omega$, denote the symmetric difference as $\mathcal{A} \doteq \mathcal{B} = (\mathcal{A} \backslash \mathcal{B}) \cup (\mathcal{B} \backslash \mathcal{A})$. Consider any $S, T \subseteq \Omega$, then we have

$$\frac{1}{2}(f(\mathcal{S}) + f(\mathcal{T})) + \epsilon = \frac{1}{2}F^L(x(\mathcal{S})) + \frac{1}{2}F^L(x(\mathcal{T})) + \epsilon$$
$$\geq F^L(\frac{1}{2}(x(\mathcal{S}) + x(\mathcal{T})))$$
$$= F^L(\frac{1}{2}(x(\mathcal{S} \cap \mathcal{T}) + x(\mathcal{S} \cap \mathcal{T}) + x(\mathcal{S} \doteq \mathcal{T})))$$
$$= F^L(\frac{1}{2}(x(\mathcal{S} \cap \mathcal{T}) + x(\mathcal{S} \cup \mathcal{T})))$$
$$= \frac{1}{2}(f(\mathcal{S} \cap \mathcal{T}) + f(\mathcal{S} \cup \mathcal{T}))$$

where the last line uses the fact that the Lovász extension is positively homogeneous Lemma $A.2$ and Lemma $A.3$.

Suppose f is submodular. Then, $D[f] = 0$ and $F^C = F^L$ (e.g., see [5, 23]). A similar linear programming duality argument (with $D[f]$ replaced with 0) proves that $F^L(x) = \max\limits_{\gamma \in \Gamma(f)} \sum\limits_{s \in \Omega} \gamma_s x_s$. $\qquad \square$

PROPOSITION 3. *Given set functions $f, g : 2^\Omega \to \mathbb{R}$, where $f(\emptyset) = g(\emptyset) = 0$, and their respective Lovász extensions $F^L, G^L : [0, 1]^{|\Omega|} \to \mathbb{R}$, $||F^L - G^L||_\infty = ||f - g||_\infty$.*

Proof: Let $x \in [0, 1]^{|\Omega|}$ and π a permutation of $(1, \ldots, |\Omega|)$ such that $x_{\pi_1} \geq x_{\pi_2} \geq \cdots \geq x_{\pi_{|\Omega|}}$, and set $x_{\pi_{|\Omega|+1}} = 0$. Then $F^L(x) = \sum\limits_{i=1}^{|\Omega|} (f(\mathcal{S}_i^\pi) - f(\mathcal{S}_{i-1}^\pi))x_{\pi_i} = -x_{\pi_{|\Omega|}} f(\Omega) + \sum\limits_{i=1}^{|\Omega|-1} (x_{\pi_i} - x_{\pi_{i+1}})f(\mathcal{S}_i^\pi)$, and similarly for G^L. Hence,

$$
\begin{aligned}
|F^L(x) - G^L(x)| &= \left| x_{\pi_{|\Omega|}}(f(\Omega) - g(\Omega)) + \sum_{i=1}^{|\Omega|-1} (x_{\pi_i} - x_{\pi_{i+1}})(f(\mathcal{S}_i^\pi) - g(\mathcal{S}_i^\pi)) \right| \\
&\leq |x_{\pi_{|\Omega|}}(f(\Omega) - g(\Omega))| + \sum_{i=1}^{|\Omega|-1} \left| (x_{\pi_i} - x_{\pi_{i+1}})(f(\mathcal{S}_i^\pi) - g(\mathcal{S}_i^\pi)) \right| \\
&= |x_{\pi_{|\Omega|}}| \cdot |(f(\Omega) - g(\Omega))| + \sum_{i=1}^{|\Omega|-1} |(x_{\pi_i} - x_{\pi_{i+1}})| \cdot |(f(\mathcal{S}_i^\pi) - g(\mathcal{S}_i^\pi))| \\
&\leq |x_{\pi_{|\Omega|}}| \cdot ||f - g||_\infty + \sum_{i=1}^{|\Omega|-1} |(x_{\pi_i} - x_{\pi_{i+1}})| \cdot ||f - g||_\infty \\
&= ||f - g||_\infty \left(x_{\pi_{|\Omega|}} + \sum_{i=1}^{|\Omega|-1} (x_{\pi_i} - x_{\pi_{i+1}}) \right) \\
&\leq ||f - g||_\infty.
\end{aligned}
$$

