# The Complexity of Partial Function Extension for Coverage Functions

**Umang Bhaskar**
Tata Institute of Fundamental Research, Mumbai, India
umang@tifr.res.in

**Gunjan Kumar**
Tata Institute of Fundamental Research, Mumbai, India
gunjan.kumar@tifr.res.in

## Abstract

Coverage functions are an important subclass of submodular functions, finding applications in machine learning, game theory, social networks, and facility location. We study the complexity of partial function extension to coverage functions. That is, given a partial function consisting of a family of subsets of $[m]$ and a value at each point, does there exist a coverage function defined on all subsets of $[m]$ that extends this partial function? Partial function extension is previously studied for other function classes, including boolean functions and convex functions, and is useful in many fields, such as obtaining bounds on learning these function classes.

We show that determining extendibility of a partial function to a coverage function is NP-complete, establishing in the process that there is a polynomial-sized certificate of extendibility. The hardness also gives us a lower bound for learning coverage functions. We then study two natural notions of approximate extension, to account for errors in the data set. The two notions correspond roughly to multiplicative point-wise approximation and additive $L_1$ approximation. We show upper and lower bounds for both notions of approximation. In the second case we obtain nearly tight bounds.

**2012 ACM Subject Classification** Theory of computation → Approximation algorithms analysis

**Keywords and phrases** Coverage Functions, PAC Learning, Approximation Algorithm, Partial Function Extension

**Digital Object Identifier** 10.4230/LIPIcs.APPROX-RANDOM.2019.30

**Category** APPROX

**Related Version** A full version of the paper is available at `https://arxiv.org/abs/1907.07230`.

**Funding** *Umang Bhaskar*: Supported in part by a Ramanujan fellowship and an Early Career Research award.

## 1 Introduction

When can a *partial function* – given as a set of points from a domain, and a value at each point – be extended to a *total function* on the domain, that lies in some particular class of functions? This is the basic question of *partial function extension*, and is studied both independently (such as in convex analysis) and as a recurring subproblem in many areas in combinatorial optimization, including computational learning and property testing.

In this paper we study the computational complexity of partial function extension for *coverage functions*. Coverage functions are a natural and widely-studied subclass of submodular functions that find many applications, including in machine learning [18], auctions [6, 19], influence maximization [8, 22], and plant location [11]. For a natural number $m$, let $[m]$ denote the set $\{1, 2, \dots, m\}$. A set function $f : 2^{[m]} \to \mathbb{R}_+$ is a coverage function if there exists a universe $U$ of elements with non-negative weights and $m$ sets $A_1, \dots, A_m \subseteq U$ such that for all $S \subseteq [m]$, $f(S)$ is the total weight of elements in $\cup_{j \in S} A_j$. A coverage function is succinct if $|U|$ is at most a fixed polynomial in $m$.

The complexity of partial function extension has been studied earlier for other function classes, with a number of important applications shown. For boolean functions, Boros et al. present complexity results for extension to a large number of boolean function classes, as well as results on approximate extension [9]. Pitt and Valiant show a direct relation between the complexity of partial function extension problem and proper PAC-learning. Informally, a class $\mathcal{F}$ of (boolean) functions on $2^{[m]}$ is said to be properly PAC-learnable if for any distribution $\mu$ on $2^{[m]}$ and any small enough $\epsilon > 0$, any function $f^* \in \mathcal{F}$ can be learned by a polynomial-time algorithm that returns a function $f \in \mathcal{F}$ with a polynomial number of samples that differs from $f^*$ with probability at most $\epsilon$. Pitt and Valiant show that if partial function extension for a class $\mathcal{F}$ of functions is NP-hard, then the class $\mathcal{F}$ cannot be PAC-learned unless NP = RP [21].[1] They show computational lower bounds for various classes of boolean functions, thereby obtaining lower bounds on the complexity for learning these classes. In this paper, we show lower bounds on partial function extension for coverage functions, which by this relation give lower bounds on proper PAC learning as well. In separate work, we present results on the computational complexity of partial function extension for submodular, subadditive, and convex functions, and show further connections with learning and property testing [5].

Characterizing partial functions extendible to convex functions is widely studied in convex analysis. Here a partial function is given defined on a non-convex set of points, and is required to be extended to a convex function on the convex hull or some other convex domain. Characterizations for extendible partial functions are given in various papers, such as [12, 26]. This finds many applications, including mechanism design [14], decision making under risk [20], and quantum computation [25].

Another example of the ubiquity of partial function extension is in property testing. Given oracle access to a function $f$, the goal of property testing is to determine by querying the oracle if the function $f$ lies in some class $\mathcal{F}$ of functions of interest, or is far from it, i.e., differs from any function in $\mathcal{F}$ at a large number of points. Partial function extension is a natural step in property testing, since at any time the query algorithm has a partial function consisting of the points queried and the values at those points. If at any time the partial function thus obtained is not extendible to a function in $\mathcal{F}$, the algorithm should reject, and should accept otherwise. Partial function extension is used to give both upper and lower bounds for property testing [5, 23]. Partial function extension is thus a basic problem that finds application in a wide variety of different fields.

## Our Contribution

Our input is a partial function $H = \{(T_1, f_1), \ldots, (T_n, f_n)\}$ with $T_i \subseteq [m]$ and $f_i \geq 0$, and the goal is to determine if there exists a coverage function $f : 2^{[m]} \to \mathbb{R}_{\geq 0}$ such that $f(T_i) = f_i$ for all $i \in [n]$. This is the Coverage Extension problem. Throughout the paper we use $[m]$ for the ground set, $n$ for the number of defined sets in the partial function, and $\mathcal{D}$ for the set of defined sets $\{T_1, \ldots, T_n\}$. We also use $d = \max_{i \in [n]} |T_i|$ to denote the maximum size of sets in $\mathcal{D}$, and $F := \sum_{i \in [n]} f_i$.

Our first result shows that Coverage Extension is NP-hard. Interestingly, we show if there exists a coverage function extending the given partial function then there is an extension by a coverage function for which the size of the universe $|U|$ is at most $n$. This shows that Coverage

[1] Randomized Polynomial (RP) is the class of problems for which a randomized algorithm runs in polynomial time, always answers correctly if the input is a "no" instance, and answers correctly with probability at least 1/2 if the input is a "yes" instance.

Extension is in NP. In contrast, it is known that minimal certificates for non-extendibility may be of exponential size [10]. Also, unlike property testing, this shows that Coverage Extension does not become easier when restricted to succinct coverage functions.

▶ **Theorem 1.** *Coverage Extension is NP-complete.*

For the hardness, we show a reduction from fractional graph colouring, a problem studied in fractional graph theory. Our hardness for extension also shows the following result for proper learning of succinct coverage functions.

▶ **Theorem 2.** *Unless RP = NP, the class of succinct coverage functions cannot be PAC-learned (i.e., cannot be PMAC-learned with approximation factor $\alpha = 1$).*

These are the first hardness results for learning coverage functions based on standard complexity assumptions. Earlier results showed a reduction from learning disjoint DNF formulas to learning coverage functions [13], however as far as we are aware, there are no known lower bounds for learning disjoint DNF formulas. The following theorem is shown in the appendix.

Given the hardness result for Coverage Extension, we study approximation algorithms for two natural optimization versions of the extension problem. In both of these problems, the goal is to determine the distance between the given partial function and the class of coverage functions. Based on the notion of the distance, we study the following two problems.

In *Coverage Approximate Extension*, the goal is to determine minimum value of $\alpha \geq 1$ such that there exists a coverage function $f : 2^{[m]} \to \mathbb{R}_{\geq 0}$ satisfying $f_i \leq f(T_i) \leq \alpha f_i$ for all $i \in [n]$.

In *Coverage Norm Extension*, the goal is to determine the minimum $L_1$ distance from a coverage function, i.e., minimize $\sum_{i\in[n]} |\epsilon_i|$ where $\epsilon_i = f(T_i) - f_i$ for all $i \in [n]$ for some coverage function $f$.

The two notions of approximation we study thus roughly correspond to the two prevalent notions of learning real-valued functions. Coverage Approximate Extension corresponds to PMAC learning, where we look for point-wise multiplicative approximations. Coverage Norm Extension corresponds to minimizing the $L_1$ distance in PAC learning.

Throughout this paper, the minimum value of $\alpha$ in Coverage Approximate Extension will be denoted by $\alpha^*$ and minimum value of $\sum_{i\in[n]} |\epsilon_i|$ in Norm Extension will be denoted by $OPT$. As both of these problems are generalisations of Coverage Extension, they are NP-hard. We give upper and lower bounds for approximation for both of these problem.

▶ **Theorem 3.** *There is a $\big(\min\{d, m^{2/3}\} \log d\big)$-approximation algorithm for Coverage Approximate Extension. If $d$ is a constant then there is a $d$-approximation algorithm.*

In Coverage Norm Extension, $OPT = 0$ iff the partial function is extendible and hence no multiplicative approximation is possible for $OPT$ unless P = NP (because of Theorem 1). We hence consider additive approximations for Coverage Norm Extension. An algorithm for Coverage Norm Extension is called an $\alpha$-approximation algorithm if for all instances (partial functions), the value $\beta$ returned by the algorithm satisfies $OPT \leq \beta \leq OPT + \alpha$. We show nearly tight upper and lower bounds on the hardness of approximation. As defined before $F = \sum_{i\in[n]} f_i$. Note that an $F$-approximation algorithm is trivial, since the function $f = 0$ is coverage and satisfies $\sum_{i\in[n]} |f(T_i) - f_i| \leq F$.

▶ **Theorem 4.** *There is a $(1-1/d)F$-approximation algorithm for Coverage Norm Extension. Moreover, a coverage function $f$ can be efficiently computed such that $\sum_{i\in[n]} |f(T_i) - f_i| \leq OPT + (1 - 1/d)F$.*

▶ **Theorem 5.** *It is NP-hard to approximate Coverage Norm Extension by a factor* $\alpha = 2^{poly(n,m)} F^{\delta}$ *for any fixed* $0 \leq \delta < 1$*. This holds even when* $d = 2$*.*

Our lower bound is roughly based on the equivalence of *validity* and *membership*, where given a convex, compact set $K$, the validity problem is to determine the optimal value of $c^T x$ given a vector $c$ over all $x \in K$, while the membership problem seeks to determine if a given point $x$ is in $K$ or not. The equivalence of optimization and separation is a widely used tool. The reduction from optimization to separation is particularly useful for, e.g., solving linear programs with exponential constraints. Our work is unusual in both the use of validity and membership rather than optimization and separation, and because of the direction – we use the equivalence to show hardness of the validity problem. We hope that our techniques may be useful in future work as well.

## Related Work

We focus here on work related to partial function extension and coverage functions. In a separate paper, we study partial function extension to submodular, subadditive, and convex functions, showing results on the complexity as well as applications to learning and property testing [5]. Previously, Seshadri and Vondrak [23] introduce the problem of extending partial functions to a submodular function, and note its usefulness in analyzing property testing algorithms. For submodular functions, partial function extension is also useful in optimization [24]. The problem of extending a partial function to a convex function is also studied in convex analysis [26, 12]. As mentioned earlier, both characterizing extendible partial functions, and the complexity of partial function extension has been studied for large classes of Boolean functions [9, 21].

Chakrabarty and Huang study property testing for coverage functions [10]. Here, the goal is to determine whether the input function (given by an oracle) is coverage or far from coverage by querying an oracle, where distance is measured by the number of points at which the function must be changed for it to be coverage. They show that succinct coverage functions can be reconstructed with a polynomial number of queries and hence can be efficiently tested. However, they conjecture that testing general coverage functions requires $2^{\Omega(m)}$ queries, and prove this lower bound under a different notion of distance. They present a particular characterization of coverage functions in terms of the $W$-transform that we use as well.

There has also been interest in sketching and learning coverage functions. Badanidiyuru et al. [1] showed that coverage functions admit a $(1+\epsilon)$-sketch, i.e., given any coverage function, there exists a succinct coverage function (of size polynomial in $m$ and $1/\epsilon$) that approximates the original function within $(1+\epsilon)$ factor with high probability. Feldman and Kothari [13] gave a fully polynomial time algorithm for learning succinct coverage functions in the PMAC model if the distribution is uniform. However, if the distribution is unknown, they show learning coverage functions is as hard as learning polynomial size DNF formulas for which no efficient algorithm is known.

Balkanski et al [3] study whether coverage functions can be optimized from samples. They consider a scenario where random samples $\{(S_i, f(S_i))\}$ of an unknown coverage function $f$ are provided and ask if it is possible to optimize $f$ under a cardinality constraint, i.e., solve $\max_{S:|S|\leq k} f(S)$. They prove a negative result: no algorithm can achieve approximation ratio better than $2^{\Omega(\sqrt{\log m})}$ with a polynomial number of sampled points.

## 2 Preliminaries

As earlier, for $m \in \mathbb{Z}_+$, define $[m] := \{1, 2, \ldots, m\}$. A set function $f$ over a ground set $[m]$ is a coverage function if there exists a universe $U$ of elements with non-negative weights and $m$ sets $A_1, ..., A_m \subseteq U$, such that for all $S \subseteq [m]$, $f(S)$ is the total weight of elements in $\cup_{j \in S} A_j$. A coverage function is *succinct* if $|U|$ is at most a fixed polynomial in $m$.

Chakrabarty and Huang [10] characterize coverage functions in terms of their $W$-transform, which we use as well. For a set function $f : 2^{[m]} \to \mathbb{R}_{\geq 0}$, the $W$-transform $w : 2^{[m]} \setminus \emptyset \to \mathbb{R}$ is defined as

$$\forall S \in 2^{[m]} \setminus \emptyset, \quad w(S) = \sum_{T: S \cup T = [m]} (-1)^{|S \cap T|+1} f(T) \,. \tag{1}$$

The set $\{w(S) | S \in 2^{[m]} \setminus \emptyset\}$ is called the set of $W$-coefficients of $f$. We can also recover the function $f$ from its $W$-coefficients.

$$\forall T \subseteq [m], \quad f(T) = \sum_{S \subseteq [m]: S \cap T \neq \emptyset} w(S) \,. \tag{2}$$

If $f$ is a coverage function induced by the universe $U$ and sets $A_1, \ldots, A_m$, then the $W$-transform $w(S)$ is precisely the weight of the set $\{(\cap_{i \in S} A_i) \setminus \cup_{j \notin S} A_j\}$, and is hence non-negative. The converse is also true. The set $\{S | w(S) > 0\}$ is the called the *support* of the coverage function, and the elements are exactly the elements of the universe $U$.

▶ **Theorem 6** ([10]). *A set function $f : 2^{[m]} \to \mathbb{R}_{\geq 0}$ is a coverage function iff all of its $W$-coefficients are non-negative.*

From Theorem 6, given a partial function $H$, there exists a coverage function $f$ satisfying $f(T_i) = f_i$ for all $i \in [n]$ iff the following linear program is feasible, where the variables are the $W$-coefficients $w(S)$ for all $S \in 2^{[m]} \setminus \emptyset$.

$$\text{Extension-P:} \qquad \sum_{S: S \cap T_i \neq \emptyset} w(S) = f_i \quad \forall i \in [n] \,, \qquad w(S) \geq 0 \quad \forall S \in 2^{[m]} \setminus \emptyset.$$

All missing proofs are in the appendix.

## 3 Coverage Extension and PAC-Learning

Our first observation is that there is a polynomial-sized certificate of extendibility to a coverage function. This is obtained by observing that at a vertex of the feasible set in Extension-P, at most $n$ of the variables are non-zero. It is interesting to compare this with Chakrabarty and Huang [10], who give an example to show that minimal certificates of nonextendibility may be of exponential size.

▶ **Proposition 7.** *If a partial function is extendible to a coverage function, then it is also extendible to a coverage function with support size $\leq n$. Hence, Coverage Extension is in NP.*

We show the NP-hardness of Coverage Extension by reduction from *fractional chromatic number*, defined as follows. Given a graph $G = (V, E)$, a set $I \subseteq V$ is called an *independent set* if no two vertices in $I$ are adjacent. Let $\mathcal{I}$ be the set of all independent sets. The fractional chromatic number $\chi^*(G)$ of a graph $G$ is the optimal value of the following linear program.

$$\chi^*(G) := \left\{ \min \sum_{I \in \mathcal{I}} x_I : \sum_{I \in \mathcal{I}: v \in I} x_I \geq 1 \quad \forall v \in V(G),\, 0 \leq x_I \leq 1 \quad \forall I \in \mathcal{I} \right\}$$

Note that if $x_I \in \{0,1\}$ then the optimal value is just the chromatic number of the graph.[2]

▶ **Theorem 8** ([15]). *For graph $G = (V, E)$, there exist nonnegative weights $\{x_I\}_{I \in \mathcal{I}}$ on independent sets such that $\chi^*(G) = \sum_{I \in \mathcal{I}} x_I$ and $\sum_{I \in \mathcal{I}: v \in I} x_I = 1 \quad \forall v \in V$.*

▶ **Corollary 9.** *For graph $G = (V, E)$ and for any value of $t$ such that $\chi^*(G) \leq t \leq |V|$, there exist nonnegative weights $\{z_I\}_{I \in \mathcal{I}}$ on independent sets such that $\sum_{I \in \mathcal{I}} z_I = t$ and $\sum_{I \in \mathcal{I}: v \in I} z_I = 1 \quad \forall v \in V$.*

▶ **Theorem 10** ([17]). *Given graph $G = (V, E)$ and $1 \leq k \leq |V|$, it is NP-hard to determine if $\chi^*(G) \leq k$.*

We now show the NP-hardness of Coverage Extension.

**Proof of Theorem 1.** Since membership in NP was shown earlier, we give the reduction from fractional chromatic number. The input is a graph $G = (V, E)$ and $1 \leq k \leq |V|$.