Hence, $||F^L - G^L||_\infty \leq ||f - g||_\infty$. Moreover, for some $\mathcal{S} \subseteq \Omega, ||f - g||_\infty = |f(\mathcal{S}) - g(\mathcal{S})| = |F^L(x(\mathcal{S})) - G^L(x(\mathcal{S}))|$, which implies $||F^L - G^L||_\infty = ||f - g||_\infty$. $\qquad \square$

LEMMA 2. *For any $\gamma \in \Gamma(g_\tau), \mathcal{S} \subseteq \Omega$, we have $-|\Omega|D[f_\tau] + \sum\limits_{s \in \Omega} \gamma_s (x(\mathcal{S}))_s \leq f_\tau(\mathcal{S}) - \phi(\tau)$.*

Proof: Recall that $g_\tau(\emptyset) = 0$ and $D[f_\tau] = D[g_\tau]$. By Theorem 2, $-|\Omega|D[f_\tau] + \sum\limits_{s \in \Omega} \gamma_s (x(\mathcal{S}))_s \leq G_\tau^L(x(\mathcal{S})) = g_\tau(\mathcal{S}) = f_\tau(\mathcal{S}) - \phi(\tau)$. $\qquad \square$

PROPOSITION 4. *For any $\gamma \in \Gamma(g_\sigma)$, the following inequality is valid for $\mathrm{conv}(H_\mathbb{B})$:*

$$
-|\Omega|D[f_\sigma] + \sum_{s \in \Omega} \gamma_s x_s \leq z - \phi(\sigma). \tag{6}
$$

Proof: This proof follows arguments similar to that of [2]. Consider $(x, z) \in H_\mathbb{B}$, which implies $x = x(\mathcal{S})$, for some $\mathcal{S} \subseteq \Omega$. From Lemma 2, $-|\Omega|D[f_\sigma] + \sum\limits_{s \in \Omega} \gamma_s (x(\mathcal{S}))_s \leq f_\sigma(\mathcal{S}) - \phi(\sigma) = g_\sigma(\mathcal{S})$. Because $(x, z) \in H_\mathbb{B}, F_\sigma(x) = f_\sigma(\mathcal{S}) = g_\sigma(\mathcal{S}) + \phi(\sigma) \leq z$, which implies $-|\Omega|D[f_\sigma] + \sum\limits_{s \in \Omega} \gamma_s x_s \leq z - \phi(\sigma)$. $\qquad \square$

PROPOSITION 6. *Let $f : 2^\Omega \to \mathbb{R}$ be increasing with $f(\emptyset) = 0$. Suppose $\gamma \in \mathbb{R}^{|\Omega|}$ and*

$$\sum_{s \in \Omega} \gamma_s x_s \le z + |\Omega| D[f] + \gamma_0 \tag{7}$$

defines a nontrivial facet of H_f. Let $\bar{f} : 2^\Omega \to \mathbb{R}$ be defined by $\bar{f}(\emptyset) = 0, \bar{f}(\mathcal{S}) = f(\mathcal{S}) + |\Omega| D[f] + \gamma_0$, for all nonempty $\mathcal{S} \subseteq \Omega$, and suppose $\gamma \in \Gamma(\bar{f})$. Then, $\gamma_0 \le 0$.

Proof: This proof follows steps similar to that of [4], with additional arguments to account for approximate submodularity. Suppose $\gamma_0 > 0$. Because (7) is a valid inequality for H_f, for any non-empty $\mathcal{S} \subseteq \Omega$, $\sum_{s \in \mathcal{S}} \gamma_s = \sum_{s \in \Omega} \gamma_s x(\mathcal{S})_s \le f(\mathcal{S}) + \gamma_0 + |\Omega| D[f] = \bar{f}(\mathcal{S})$, and $\sum_{s \in \emptyset} \gamma_s = 0 = \bar{f}(\emptyset)$. Thus, $\gamma \in P_{\bar{f}}$, and by Proposition 5, (8) is valid for $H_{\bar{f}}$:

$$\sum_{s \in \Omega} \gamma_s x_s \le z. \tag{8}$$

We show that (8) is facet-defining for $H_{\bar{f}}$. Observe that the solutions $\{(x(\{s\}), f(\{s\}))\}_{s \in \Omega} \cup \{(x(\emptyset), 1),$ $(x(\emptyset), 0)\}$ are $|\Omega| + 2$ affinely independent solutions, so the dimension of H_f is $|\Omega| + 1$. By the hypothesis, $\sum_{s \in \Omega} \gamma_s x_s \le z + |\Omega| D[f] + \gamma_0$ is facet-defining for H_f. Thus, there exist $|\Omega| + 1$ affinely independent solutions $\{(x^k, z^k)\}_{k=1}^{|\Omega|+1}$ such that $(x^k, z^k) \in H_f$ and $\sum_{s \in \Omega} \gamma_s x_s^k = z^k + |\Omega| D[f] + \gamma_0$. By Carathéodory's theorem, each of the affinely independent solutions (x^k, z^k) can be represented by a convex combination of $|\Omega| + 2$ (integral) extreme points of H_f: $(x^k, z^k) = \sum_{\ell=1}^{|\Omega|+2} \lambda^{k,\ell}(x^{k,\ell}, z^{k,\ell})$, where $\sum_{\ell=1}^{|\Omega|+2} \lambda^{k,\ell} = 1, \lambda^k \in \mathbb{R}_+^{|\Omega|+2}$, and $(x^{k,\ell}, z^{k,\ell})$ is an integral extreme point of H_f, for all $\ell \in \{1, \dots, |\Omega| + 2\}$. Consider $(x^{k,\ell}, z^{k,\ell})$ for some $k \in \{1, \dots, |\Omega| + 1\}$ and $\ell \in \{1, \dots, |\Omega| + 2\}$. Suppose that $x^{k,\ell} = 0$; because $(x^{k,\ell}, z^{k,\ell}) \in H_f$, $0 \le z^{k,\ell}$. Thus, $\bar{f}(x^{k,\ell}) = 0 \le z^{k,\ell} < z^{k,\ell} + |\Omega| D[f] + \gamma_0$. If instead $x^{k,\ell} \ne 0$, then $x^{k,\ell} = x(\mathcal{S})$ for some nonempty $\mathcal{S} \subseteq \Omega$. Notice that by $(x^{k,\ell}, z^{k,\ell}) \in H_f$, $f(x^{k,\ell}) \le z^{k,\ell}$, so $\bar{f}(x^{k,\ell}) \le z^{k,\ell} + |\Omega| D[f] + \gamma_0$. Hence, $(x^{k,\ell}, z^{k,\ell} + |\Omega| D[f] + \gamma_0) \in H_{\bar{f}}$; moreover, $(x^k, z^k + |\Omega| D[f] + \gamma_0) = \sum_{\ell=1}^{|\Omega|+2} \lambda^{k,\ell}(x^{k,\ell}, z^{k,\ell} + |\Omega| D[f] + \gamma_0) \in H_{\bar{f}}$.

Suppose that the points $(x^k, z^k + |\Omega| D[f] + \gamma_0)_{k=1}^{|\Omega|+1}$ are not affinely independent. Then there exists $\sigma \in \mathbb{R}^{|\Omega|+1} \backslash \{0\}$ such that $\sum_{k=1}^{|\Omega|+1} \sigma_k = 0$ and $\sum_{k=1}^{|\Omega|+1} \sigma_k(x^k, z^k + |\Omega| D[f] + \gamma_0) = (0, 0)$. Let $j \in \{1, \dots, |\Omega| + 1\}$ be such that $\sigma_j \ne 0$; without loss of generality, let $\sigma_j = 1$. Thus, $\sum_{k \ne j} \sigma_k(x^k, z^k + |\Omega| D[f] + \gamma_0) = -(x^j, z^j + |\Omega| D[f] + \gamma_0)$. Because $\sigma_j = 1$, $\sum_{k \ne j} \sigma_k = -1$; thus,