We identify $[n']$ with the set of vertices $V$, and therefore $E(G) \subseteq \{\{i, j\} | i, j \in [n']\}$, and any set $S \subseteq [n']$ can be viewed as a set of vertices. The partial function construction is as follows. The ground set is $[n']$ and therefore $m = n'$. The partial function is defined at all vertices, all edges, and the set consisting of all vertices. Hence $\mathcal{D}$, the set of defined points for the partial function, is $\{\{i\} | i \in [n']\} \cup E(G) \cup \{[n']\}$ and $|\mathcal{D}| = n' + |E(G)| + 1$. The value of the partial function $h$ at these defined sets is given by

$$h(S) = \begin{cases} 1 & \text{if} \quad S = \{i\},\, i \in [n']\,, \\ 2 & \text{if} \quad S \in E(G)\,, \\ k & \text{if} \quad S = \{[n']\}\,. \end{cases}$$

Intuitively, the function $h(S)$ can be interpreted as the (fractional) number of colours used to colour the subset $S$.

We claim that $\chi^*(G) \leq k$ iff the above partial function is extendible. Suppose $\chi^*(G) \leq k$. Therefore by Corollary 9, there exist nonnegative weights $\{x_I\}_{I \in \mathcal{I}}$ such that $\sum_{I \in \mathcal{I}} x_I = k$ and $\sum_{I \in \mathcal{I}: v \in I} x_I = 1 \quad \forall v \in V(G)$. For all $S \in 2^{[m]} \setminus \emptyset$, define the function $w(S)$ as $x_S$ if $S \in \mathcal{I}$ and 0 otherwise. Since $w(S) \geq 0$, this defines the $W$-transform for a coverage function $g$. We have, for any $i \in [n']$,

$$g(\{i\}) = \sum_{S: S \cap \{i\} \neq \emptyset} w(S) = \sum_{I \in \mathcal{I}: i \in I} x_I = 1\,,$$

for any $\{i, j\} \in E(G)$,

$$g(\{i, j\}) = \sum_{S: S \cap \{i,j\} \neq \emptyset} w(S) = \sum_{I \in \mathcal{I}: i \in I} x_I + \sum_{I \in \mathcal{I}: j \in I} x_I = 2$$

as no independent set $I$ can contain both $i$ and $j$; and finally $g(\{[n']\}) = \sum_{S: S \cap \{[n']\} \neq \emptyset} w(S) = \sum_{I \in \mathcal{I}} x_I = k$. Therefore $g$ is an extension of the above partial function $h$.

[2] The chromatic number of a graph is the minimum number of colours required to colour the vertices so that no two adjacent vertices get the same colour.

Now suppose there is an extension, i.e., there exists $w(S) \geq 0$ for all $S \in 2^{[m]} \setminus \emptyset$ such that for any $i \in [n']$, $\sum_{S:S\cap\{i\}\neq\emptyset} w(S) = 1$; for any $\{i,j\} \in E(G), \sum_{S:S\cap\{i,j\}\neq\emptyset} w(S) = 2$; and finally $\sum_{S:S\cap\{[n']\}\neq\emptyset} w(S) = k$. For any $\{i,j\} \in E(G)$, we have

$$\sum_{S:S\cap\{i,j\}\neq\emptyset} w(S) = \sum_{S:S\cap\{i\}\neq\emptyset} w(S) + \sum_{S:S\cap\{j\}\neq\emptyset} w(S) - \sum_{S:S\supseteq\{i,j\}} w(S)\,.$$

Therefore,$\sum_{S:S\supseteq\{i,j\}} w(S) = 0$, i.e., if $w(S) > 0$ then $S$ must be an independent set. It now follows that $\chi^*(G) \leq \sum_{S:S\cap\{[n']\}\neq\emptyset} w(S) = k$. ◀

## Proper PAC-learning of Coverage functions

We now prove Theorem 2. We first recall the definition of proper PAC-learning.

▶ **Definition 11** ([2]). *An algorithm $\mathcal{A}$ properly PAC-learns a family of functions $\mathcal{F}$, if for* any *distribution $\mu$ (on $2^{[m]}$) and* any *target function $f^* \in \mathcal{F}$, and for* any *sufficiently small $\epsilon, \delta > 0$:*

1. *$\mathcal{A}$ takes the sequence $\{(S_i, f^*(S_i))\}_{1\leq i\leq l}$ as input where $l$ is poly$(m, 1/\delta, 1/\epsilon)$ and the sequence $\{S_i\}_{1\leq i\leq l}$ is drawn i.i.d. from the distribution $\mu$,*
2. *$\mathcal{A}$ runs in poly$(m, 1/\delta, 1/\epsilon)$ time, and*
3. *$\mathcal{A}$ returns a function $f : 2^{[m]} \to \mathbb{R} \in \mathcal{F}$ such that*

$$Pr_{S_1,\ldots,S_l\sim\mu}\big[Pr_{S\sim\mu}[f(S) = f^*(S)] \geq 1 - \epsilon\big] \geq 1 - \delta$$

We use the reconstruction algorithm for coverage functions given by Chakrabarty and Huang [10] in our proof. Given a coverage function $f$ as an input, this reconstruction algorithm terminates in $O(m\,s)$ steps where $s$ is the support size of $f$, i.e., the number of non-zero $W$-coefficients of $f$, and returns these non-zero $W$-coefficients.

Recall the reduction from fractional chromatic number to Coverage Extension (Theorem 1). Given an instance of fractional chromatic number (graph $G = (V, E)$ and rational $k'$ with $|V| = n'$), the instance of Coverage Extension is a set of defined points $\mathcal{D} = \{\{i\}|i \in [n']\} \cup E(G) \cup \{[n']\}$ and a function $h$ on $\mathcal{D}$. Let $k = |\mathcal{D}| = |V| + |E| + 1$. From Theorem 1 and Proposition 7, $\chi^*(G) \leq k'$ iff $h$ is extendible to a coverage function with support size at most $k$.

Let $\mathcal{F}$ be a family of coverage functions with support size at most $k$. Let $\epsilon = 1/k^3$ (and hence $\epsilon < 1/|\mathcal{D}|$) and $\mu$ be a uniform distribution over $\{(S, h(S))|S \in \mathcal{D}\}$. Now suppose a (randomized) algorithm $A$ properly PAC-learns $\mathcal{F}$. We will show that in this case, we can determine efficiently if the partial function is extendible to a coverage function, and hence RP = NP.

Suppose the algorithm $A$ returns a function $g$. If the partial function is extendible then there exists a function in $\mathcal{F}$ that has the same value on samples seen by $A$. Therefore, if the partial function is extendible then $g(S)$ must be equal to $h(S)$ for all $S \in \mathcal{D}$ (since $\epsilon < 1/|\mathcal{D}|$ and $A$ must satisfy$Pr_{S\sim D^*}[f(S) = f^*(S)] \geq 1 - \epsilon$). We run the reconstruction algorithm on input $g$. If the partial function is extendible then $g$ must be in $\mathcal{F}$ and hence the reconstruction algorithm must terminate in $O(mk)$ steps. Further, if $\{w(S)\}_{S\in\mathcal{S}}$ is the output of the algorithm then (i) $w(S) > 0$ for all $S \in \mathcal{S}$, (ii) $|\mathcal{S}| \leq k$ (iii) the coverage function $f'$ given by the $W$-coefficients $w'(S) = w(S)$ if $S \in \mathcal{S}$ and 0 otherwise is an extension of the partial function $h$. Condition (iii) should hold because $f'$ must be the same as $g$ which we have shown earlier is an extension of $h$.

The converse is also true – if the reconstruction algorithm terminates and (i), (ii), (iii) hold then clearly $h$ is extendible (by $f'$). Since all the steps require polynomial time to check, we can efficiently determine if the partial function is extendible.

## 4 Coverage Approximate Extension

We now build the framework for Theorem 3. We start with the following lemma.

▶ **Lemma 12.** *Given a partial function $H$ and $\alpha \geq 1$, there is no coverage function $f$ satisfying $f_i \leq f(T_i) \leq \alpha f_i$ for all $i \in [n]$ iff the following program, with variables $l_i$ for all $i \in [n]$ is feasible:*

$$-\alpha \sum_{i:l_i<0} f_i l_i < \sum_{i:l_i>0} f_i l_i \tag{3}$$

$$\sum_{i:S\cap T_i \neq \emptyset} l_i \leq 0 \quad \forall S \subseteq [m] \tag{4}$$

Thus the optimal approximation ratio $\alpha^*$ is the minimum value of $\alpha$ for which (3) and (4) are not feasible together.

A natural representation of the partial function $H = \{(T_1, f_1), \dots, (T_n, f_n)\}$ is as a weighted bipartite graph $H = (A \cup [m], E)$ with $|A| = n$, and an edge between $a_i \in A$ and $j \in [m]$ if the set $T_i$ contains element $j \in [m]$. Each vertex $a_i \in A$ also has weight $f_i$. Then $d = \max_i |T_i|$ is the maximum degree of any vertex in $A$. For the remainder of this section, we will use this representation of partial functions.