$$-(x^j, z^j + |\Omega| D[f] + \gamma_0) = \sum_{k \ne j} \sigma_k(x^k, z^k + |\Omega| D[f] + \gamma_0)$$

$$= -(0, |\Omega| D[f] + \gamma_0) + \sum_{k \ne j} \sigma_k(x^k, z^k)$$

$$\iff -(x^j, z^j) = \sum_{k \neq j} \sigma_k(x^k, z^k),$$

which contradicts the affine independence of $(x^k, z^k)_{k=1}^{|\Omega|+1}$, so $(x^k, z^k + |\Omega|D[f] + \gamma_0)_{k=1}^{|\Omega|+1}$ are affinely independent. Also, $\sum_{s \in \Omega} \gamma_s x_s^k = z^k + |\Omega|D[f] + \gamma_0$, for each $k \in \{1, \dots, |\Omega|+1\}$, which implies (8) is facet-defining for $H_{\bar{f}}$.

By Proposition 5, γ is an extreme point of $P_{\bar{f}}$, and by the hypothesis, there exists a permutation $(\rho_1, \dots, \rho_{|\Omega|})$ such that $\gamma_{\rho_s} = \bar{f}(\mathcal{S}_s^\rho) - \bar{f}(\mathcal{S}_{s-1}^\rho)$, for all $s \in \Omega$. Define $\hat{\gamma}$ by $\hat{\gamma}_{\rho_1} = \gamma_{\rho_1} - \gamma_0, \hat{\gamma}_{\rho_s} = \gamma_{\rho_s}$, otherwise. Thus, $\hat{\gamma}_{\rho_s} = f(\mathcal{S}_s^\rho) - f(\mathcal{S}_{s-1}^\rho)$ and $\hat{\gamma} \in \Gamma(f)$. Because $\gamma_0 > 0$, $\bar{f}$ is increasing, so that by Lemma A.4, $\sum_{s \in \mathcal{S}} \hat{\gamma}_s \leq f(\mathcal{S}) + |\mathcal{S}|D[f]$. Hence, if $\gamma_s' = \hat{\gamma}_s - D[f]$, then $\gamma' \in P_f$. By Proposition 5, we have $\sum_{s \in \Omega} \gamma_s' x_s \leq z$ is valid for H_f, which implies $(\gamma_{\rho_1} - \gamma_0)x_{\rho_1} + \sum_{s \in \Omega \setminus \{\rho_1\}} \hat{\gamma}_s x_s \leq z + |\Omega|D[f]$ is also valid for H_f. Because $\gamma_0 > 0$, we also have the valid inequality $\gamma_0 x_{\rho_1} \leq \gamma_0$.

Combining these last two inequalities implies $\sum_{s \in \Omega} \gamma_s x_s \leq z + |\Omega|D[f] + \gamma_0$, thus the facet-defining inequality (7) is dominated, a contradiction. $\qquad\square$

Proposition A.1. *If $f(\{s\}) \leq b$, for all $s \in \Omega$, then X is full-dimensional.*

Proof: By the hypothesis, the zero vector and $x(\{s\})$ are feasible for each $s \in \Omega$. Hence there are $|\Omega| + 1$ affinely independent points in X, implying the dimension of X is $|\Omega|$. $\qquad\square$

Proposition A.2. (Hammer et al. 1975 [16], Atamtürk and Narayanan 2009 [3])

1. *The inequality $x(\{s\}) \geq 0$ is facet-defining for $\mathrm{conv}(X)$, for all $s \in \Omega$.*

2. *The inequality $x(\{s\}) \leq 1$ is facet-defining for $\mathrm{conv}(X)$ if and only if $f(\{s, t\}) \leq b$ for all $t \in \Omega \setminus s$.*

PROPOSITION 7. *If $\mathcal{S} \subseteq \Omega$ is a cover for X, the extended cover inequality $\sum_{s \in E_\pi(\mathcal{S})} x_s \leq |\mathcal{S}| - 1$ is valid for X if $f(\mathcal{S}) > (|\mathcal{S}| + |U_\pi(\mathcal{S})|)D[f] + b$. In addition, the inequality defines a facet of $\{x \in X \mid x_s = 0, \forall s \notin E_\pi(\mathcal{S})\}$ if $\mathcal{S}$ is also a minimal cover and for each $s \in U_\pi(\mathcal{S})$, there exist $t_s, u_s \in \mathcal{S}$ such that $t_s \neq u_s$, and $f(\mathcal{S} \cup \{s\} \setminus \{t_s, u_s\}) \leq b$.*

Proof: This proof uses steps similar to those in [3], who establish the submodular case. We show that if $x \in [0,1]^{|\Omega|}$ with $\sum_{s \in E_\pi(\mathcal{S})} x_s > |S| - 1$, then $x_s \notin X$. Because X is the convex hull of characteristic vectors, it suffices to consider such characteristic vectors. That is $x(\tilde{S})$, where $\tilde{S} \subseteq \Omega$ and there exists $\mathcal{T} \subseteq \tilde{S}$ such that $\mathcal{T} \subseteq E_\pi(\mathcal{S})$ with $|\mathcal{T}| \geq |\mathcal{S}|$. In this case, $\sum_{s \in E_\pi(\mathcal{S})} x(\tilde{S})_s \geq \sum_{s \in \mathcal{T}} x(\tilde{S})_s \geq |\mathcal{S}|$. Let $\mathcal{K} = \mathcal{S} \setminus \mathcal{T}$, and $\mathcal{L} = U_\pi(\mathcal{S}) \cap \mathcal{T} = \{\ell_1, \dots, \ell_{|\mathcal{L}|}\}$, with indexing consistent with π.

Observe that $\mathcal{S} \setminus \mathcal{K} = \mathcal{S} \cap \mathcal{T}$ and $(\mathcal{S} \cup \mathcal{L}) \setminus \mathcal{K} = (\mathcal{S} \cup (U_\pi(\mathcal{S}) \cap \mathcal{T})) \setminus (\mathcal{S} \setminus \mathcal{T}) = (\mathcal{S} \cap \mathcal{T}) \cup (U_\pi(\mathcal{S}) \cap \mathcal{T}) = \mathcal{T}$. Hence, $f(\mathcal{T}) = f(\mathcal{S} \setminus \mathcal{K}) + \sum_{\ell_i \in \mathcal{L}} f((\mathcal{S} \cup \{\ell_1, \dots, \ell_i\}) \setminus \mathcal{K}) - f((\mathcal{S} \cup \{\ell_1, \dots, \ell_{i-1}\}) \setminus \mathcal{K})$.

Given $\ell_i \in \mathcal{L}$, let $\pi_j = \ell_i$. Then $(\mathcal{S} \cup \{\ell_1, \ldots, \ell_{i-1}\}) \backslash \mathcal{K} \subseteq \mathcal{S} \cup \{\pi_1, \ldots, \pi_{j-1}\}$; it follows from the definition of $D[f]$ that

$$f((\mathcal{S} \cup \{\ell_1, \ldots, \ell_i\}) \backslash \mathcal{K}) - f((\mathcal{S} \cup \{\ell_1, \ldots, \ell_{i-1}\}) \backslash \mathcal{K}) \geq f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_j\}) - f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_{j-1}\}) - D[f].$$

Therefore,

$$f(\mathcal{T}) \geq f(\mathcal{S} \backslash \mathcal{K}) - |\mathcal{L}|D[f] + \sum_{\pi_j \in \mathcal{L}} f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_j\}) - f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_{j-1}\}).$$

By the definition of $U_\pi(\mathcal{S})$, for all $s \in \mathcal{S}$, $f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_j\}) - f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_{j-1}\}) \geq f(\{s\})$. Because $\mathcal{T} = (\mathcal{S} \cup \mathcal{L}) \backslash \mathcal{K}$ and $|\mathcal{T}| \geq |\mathcal{S}|$, $|\mathcal{S} \cap \mathcal{T}| + |\mathcal{K}| = |\mathcal{S}| \leq |\mathcal{T}| = |\mathcal{S} \cap \mathcal{T}| + |\mathcal{L} \cap \mathcal{T}| = |\mathcal{S} \cap \mathcal{T}| + |\mathcal{L}|$; thus, $|\mathcal{K}| \leq |\mathcal{L}|$. This implies that

$$f(\mathcal{S} \backslash \mathcal{K}) - |\mathcal{L}|D[f] + \sum_{\pi_j \in \mathcal{L}} f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_j\}) - f(\mathcal{S} \cup \{\pi_1, \ldots, \pi_{j-1}\}) \geq f(\mathcal{S} \backslash \mathcal{K}) - |\mathcal{L}|D[f] + \sum_{s \in \mathcal{K}} f(\{s\}).$$