We use the following notation given a bipartite graph $H = (A \cup [m], E)$. For any $S \subseteq [m]$, let $N(S) = \{v \in A : (v, j) \in E \quad \text{for some} \quad j \in S\}$ be the set of neighbours of set $S$. Similarly for set $R \subseteq A$, $N(R) = \{j \in [m] : (v, j) \in E \quad \text{for some} \quad v \in R\}$ be the set of neighbours of set $R$. For any vertex $v$ in $H$, we use $N(v)$ for $N(\{v\})$. In this bipartite graph representation, the inequality (4) is equivalent to $\sum_{i \in N(S)} l_i \leq 0$ for all $S \subseteq [m]$.

We now define a parameter $\kappa$ called the *replacement ratio* for a partial function $H$.

▶ **Definition 13.** *Let $H = (A \cup [m], E)$ be a bipartite graph with weights $f_v$ on each $v \in A$. For $v \in A$, let $\mathcal{F}_v = \{R \subseteq A \setminus \{v\} \mid N(R) \supseteq N(v)\}$ be the set of all subsets of $A \setminus \{v\}$ that cover all the neighbours of $v$. We call each $R \in \mathcal{F}_v$ a* replacement *for $v$. The* replacement ratio *$\kappa$ is then the minimum of $\frac{\sum_{w \in R} f_w}{f_v}$ over all vertices $v \in A$ and replacements $R \in \mathcal{F}_v$.*

The proof of the upper bound in Theorem 3 will follow from the bounds on $\alpha^*$ shown in Lemma 14, 16 and 18.

▶ **Lemma 14.** *For any partial function $H$, $\alpha^* \geq \frac{1}{\kappa}$.*

**Proof.** By definition of $\kappa$, there exists a vertex $v \in A$ and a replacement $R$ for $v$ such that $\sum_{w \in R} f_w = c f_v$. Note that setting $l_w = -1 \quad \forall w \in R$, $l_v = 1$ and all other $l_w$'s to be zero results in feasibility of the inequalities $\sum_{w \in N(S)} l_w \leq 0$ for all $S \in 2^{[m]} \setminus \emptyset$. From the definition of $\alpha^*$ and Lemma 12, $\alpha^* \sum_{w \in R} f_w \geq f_v$, and hence $\alpha^* \geq 1/\kappa$. ◀

Let $\beta = \frac{\min\{d, m^{2/3}\}}{\kappa}$. Given values $\{l_v\}_{v \in A}$ on the vertices in $A$ such that $\sum_{v \in N(S)} l_v \leq 0$ for all $S \subseteq [m]$, we will show that $\beta \sum_{v:l_v<0} f_v l_v \geq \sum_{v:l_v>0} f_v l_v$ and hence $\alpha^* \leq \beta$. If $l_v = 0$ for any vertex, we simply ignore such a vertex, since it does not affect either (4) or (3).

By scaling, we can assume that $l_v \in \mathbb{Z}$ for all $v \in A$. At some point, we will use Hall's theorem to show a perfect matching. To simplify exposition, we replace each $v \in A$ with $|l_v|$ identical copies, each of which is adjacent to the same vertices as $v$. Each such new vertex $v'$ has $l_{v'} = 1$ if $l_v > 0$ and $l_{v'} = -1$ if $l_v < 0$, and $f_{v'} = f_v$. Let the new bipartite graph be $H' = (A' \cup [m], E')$. It is easy to check that in the new bipartite graph, the degree of vertices in $A'$ and the values $\kappa$, $\sum_{v \in A':l_v>0} f_v l_v$, $\sum_{v \in A':l_v<0} f_v l_v$ and $\sum_{v \in N(S)} l_v$ remain unchanged for all $S \subseteq [m]$.

Let $\mathcal{N} = \{v \in A' | l_v = -1\}$ and $\mathcal{P} = \{v \in A' | l_v = 1\}$, and let $E^-$ be the set of edges with one end-point in $\mathcal{N}$, while $E^+$ are the edges with one end-point in $\mathcal{P}$. For any $S \subseteq [m]$, let $N^-(S) = N(S) \cap \mathcal{N}$ and $N^+(S) = N(S) \cap \mathcal{P}$ (so $N(S) = N^+(S) \cup N^-(S)$). Finally, define $E^+(S)$ ($E^-(S)$) as the set of edges with one end-point in $S$ and the other end-point in $\mathcal{P}$ ($\mathcal{P}$). If $S = \{j\}$, we abuse notation slightly and use $N^-(j)$, $N^+(j)$, $E^-(j)$ and $E^+(j)$. Note that $|N^-(S)| \geq |N^+(S)|$ for all $S \subseteq [m]$ in $H'$, since in $H$, $\sum_{v \in N(S)} l_v \leq 0$ for all $S \subseteq [m]$. Our goal is to show $\beta \sum_{v \in \mathcal{N}} f_v \geq \sum_{v \in \mathcal{P}} f_v$.

▶ **Lemma 15.** *Suppose for some $\beta' \geq 1$, $\beta' |N^-(S)| \geq \sum_{j \in S} |N^+(j)|$ for all $S \subseteq [m]$. Then for each vertex $v \in \mathcal{P}$, there exists a replacement $F_v \subseteq \mathcal{N}$ such that each vertex in $\mathcal{N}$ is contained in $F_v$ for at most $\beta'$ vertices $v \in \mathcal{P}$. Hence, $\beta' \sum_{v \in \mathcal{N}} f_v \geq \kappa \sum_{v \in \mathcal{P}} f_v$ and so $\alpha^* \leq \frac{\beta'}{\kappa}$.*

**Proof.** By Hall's theorem, there exists a set of edges $M \subseteq E^-$ such that (i) the degree in $M$ of each vertex $j \in [m]$ is at least $|N^+(j)|$, and (ii) the degree in $M$ of each vertex $v \in \mathcal{N}$ is at most $\beta'$. Because of (i), for each $j \in [m]$ there is an injection $h_j$ from edges in $E^+(j)$ to edges in $E^-(j) \cap M$, i.e., each edge in $E^+(j)$ maps to a distinct edge in $E^-(j) \cap M$. Now for a vertex $v \in \mathcal{P}$, consider a neighbouring vertex $j \in N(v)$. Each such edge $(v, j)$ is in $E^+(j)$, and is hence mapped by $h_j$ to an edge in $E^-(j) \cap M$. Let $F_v$ be the end-points in $\mathcal{N}$ of these mapped edges. That is, $w \in F_v$ iff there exists $j \in N(v)$ such that $(w, j) = h_j(v, j)$. Then $F_v$ is a replacement for $v$, and hence, $\sum_{w \in F_v} f_w \geq \kappa f_v$. Further, because of (ii), and since each $h_j$ is an injection, each vertex in $\mathcal{N}$ is contained in $F_v$ for at most $\beta'$ vertices $v \in \mathcal{P}$. Then summing the inequality $\sum_{w \in F_v} f_w \geq \kappa f_v$ over all $v \in \mathcal{P}$, we get that $\beta' \sum_{v \in \mathcal{N}} f_v \geq \kappa \sum_{v \in \mathcal{P}} f_v$ as required. ◀

▶ **Lemma 16.** *For any partial function $H$, $\alpha^* \leq \frac{d}{\kappa}$.*

**Proof.** Fix $S \subseteq [m]$. Since $|N^-(j)| \geq |N^+(j)|$ for all $j \in [m]$, $\sum_{j \in S} |N^-(j)| \geq \sum_{j \in S} |N^+(j)|$, and since $d$ is the maximum degree of any vertex in $A'$, $d|N^-(S)| \geq \sum_{j \in S} |N^-(j)|$. The proof follows from Lemma 15. ◀

If we can show $m^{2/3}|N^-(S)| \geq \sum_{j \in S} |N^+(j)|$ for all $S \subseteq [m]$ then by Lemma 15, $\alpha^* \leq \frac{m^{2/3}}{\kappa}$. Unfortunately this may not be true. Let $\mathcal{N} = \{v_1\}, \mathcal{P} = \{v_2\}, E^- = \{(v_1, j) | j \in [m]\}$ and $E^+ = \{(v_2, j) | j \in [m]\}$. Note that $\sum_{j \in [m]} |N^+(j)| = m$ whereas $|N^-([m])| = 1$. Notice that in this bad example, the bipartite graph contains a 4-cycle $v_1, j_1, v_2, j_2, v_1$ where $v_1 \in \mathcal{N}$ and $v_2 \in \mathcal{P}$. We now define a subgraph called a *diamond* which generalises such a 4-cycle. A diamond $(v_p, v_n, J)$ of size $k$ is a subgraph of $H'$ where $v_p \in \mathcal{P}, v_n \in \mathcal{N}, J \subseteq [m]$ ($|J| = k$) such that for all $j \in J$, both $(v_p, j)$ and $(v_n, j)$ are contained in $E'$. Note that a 4-cycle is a diamond of size two (and the bad example considered above is a diamond of size $m$).