By the definition of $D[f]$, $f(\mathcal{T}) \geq f(\mathcal{S} \backslash \mathcal{K}) - (|\mathcal{L}| + |\mathcal{K}|)D[f] + \sum_{s \in \mathcal{K}} f((\mathcal{S} \backslash \mathcal{K}) \cup \{s\}) - f(\mathcal{S} \backslash \mathcal{K})$. By the monotonicity of f, $f(\mathcal{S}) \geq f(\mathcal{S} \backslash \mathcal{K} \cup \{s\}) \geq f(\mathcal{S} \backslash \mathcal{K})$, for all $s \in \mathcal{K}$. Also, $|\mathcal{L}| \leq |U_\pi(\mathcal{S})|$ and $|\mathcal{K}| \leq |\mathcal{S}|$. Thus, $f(\mathcal{T}) \geq f(\mathcal{S}) - (|\mathcal{S}| + |U_\pi(\mathcal{S})|)D[f] > b$, which follows from the hypothesis. It follows that $f(x(\tilde{\mathcal{S}})) > b$.

To prove the facet claim, observe that each of the points $x(\mathcal{S} \backslash \{s\})$, for all $s \in \mathcal{S}$ and $x(\mathcal{S} \cup \{s\} \backslash \{t_s, u_s\})$, for all $s \in U_\pi(\mathcal{S})$, are $|E_\pi(\mathcal{S})|$ affinely independent points in $\{x \in X \mid x_s = 0, \forall s \notin E_\pi(\mathcal{S})\}$, and the valid inequality holds with equality for these points. Thus, the valid inequality defines a facet of $\{x \in X \mid x_s = 0, \forall s \notin E_\pi(\mathcal{S})\}$. $\qquad \square$

PROPOSITION 8. *We have $D[f] \leq p||w||_1^{p-1}||w||_\infty$.*

Proof: Let $\mathcal{W} = \{s \in \Omega \mid w_s > 0\}$. Let $\ell \in \{0, \ldots, |\Omega| - 1\}, k \in \{0, \ldots, |\Omega|\}, \mathcal{A}, \mathcal{B} \subseteq \Omega, s \in \Omega, |\mathcal{A}| = \ell, |\mathcal{B}| = k$. Suppose that $s \notin \mathcal{W}$. Then $(f(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f(\mathcal{A} \cup \mathcal{B})) - (f(\mathcal{A} \cup \{s\}) - f(\mathcal{A})) = u_s - u_s \leq p||w||_1^{1-p}||w||_\infty$.

Next, suppose $s \in \mathcal{W}$. Observe that H is Lipschitz continuous on $[0, ||w||_1]$, the codomain of G: for any $z_1, z_2 \in [0, ||w||_1]$, we have $|H(z_2) - H(z_1)| \leq ||H'||_\infty |z_2 - z_1| \leq p||w||_1^{p-1}|z_2 - z_1|$. Hence, $f(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f(\mathcal{A} \cup \mathcal{B}) \leq u_s + p||w||_1^{p-1}w_s \leq u_s + p||w||_1^{p-1}||w||_\infty$. Also, $f(\mathcal{A} \cup \{s\}) - f(\mathcal{A}) \geq u_s$. Hence, $(f(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f(\mathcal{A} \cup \mathcal{B})) - (f(\mathcal{A} \cup \{s\}) - f(\mathcal{A})) \leq p||w||_1^{p-1}||w||_\infty$, so $D[f] \leq p||w||_1^{p-1}||w||_\infty$. $\qquad \square$

B   Example of Bounds on the Marginal Violation: The Cooperative Uncapacitated Facility Location Problem

We present a generalization of the well-known uncapacitated facility location problem (see [24] for a detailed overview). We choose uncapacitated facility location as a demonstrative example because of its historical importance (e.g., [9]). In this generalized facility location problem that we consider, the objective function is not submodular in general. We show that the marginal violation metric D (Definition 3) can be bounded analytically by exploiting the problem structure and that the objective function's proximity to submodularity is influenced by certain problem parameters.