Let $k_{max} = m^a$ $(0 \leq a \leq 1)$ be the maximum size of any diamond in $H'$.

▶ **Lemma 17.** *For all $S \subseteq [m]$, $m^{\frac{1+a}{2}} |N^-(S)| \geq \sum_{j \in S} |N^+(j)|$, where $m^a$ is the size of the largest diamond in $H'$.*

**Proof.** Recall that for all $j \in [m]$, $|N^+(j)| \leq |N^-(j)|$, hence there is an injection $h_j$ from $N^+(j)$ to $N^-(j)$, i.e, $h_j$ maps each vertex in $N^+(j)$ to a unique vertex in $N^-(j)$. Fix $S \subseteq [m]$ and vertex $v \in \mathcal{P}$, and let $S_v := N(v) \cap S$ be the neighbourhood of $v$ in $S$. We will consider $N^+(S_v)$ and $N^-(S_v)$, the negative and positive neighbourhoods of $S_v$. Note that since all vertices in $S_v$ are adjacent to $v \in \mathcal{P}$, a vertex in $N^-(S_v)$ is adjacent to at most $m^a$ vertices in $S_v$, by definition of $a$. Thus for a vertex $v' \in N^-(S_v)$, there are at most $m^a$ different vertices $j \in S_v$ for which $h_j$ maps a vertex in $N^+(j)$ to $v'$, and hence $m^a|N^-(S_v)| \geq \sum_{j \in S_v} |N^+(j)|$.

Now if there is a vertex $v \in \mathcal{P}$ such that $m^{\frac{1-a}{2}} \sum_{j \in S_v} |N^+(j)| \geq \sum_{j \in S} |N^+(j)|$ then we are done, since

$$|N^-(S)| \geq |N^-(S_v)| \geq \frac{\sum_{j \in S_v} |N^+(j)|}{m^a} \geq \frac{\sum_{j \in S} |N^+(j)|}{m^{\frac{1+a}{2}}} .$$

So assume that for all $v \in \mathcal{P}$, $\sum_{j \in S_v} |N^+(j)| \leq \frac{\sum_{j \in S} |N^+(j)|}{m^{\frac{1-a}{2}}}$. In this case, note that by reversing the order of summation,

$$\sum_{j \in S} |N^+(j)|^2 = \sum_{j \in S} \sum_{v \in N^+(j)} |N^+(j)| = \sum_{v \in N^+(S)} \sum_{j \in S_v} |N^+(j)| \leq |N^+(S)| \frac{\sum_{j \in S} |N^+(j)|}{m^{\frac{1-a}{2}}} .$$

Therefore, using the above inequality for $|N^+(S)|$,

$$|N^-(S)| \geq |N^+(S)| \geq m^{\frac{1-a}{2}} \frac{\sum_{j \in S} |N^+(j)|^2}{\sum_{j \in S} |N^+(j)|} \geq \frac{m^{\frac{1-a}{2}}}{|S|} \frac{\left(\sum_{j \in S} |N^+(j)|\right)^2}{\sum_{j \in S} |N^+(j)|} \geq \frac{\sum_{j \in S} |N^+(j)|}{m^{\frac{1+a}{2}}}$$

as required by the lemma. The third inequality follows from Cauchy-Schwarz. ◀

From Lemmas 15 and 17, if $a \leq 1/3$ then $\alpha^* \leq \frac{m^{2/3}}{\kappa}$. Next we show this is true in general.

▶ **Lemma 18.** *For any partial function $H$, $\alpha^* \leq \frac{m^{2/3}}{\kappa}$.*

**Proof.** If $k_{max} \leq m^{1/3}$ then by Lemma 17 and 15, $\alpha^* \leq \frac{m^{2/3}}{\kappa}$. So we assume $k_{max} > m^{1/3}$. In this case, we pick a diamond $(v_p, v_n, J)$ of size $> m^{1/3}$. We remove, for all $j \in J$, the edges $(v_p, j)$ and $(v_n, j)$. We repeat the above procedure (in the new graph) until we are left with a bipartite graph where all diamonds are of size at most $m^{1/3}$. Note that if a diamond $(v_p, v_n, J)$ of size $k$ is removed then the degree of $v_n$ decreases by $k$. Hence, for a fixed vertex $v_n$, number of removed diamonds is at most $m^{2/3}$ (as at any step we remove diamonds of size at least $m^{1/3}$). It is easy to see that after every step, $|N^-(S)| \geq |N^+(S)|$ (for all $S \in 2^{[m]} \setminus \emptyset$) still holds in the bipartite graph. Let $H^*$ be the bipartite graph at the end (all diamonds of size at most $m^{1/3}$).

Note that we do not remove any vertex in the above procedure. Fix vertex $v \in \mathcal{P}$. By Lemmas 17 and 15 with $a = 1/3$, there exists $F_v \subseteq \mathcal{N}$ such that $F_v$ covers all neighbours of $v$ in $H^*$ and each vertex in $\mathcal{N}$ is contained in $F_v$ for at most $m^{2/3}$ vertices $v \in \mathcal{P}$. Since we have removed edges, $F_v$ may not cover all the neighbours of $v$ in $H'$. Let $v^1, \ldots, v^s \in \mathcal{N}$ be the set of all vertices such that for each $i \in [s]$, a diamond $(v, v^i, J^i)$ was removed in a removal step. Clearly $\{v^1, \ldots, v^s\} \cup F_v$ cover all the neighbour of $v$ in $H'$. Therefore, we have $\sum_{i=1}^{s} f_{v^i} + \sum_{w \in F_v} f_w \geq \kappa f_v$. Since any $v^i$ $(1 \leq i \leq s)$ is a part of at most $m^{2/3}$ removed diamonds and each vertex in $\mathcal{N}$ is contained in $F_v$ for at most $m^{2/3}$ vertices $v \in \mathcal{P}$ , summing the above inequality for each $v \in \mathcal{P}$, we get $m^{2/3} \sum_{v \in \mathcal{N}} f_v \geq \kappa \sum_{v \in \mathcal{P}} f_v$ as required. ◀

It follows from Lemmas 14, 16 and 18 that an algorithm that returns $\frac{\min\{d, m^{2/3}\}}{\kappa}$ is a $\min\{d, m^{2/3}\}$-approximation algorithm. However, computing $\kappa$ corresponds to solving a general set cover instance, and is NP-hard. This connection however allows us to show the following result.

▶ **Lemma 19.** *Given a partial function, the replacement ratio $\kappa$ can be efficiently approximated by $\kappa'$ such that $\kappa \leq \kappa' \leq \kappa \log d$. If $d$ is a constant, the replacement ratio $\kappa$ can be determined efficiently.*

This completes the proof of the upper bound in Theorem 3. In the full version of the paper, we show there exist partial functions such that (i) $\alpha^* = 1/\kappa$ for any value of $\kappa$, and (ii) with $d = \sqrt{m}$ and $\alpha^* = \Omega(\frac{\sqrt{m}}{\kappa \log m})$. The bounds shown on $\alpha^*$ thus cannot be substantially improved.

## 5 Coverage Norm Extension

From Theorem 6, the Norm Extension problem can be stated as the convex program Norm-P. It can be equivalently transformed to a linear program whose dual is Norm-D.

Norm-P: $\min \sum_{i=1}^{n} |\epsilon_i|$

$$\sum_{S: S\cap T_i \neq \emptyset} w(S) = f_i + \epsilon_i \quad \forall i \in [n]$$

$$w(S) \geq 0 \quad \forall S \in 2^{[m]} \setminus \emptyset$$

Norm-D: $\max \sum_{i=1}^{n} f_i y_i$

$$\sum_{i: S\cap T_i \neq \emptyset} y_i \leq 0 \quad \forall S \in 2^{[m]} \setminus \emptyset \tag{5}$$

$$-1 \leq y_i \leq 1 \quad \forall i \in [n] \tag{6}$$

Both Norm-P and Norm-D are clearly feasible. We use $OPT$ for the optimal value of Norm-P (and Norm-D). As stated earlier, no multiplicative approximation is possible for $OPT$ unless P = NP. Therefore, we consider additive approximations for Norm Extension.

An algorithm for Norm Extension is called an $\alpha$-approximation algorithm if for all instances (partial functions), the value $\beta$ returned by the algorithm satisfies $OPT \leq \beta \leq OPT + \alpha$. First we prove our upper bound in Theorem 4. Recall that $d = \max_{i\in[n]} |T_i|$ and $F = \sum_{i\in[n]} f_i$. As noted earlier, the function $f(\cdot) = 0$ is trivially an $F$-approximation algorithm for Norm Extension, since $\sum_{i\in[n]} |f(T_i) - f_i| = F$.