The objective function of the uncapacitated facility location problem (UFLP) provides an example of a submodular function. An instance of UFLP is defined by m facility locations ($\Omega = \{1, ..., m\}$), n clients, demands $b \in \mathbb{R}^n_+$, fixed costs $w \in \mathbb{R}^m_+$, and facility-client revenues $v \in \mathbb{R}^{m \times n}$. We consider instances in which v is nonnegative. Additionally, we assume that $w = 0$ so that the firm only assigns facilities to clients based on the variable revenue. We note that [9] consider similar conditions. Let $f : 2^\Omega \to \mathbb{R}$ be the objective function of the UFLP with cardinality parameter $K \in \{1, \dots, |\Omega|\}$

$$f(\mathcal{S}) := \begin{cases} \sum_{j=1}^n b_j \max_{i \in \mathcal{S}} v_{ij}, & \text{if } \mathcal{S} \neq \emptyset \\ 0, & \text{if } \mathcal{S} = \emptyset. \end{cases} \qquad \text{UFLP: } \max_{\mathcal{S} \subseteq \Omega}\{f(\mathcal{S}) \text{ subject to } |\mathcal{S}| \leq K\}.$$

Here, $\mathcal{S}$ is a subset of facility locations. Under these conditions, f is nonnegative, increasing, and submodular. We consider a generalization of UFLP where the objective function is approximately submodular function. Let $\mathcal{S}^2 = \{(p, q) \in \{1, ..., m\}^2$, for any $\mathcal{S} \subseteq \Omega$. We introduce a nonnegative reward u_{pq} associated with the simultaneous selection of facilities p and q, where $(p, q) \in \Omega^2$. We assume that $u_{pp} = 0$ for all $p \in \Omega$. Define $h : 2^\Omega \to \mathbb{R}$ as the objective function of the *cooperative uncapacitated facility location problem* (CUFLP) with maximum cardinality parameter K.

$$h(\mathcal{S}) := \begin{cases} \sum_{j=1}^n b_j \max_{i \in \mathcal{S}} v_{ij} + \sum_{(p,q) \in \mathcal{S}^2} u_{pq}, & \text{if } \mathcal{S} \neq \emptyset \\ 0 & \text{if } \mathcal{S} = \emptyset. \end{cases} \qquad \text{CUFLP : } \max_{\mathcal{S} \subseteq \Omega}\{h(\mathcal{S}) \text{ subject to } |\mathcal{S}| \leq K\}.$$

Remark B.1. *It is well known that UFLP is NP-hard [10]; thus, CUFLP (which includes UFLP as a special case) is also NP-hard. In addition, the objective function of CUFLP, h is not submodular in general.*

Example B.1. *To illustrate the second statement of Remark B.1, consider an instance of the cooperative uncapacitated facility location problem in which $m = 3, n = 1$, and $v_{i1} = 0$, for $i = 1, 2, 3$, $b_1 = 1$, $u_{2,3} = 1$,*

and $u_{pq} = 0$ otherwise. The fixed costs are zero so h is increasing. Consider $\mathcal{A} = \{1\}, \mathcal{B} = \{2\}$, and $s = \{3\}$. Then, $h(\mathcal{A} \cup \mathcal{B} \cup \{s\}) = 1, \quad h(\mathcal{A} \cup \mathcal{B}) = 0, h(\mathcal{A} \cup \{s\}) = 0, \quad and \quad h(\mathcal{A}) = 0 \Rightarrow 1 = h(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - h(\mathcal{A} \cup \mathcal{B}) - h(\mathcal{A} \cup \{s\}) + h(\mathcal{A})$. By Lemma 1, h is not submodular.

Let $\operatorname{supp}(u) := \{(p, q) \in \Omega^2 \mid u_{pq} > 0\}$.