**Proof of Theorem 4.** Consider the linear programs obtained by restricting Norm-P to variables $w(S)$ for $S \in [m]$, and similarly restricting the constraints (5) in Norm-D to sets $S \in [m]$ only. They are clearly the primal and dual of each other. The optimal values of these modified problems (say $OPT^R$, $w^R$ and $y^R$) can be computed in polynomial time. We will show that $OPT \leq OPT^R \leq OPT + (1 - 1/d)F$ for the proof of the theorem. The first inequality is obvious, since $OPT^R$ is the optimal solution to a relaxed (dual) linear program.

For the second inequality, define $y^A = (y_1^A, \ldots, y_n^A)$ as the vector such that for all $i \in [n]$, $y_i^A = y_i^R$ if $y_i^R \leq 0$ and $y_i^R/d$ otherwise. Then note that

$$OPT^R = \sum_{i\in[n]} f_i y_i^R = \sum_{i\in[n]} f_i y_i^A + (1 - 1/d) \sum_{i: y_i^R \geq 0} f_i y_i^R \leq \sum_{i\in[n]} f_i y_i^A + (1 - 1/d)F\,, \tag{7}$$

where the last inequality is because each $y_i^R \leq 1$. We now show that $y^A$ is a feasible solution for Norm-D, and hence $\sum_{i\in[n]} f_i y_i^A \leq OPT$. Together with (7) this completes the proof.

Clearly $y^A$ satisfies the constraints (6). We will show that $y^A$ also satisfies the constraints (5) for all $S \in 2^{[m]} \setminus \emptyset$. Consider any $S \in 2^{[m]} \setminus \emptyset$. Let $P = \{i \in [n] | S \cap T_i \neq \emptyset \quad \text{and} \quad y_i^R > 0\}$ and $N = \{i \in [n] | S \cap T_i \neq \emptyset \quad \text{and} \quad y_i^R \leq 0\}$. Thus $P \cup N$ are all sets in $\mathcal{D}$ that have nonempty intersection with $S$. We have for any $j \in S$ that $\sum_{i: j\in T_i} y_i^R \leq 0$. Summing these inequalities over $j \in S$, we obtain $\sum_{i\in P\cup N} |T_i \cap S| y_i^R \leq 0$. Thus $\sum_{i\in P} y_i^R + d \sum_{i\in N} y_i^R \leq 0$. From the definition of $y_i^A$, we get $\sum_{i: S\cap T_i \neq \emptyset} y_i^A \leq 0$, as required. ◀

We now prove the lower bound in Theorem 4. We start with an outline of the proof. In a nutshell, the proof shows the following reductions (for brevity, WM stands for Weak Membership and WV for Weak Validity):

$$\text{Densest-Cut} \leq_p \text{Cut WM} \leq_p \text{Span WM} \equiv \text{Coverage WM} \leq_p \text{Coverage WV} \\ \leq_p \text{Norm Extension}\,.$$

Given a graph $G = (V, E)$ and a positive rational $M$, the *Densest-Cut* problem asks if there is a cut $S \subset V$ such that $\frac{|\delta(S)|}{|S|\,|V \setminus S|} > M$. The Densest-Cut problem is known to be NP-hard [7], and ultimately we reduce the Densest-Cut problem to the problem of approximating the optimal value for Norm-P. We formally define the other problems later. However, to show this reduction, we need to utilize the equivalence of optimization (or validity) over a polytope and membership in the polytope. Typically optimization algorithms use the equivalence of optimization and separation to show upper bounds, e.g., that a linear program with an exponential number of constraints can be optimized. Our work is unique in that we use the less-utilized equivalence of validity and membership; and secondly, we use it to show hardness. In fact, since we are looking for hardness of approximation algorithms, our work is complicated further by the need to use *weak* versions of this equivalence.

Given a convex and compact set $K$ and a vector $c$, the *Strong Validity* problem, given a vector $c$, is to find the maximum value of $c^T x$ such that $x \in K$ (the $x$ which obtains this maximum is not required). In the *Strong Membership* problem, the goal is to determine if a given vector $y$ is in $K$ or not. The *Weak Validity* and *Weak Membership* problems are weaker versions of the Strong Validity and Strong Membership problems respectively, formally defined later. Then Theorem 4.4.4 in [16] says that for a convex and compact body $K$, there is an oracle polynomial time reduction from the Weak Membership problem for $K$ to the Weak Validity problem for $K$.

To formally state Theorem 4.4.4 from [16], which will form the basis of our reduction, we need the following notations and definitions.

We use $||.||$ for the Euclidean norm. Let $K \subseteq \mathbb{R}^{n'}$ be a convex and compact set. A ball of radius $\epsilon > 0$ around $K$ is defined as

$$S(K, \epsilon) := \{x \in \mathbb{R}^{n'} \,|\, ||x - y|| \leq \epsilon \text{ for some } y \text{ in } K\}\,.$$

Thus, for $x \in \mathbb{R}^{n'}$, $S(x, \epsilon)$ is the ball of radius $\epsilon$ around $x$. The interior $\epsilon$-ball of $K$ is defined as

$$S(K, -\epsilon) := \{x \in K | S(x, \epsilon) \subseteq K\}$$

Thus $S(K, -\epsilon)$ can be seen as points deep inside $K$.

▶ **Definition 20** ([16]). *Given a vector $c \in \mathbb{Q}^{n'}$, a rational number $\gamma$ and a rational number $\epsilon > 0$, the Weak Validity problem is to assert either (1) $c^T x \leq \gamma + \epsilon$ for all $x \in S(K, -\epsilon)$, or (2) $c^T x \geq \gamma - \epsilon$ for some $x \in S(K, \epsilon)$. Note that the vector $x$ satisfying the second inequality is not required.*

▶ **Definition 21** ([16]). *Given a vector $y \in \mathbb{R}^{n'}$ and $\delta > 0$, the Weak Membership problem is to assert either (1) $y \in S(K, \delta)$, or (2) $y \notin S(K, -\delta)$.*

Intuitively, in the Weak Membership problem, it is required to distinguish between the cases when the given point $y$ is far from the polyhedron $K$ (in which case, the algorithm should return $y \notin S(K, -\delta)$) and $y$ is deep inside $K$ (which case the algorithm should return $y \in S(K, \delta)$). If $y$ is near the boundary of $K$, then either output can be returned. Our reduction crucially uses the following result.

▶ **Theorem 22** (Theorem 4.4.4 of [16]). *Given a weak validity oracle for $K \subseteq \mathbb{R}^{n'}$ that runs in polynomial time and a positive $R$ such that $K \subseteq S(0, R)$, the Weak Membership problem for the polyhedron $K$ can be solved in polynomial time.*

For our problem $K$ is the polytope of linear program Norm-D.

$$K := \left\{ y \in \mathbb{R}^n : \sum_{i:S\cap T_i \neq \emptyset} y_i \leq 0 \quad \forall S \subseteq [m], \quad ||y||_\infty \leq 1 \right\}. \tag{8}$$

## Coverage WM $\leq_p$ Coverage WV $\leq_p$ Coverage Norm Extension

Coverage Weak Membership is the Weak Membership problem for polytope $K$ (8). Given a set $\mathcal{D} = \{T_1, \ldots, T_n\}$ (where $T_i \subseteq [m]$) with weights $\hat{y}_i$ ($\hat{y}_i \in \mathbb{R}$) associated with $T_i$ for all $i \in [n]$ and a $\delta > 0$, the goal in this problem is to assert either $(\hat{y}_1, \ldots, \hat{y}_n) \in S(K, \delta)$ or $(\hat{y}_1, \ldots, \hat{y}_n) \notin S(K, -\delta)$.

Note that Coverage Norm Extension is the Strong Validity problem for $K$ with $c_i = f_i$. We show the following lemma (Coverage WV $\leq_p$ Coverage Norm Extension).

▶ **Lemma 23.** *If there is an $\alpha = 2^{poly(n,m)} F^\delta$ efficient approximation algorithm (for any fixed $0 \leq \delta < 1$) for Coverage Norm Extension then there is an efficient algorithm for Weak Validity problem for $K$.*

Theorem 22 immediately gives Coverage WM $\leq_p$ Coverage WV.

## Span WM $\equiv$ Coverage WM

In fact, we show that Coverage Weak Membership is NP-hard even for the case when $|T_i| = 2$ for all $i \in [n]$.[3] The restriction $|T_i| = 2$ gives us a graphical representation of the membership problems. We first introduce some notations, which will be used in the remainder. Given a weighted graph $G = (V, E)$ and a set $S \subseteq V$, the span $E^+_G(S)$ and cut $\delta_G(S)$ of set $S$ are the set of edges with at least one endpoint and exactly one endpoint in $S$ respectively. We use $w(E^+_G(S)), w(\delta_G(S))$ and $w(E_G(S))$ for the sum of weight of edges with at least one endpoint, exactly one endpoint and both endpoints in $S$ respectively. If the set $S$ is a single vertex $v$ then we use $v$ instead of $\{v\}$. If the graph $G$ is understood from the context we drop the subscript $G$.

Given a set $\mathcal{D} = \{T_1, \ldots, T_n\}$ ($T_i \subseteq [m]$) with the property that $|T_i| = 2$ for all $i \in [n]$, we construct a weighted graph $G = (V, E)$ as follows: vertex set $V = [m]$ and $\{i, j\} \in E$ ($i, j \in [m]$) iff there exists a $T_k \in \mathcal{D}$ such that $T_k = \{i, j\}$. The weight $\hat{y}_k$ associated with $T_k = \{i, j\}$ is now associated to the edge $\{i, j\}$. Now the constraint $\sum_{i:T_i \cap S \neq \emptyset} y_i \leq 0$ (in the polyhedron $K$) translates to $\sum_{e \in E^+(S)} y_e \leq 0$ for all $S \subseteq V$. Thus Coverage-Weak-Membership for $|T_i| = 2$ case is equivalent to following problem, which we call *Span Weak Membership.*

Given a weighted graph $G = (V, E)$ with weights $\hat{y}_e$ on the edges and $\delta > 0$, assert either $\hat{y} = (\hat{y}_e)_{e \in E}$ is in $S(K_s, \delta)$ or $\hat{y}$ is not in $S(K_s, -\delta)$, where

$$K_s = \left\{ \sum_{e \in E^+(S)} y_e \leq 0 \quad \forall S \subseteq V, \, ||y||_\infty \leq 1 \right\}. \tag{9}$$

[3] There is a relatively easier proof for unrestricted $d$ by reduction from Set Cover, which we show in the full version.

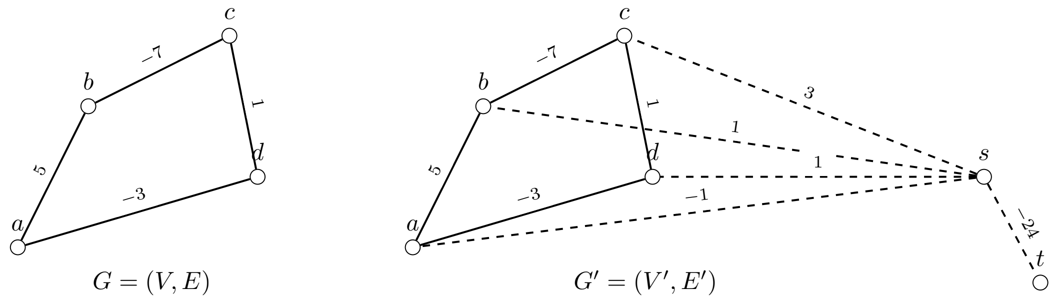


**Figure 1** Reduction from Cut Strong Membership to Span Strong Membership. The number shown on the edges in $E$ is the weight $y_e$, while on edges in $E'$ is the product of $L = 24$ and weight $y'_e$.

## Densest-Cut $\leq_p$ Cut WM $\leq_p$ Span WM

We now show that the Span Weak Membership is NP-Hard thereby showing Coverage Weak Membership is also NP-Hard for the restricted setting with $|T_i| = 2$ for all $i \in [n]$. We first define Cut Weak Membership.

Given a weighted graph $G = (V, E)$ with weights $\hat{y}_e$ on the edges and $\delta > 0$, the goal in Cut Weak Membership is to assert either $\hat{y} = (\hat{y}_e)_{e\in E}$ is in $S(K_c, \delta)$ or $\hat{y}$ is not in $S(K_c, -\delta)$ where

$$K_c = \left\{ \sum_{e\in\delta(S)} y_e \leq 0 \quad \forall S \in 2^V \setminus \emptyset, ||y||_\infty \leq 1 \right\}. \tag{10}$$

Note that in the Cut Weak membership problem, we have constraints $\sum_{e\in\delta(S)} y_e \leq 0$ instead of $\sum_{e\in E^+(S)} y_e \leq 0$ for all $S$.

▶ **Lemma 24.** *There is a reduction from Densest-Cut to Cut Weak Membership and from Cut Weak Membership to Span Weak Membership. Therefore, Coverage Weak Membership is NP-hard even when $d = 2$.*

We can now complete the proof of Theorem 5.

**Proof of Theorem 5.** Suppose there is an efficient $\alpha$-approximation algorithm for Coverage Norm Extension. Then by Lemma 23 there is an efficient algorithm for Weak Validity problem for polytope $K$ (8) and then by Theorem 22 we have an efficient algorithm for Coverage Weak Membership. But by Lemma 24, this is not possible unless $P = NP$. ◀

We here prove Lemma 25, which is a weaker statement than Lemma 24 to convey the main ideas. Recall that in Strong Membership problem, the goal is to decide if given vector $y$ is in polyhedron $K$. Following our nomenclature, we define the following Strong Membership problems.

An instance of Span Strong Membership and Cut Strong Membership is given by a weighted graph $G = (V, E)$ with weights $\hat{y}_e$ on the edges, and the goal is to decide if vector $y = (y_e)_{e\in E}$ is in $K_s$ and $K_c$ respectively, with $K_s$ and $K_c$ as defined in (9), (10).

▶ **Lemma 25.** *There is a reduction from Densest-Cut to Cut Strong Membership, and from Cut Strong Membership to Span Strong Membership.*

**Proof.** For the second reduction, the instance of Cut Strong Membership is weighted graph $G = (V, E)$ with weights $y_e$ on the edges. We assume $||y||_\infty \leq 1$ as otherwise clearly $y \notin K_c$.

Let $L = 2|E| + |V||E|$. We construct an instance of Span Strong Membership (see Figure 1), i.e., graph $G' = (V', E')$ and weights $y'_e$ as follows:

$$V' = V \cup \{s, t\}\,,\; E' = E \cup \{s, t\} \cup \{v, s\} \quad \forall v \in V\,,\; y'_e = \begin{cases} \frac{y_e}{L} & \text{if } e \in E(G) \\ -\frac{1}{2L} w(\delta_G(v)) & \text{if } e = \{v, s\}, v \neq t \\ -1 & \text{if } e = \{s, t\}. \end{cases}$$

Then $||y'||_\infty \leq 1$.

Assume $y \notin K_c$, i.e., there exists $S \subseteq V$ s.t. $w(\delta_G(S)) > 0$. We need to show there exists $S' \subseteq V'$ s.t. $\sum_{e \in E^+(S')} y'_e > 0$. For $S' = S$, $L \sum_{e \in E^+(S')} y'_e = w(E_G(S)) + w(\delta_G(S)) + \sum_{v \in S} -\frac{1}{2} \cdot w(\delta_G(v)) = w(E_G(S)) + w(\delta_G(S)) - \frac{1}{2} \cdot (2w(E_G(S)) + w(\delta_G(S))) = \frac{w(\delta_G(S))}{2} > 0$.

Now assume $y \in K_c$, i.e., $\forall S \subseteq V, w(\delta_G(S)) \leq 0$. We need to show $\forall S' \subseteq V'$, $\sum_{e \in E^+(S')} y'_e \leq 0$. Since $y'_{\{s,t\}} = -1$ (and $L$ is sufficiently large), we need to consider only those $S'$ which do not contain either $s$ or $t$. But we have shown that for such $S'$, $\sum_{e \in E^+(S')} y'_e = \frac{w(\delta_G(S'))}{2L} \leq 0$.

Now we finish the proof by giving a reduction from Densest-Cut to Cut Strong Membership. Given an undirected graph $G = (V, E)$ and rational $M$, we want to know if there exists $S \subset V$ s.t. $\frac{\delta_G(S)}{|S||V \setminus S|} > M$. Consider the complete graph $G' = (V, E')$ where the weight of an edge is $\frac{1-M}{L}$ if it existed in $E$, and is $-\frac{M}{L}$ otherwise (note that edges may now have positive, negative, or zero weight). Let $L' = 2\max\{M, |1 - M|\}$ be a sufficiently large quantity so that $||\hat{y}||_\infty < 1$ . It is easy to see that $L\, w(\delta_{G'}(S)) = |\delta_G(S)| - M|S||V \setminus S|$. Therefore, $\exists S \subset V$ s.t. $w(\delta_{G'}(S)) > 0 \Leftrightarrow \exists S \subset V$ s.t. $\frac{|\delta_G(S)|}{|S||V \setminus S|} > M$. ◀

## References


1. Ashwinkumar Badanidiyuru, Shahar Dobzinski, Hu Fu, Robert Kleinberg, Noam Nisan, and Tim Roughgarden. Sketching valuation functions. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 1025–1035. Society for Industrial and Applied Mathematics, 2012.
2. Maria-Florina Balcan and Nicholas J. A. Harvey. Learning submodular functions. In *Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC 2011, San Jose, CA, USA, 6-8 June 2011*, pages 793–802, 2011.
3. Eric Balkanski, Aviad Rubinstein, and Yaron Singer. The limitations of optimization from samples. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1016–1027. ACM, 2017.
4. Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997.
5. Umang Bhaskar and Gunjan Kumar. Partial Function Extension with Applications to Learning and Property Testing. *arXiv preprint*, 2018. `arXiv:1812.05821`.
6. Liad Blumrosen and Noam Nisan. Combinatorial auctions. *Algorithmic game theory*, 267:300, 2007.
7. Paul S. Bonsma, Hajo Broersma, Viresh Patel, and Artem V. Pyatkin. The Complexity Status of Problems Related to Sparsest Cuts. In *Combinatorial Algorithms - 21st International Workshop, IWOCA 2010, London, UK, July 26-28, 2010, Revised Selected Papers*, pages 125–135, 2010.
8. Christian Borgs, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. Maximizing social influence in nearly optimal time. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 946–957. SIAM, 2014.
9. Endre Boros, Toshihide Ibaraki, and Kazuhisa Makino. Error-Free and Best-Fit Extensions of Partially Defined Boolean Functions. *Inf. Comput.*, 140(2):254–283, 1998.