Proposition B.1. *Given an instance of CUFLP, we have $d^{\ell, k}[h] \leq |\operatorname{supp}(u)| \max\{u_{pq} \mid (p, q) \in \Omega^2\}$ for all $\ell \in \{0, ..., m-1\}, k \in \{0, ..., m\}$. Hence, $D[h] \leq |\operatorname{supp}(u)| \max\{u_{pq} \mid (p, q) \in \Omega^2\}$.*

Proof. Because u is nonnegative, $f(\mathcal{S}) \leq h(\mathcal{S})$ for all $\mathcal{S} \subseteq \Omega$. Further, $f(\mathcal{S}) \geq h(\mathcal{S}) - |\operatorname{supp}(u)| \max\{u_{pq} \mid (p, q) \in \Omega^2\}$. Let $\mathcal{A}, \mathcal{B} \subseteq \Omega, s \in \Omega \backslash \mathcal{A}$, where $|\mathcal{A}| = \ell \in \{0, ..., m-1\}$ and $|\mathcal{B}| = k \in \{0, ..., m\}$.

$$h(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - h(\mathcal{A} \cup \mathcal{B}) - h(\mathcal{A} \cup \{s\}) + h(\mathcal{A})$$
$$\leq h(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f(\mathcal{A} \cup \mathcal{B}) - f(\mathcal{A} \cup \{s\}) + h(\mathcal{A})$$
$$\leq f(\mathcal{A} \cup \mathcal{B} \cup \{s\}) - f(\mathcal{A} \cup \mathcal{B}) - f(\mathcal{A} \cup \{s\}) + f(\mathcal{A})$$
$$+ |\operatorname{supp}(u)| \max\{u_{pq} \mid (p, q) \in \Omega^2\}$$
$$\leq |\operatorname{supp}(u)| \max\{u_{pq} \mid (p, q) \in \Omega^2\}.$$

It follows that $d^{\ell, k}[f] \leq |\operatorname{supp}(u)| \max\{u_{pq} \mid (p, q) \in \Omega^2\}$. The bound for $D[h]$ follows immediately. $\qquad\square$

Proposition $B.1$ is an example of how one can use the structure of the approximately submodular function in order to derive bounds for the metrics. This allows one to use these bounds immediately as a substitute for the exact value of the metric and avoid exact computation.

C   Solution Times for Randomly Generated Approximately Submodular Packing Problem Instances

Table C.1 shows solution times for the approximately submodular packing problem instances with and without the addition of the proposed valid inequalities.

No.	Time w/o Ineqs (s)	Time w/ Ineqs (s)	Time Ratio	# of Inequalities Added
1	4.814	2.685	0.558	1244
2	2.780	4.365	1.570	958
3	4.583	5.138	1.121	1040
4	3.080	2.800	0.909	1127
5	27.545	2.938	0.107	1155
6	6.146	5.898	0.960	1456
7	19.199	3.740	0.195	732
8	2.649	3.650	1.378	957
9	9.646	2.853	0.296	853
10	7.166	2.044	0.285	934
11	11.037	9.796	0.888	1018
12	23.986	4.561	0.190	887
13	24.065	5.406	0.225	987
14	32.195	27.333	0.849	1222
15	20.879	5.070	0.243	825
16	9.330	4.820	0.517	1113
17	20.906	4.503	0.215	864
18	9.292	4.153	0.447	1151
19	3.133	10.263	3.276	1133
20	29.028	6.205	0.214	965
21	9.651	3.758	0.389	890
22	24.037	7.667	0.319	1035
23	5.723	4.439	0.776	1020
24	23.952	6.858	0.286	917
25	20.757	7.636	0.368	887
26	19.381	2.596	0.134	693
27	6.565	6.536	0.996	1439
28	19.347	3.131	0.162	938
29	19.252	4.374	0.227	786
30	27.682	4.845	0.175	901
31	19.386	3.967	0.205	755
32	11.525	8.415	0.730	749
33	27.089	10.227	0.378	787
34	5.915	4.911	0.830	959
35	2.870	5.449	1.899	1126
36	2.818	2.337	0.829	902
37	18.283	6.066	0.332	1301
38	4.910	2.463	0.502	1242
39	1.378	1.884	1.367	739
40	10.733	3.375	0.314	826

Table C.1: Solution times for randomly generated approximately submodular packing problems.