10 Deeparnab Chakrabarty and Zhiyi Huang. Recognizing Coverage Functions. *SIAM J. Discrete Math.*, 29(3):1585–1599, 2015.

11 Gerard Cornuejols, Marshall L Fisher, and George L Nemhauser. Exceptional paper—Location of bank accounts to optimize float: An analytic study of exact and approximate algorithms. *Management science*, 23(8):789–810, 1977.

12 F Dragomirescu and C Ivan. The smallest convex extensions of a convex function. *Optimization*, 24(3-4):193–206, 1992.

13 Vitaly Feldman and Pravesh Kothari. Learning coverage functions and private release of marginals. In *Conference on Learning Theory*, pages 679–702, 2014.

14 Rafael M. Frongillo and Ian A. Kash. General Truthfulness Characterizations via Convex Analysis. In *Web and Internet Economics - 10th International Conference, WINE 2014, Beijing, China, December 14-17, 2014. Proceedings*, pages 354–370, 2014.

15 Chris Godsil and Gordon F Royle. *Algebraic graph theory*, volume 207. Springer Science & Business Media, 2013.

16 Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012.

17 Subhash Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In *Proceedings 2001 IEEE International Conference on Cluster Computing*, pages 600–609. IEEE, 2001.

18 Andreas Krause, H Brendan McMahan, Carlos Guestrin, and Anupam Gupta. Robust submodular observation selection. *Journal of Machine Learning Research*, 9(Dec):2761–2801, 2008.

19 Benny Lehmann, Daniel Lehmann, and Noam Nisan. Combinatorial auctions with decreasing marginal utilities. *Games and Economic Behavior*, 55(2):270–296, 2006.

20 Hans JM Peters and Peter P Wakker. Convex functions on non-convex domains. *Economics letters*, 22(2-3):251–255, 1986.

21 Leonard Pitt and Leslie G. Valiant. Computational limitations on learning from examples. *J. ACM*, 35(4):965–984, 1988.

22 Lior Seeman and Yaron Singer. Adaptive seeding in social networks. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 459–468. IEEE, 2013.

23 C. Seshadhri and Jan Vondrák. Is Submodularity Testable? *Algorithmica*, 69(1):1–25, 2014.

24 Donald M Topkis. Minimizing a submodular function on a lattice. *Operations research*, 26(2):305–321, 1978.

25 Armin Uhlmann. Roofs and convexity. *Entropy*, 12(7):1799–1832, 2010.

26 Min Yan. Extension of convex function. *arXiv preprint*, 2012. To appear in the Journal of Convex Analysis. `arXiv:1207.0944`.


# A Appendix

## Proof of Proposition 7

Consider the polyhedron Extension-P. If the partial function is extendible, then Extension-P is nonempty. Since the variables are non-negative, the polyhedron must have a vertex [4], and in particular there is a vertex in which at most $n$ variables $w(S)$ are non-zero. This is because the dimension of the problem is $2^m$, hence at a vertex at least $2^m$ constraints must be tight. But then at least $2^m - n$ of constraints $w(S) \geq 0$ must be tight.

## Proof of Corollary 9

Consider the polytope $P = \{\sum_{I\in\mathcal{I}:v\in I} x_I = 1 \quad \forall v \in V(G), 0 \leq x_I \leq 1 \quad \forall I \in \mathcal{I}\}$. By the Theorem 8, there exists $x = \{x_I\}_{I\in\mathcal{I}}$ in $P$ such that $\chi^*(G) = \sum_{I\in\mathcal{I}} x_I$. Consider $y = \{y_I\}_{I\in\mathcal{I}}$ given by $y_{\{v\}} = 1$ for all $v \in V(G)$ and 0 otherwise. Therefore, $y \in P$ and $|V(G)| = \sum_{I\in\mathcal{I}} y_I$. Consider $z = \lambda x + (1-\lambda)y$ where $\lambda = \frac{|V(G)|-t}{|V(G)|-\chi^*(G)}$. Therefore, $z \in P$ and $\sum_{I\in\mathcal{I}} z_I = \lambda \sum_{I\in\mathcal{I}} x_I + (1-\lambda)\sum_{I\in\mathcal{I}} y_I = t$.

## Proof of Lemma 12

From Theorem 6, given a partial function $H$ and $\alpha \geq 1$, there exists a coverage function $f$ satisfying $f_i \leq f(T_i) \leq \alpha f_i$ for all $i \in [n]$ iff the following linear program is feasible, where the variables are the $W$-coefficients $w(S)$ for all $S \in 2^{[m]} \setminus \emptyset$ :

$$f_i \leq \sum_{S:S\cap T_i \neq \emptyset} w(S) \leq \alpha f_i \quad \forall i \in [n]$$

$$w(S) \geq 0 \quad \forall S \in 2^{[m]} \setminus \emptyset.$$

By Farkas' Lemma, it follows that the above linear program is feasible iff the following linear program is infeasible, with variables $y_i$ and $z_i$ for all $i \in [n]$:

$$\alpha \sum_{i=1}^{n} f_i y_i < \sum_{i=1}^{n} f_i z_i \tag{11}$$

$$\sum_{i:S\cap T_i \neq \emptyset} y_i \geq \sum_{i:S\cap T_i \neq \emptyset} z_i \quad \forall S \in 2^{[m]} \setminus \emptyset \tag{12}$$

$$y_i, z_i \geq 0.$$

Now we proceed towards proving the claim. Suppose $l_i$'s satisfy (4) and (3). Set $y_i$ and $z_i$ as follows: If $l_i \leq 0$ then let $y_i = -l_i$ and $z_i = 0$. Else if $l_i > 0$ then let $y_i = 0$ and $z_i = l_i$. It is easy to see that $y_i, z_i \geq 0$ and $l_i = z_i - y_i$ and hence (12) is satisfied by $y_i$'s and $z_i$'s. Further, $\alpha \sum_{i=1}^{n} f_i y_i = \alpha(\sum_{i:l_i \leq 0} f_i y_i + \sum_{i:l_i>0} f_i y_i) = -\alpha \sum_{i:l_i \leq 0} f_i l_i$ and similarly $\sum_{i=1}^{n} f_i z_i = \sum_{i:l_i>0} f_i l_i$. Thus (11) is also satisfied by $y_i$'s and $z_i$'s.

For the other direction observe that if the vector $y = (y_1, .., y_n), z = (z_1, ..., z_n) \geq 0$ satisfy (11) and (12) then wlog we can assume for any $i$, the minimum of $y_i$ and $z_i$ is 0 (otherwise we can decrease both $y_i$ and $z_i$ by the minimum of $y_i$ and $z_i$, and $\alpha \geq 1$ allows (11) to remain true). Note that $\sum_i f_i y_i = \sum_{i:y_i \leq z_i} f_i y_i + \sum_{i:y_i>z_i} f_i y_i = \sum_{i:y_i>z_i} f_i y_i$, since $\min\{y_i, z_i\} = 0$ by the previous observation. Now suppose $y, z \geq 0$ satisfy (11) and (12). We thus have $\alpha \sum_{i=1}^{n} f_i y_i < \sum_{i=1}^{n} f_i z_i \Leftrightarrow \alpha \sum_{y_i>z_i} f_i y_i < \sum_{z_i>y_i} f_i z_i$. Now let $l_i = z_i - y_i$. This makes both (4) and (3) true.

## Proof of Lemma 19

Suppose we are given a weighted bipartite graph $G = (A \cup [m], E)$ with weight $f_v$ on each $v \in A$. Recall that $\kappa$ is the minimum of $\frac{\sum_{w \in R} f_w}{f_v}$ over vertices $v \in A$ and $R \in \mathcal{F}_v$ where $\mathcal{F}_v = \{R \subseteq A \setminus \{v\} | N(R) \supseteq N(v)\}$ is the set of all $R \subseteq A \setminus \{v\}$ that covers all the neighbours of $v$.

We will use $f(R)$ $(R \subseteq A)$ to denote the summation $\sum_{v \in R} f_v$. If $d$ is a constant then for each $v \in A$, we can find minimum of $f(R)$ over all $R \subseteq \mathcal{F}_v$ in $O(n^d)$ time where $n = |A|$. Therefore, by taking the minimum of the above minimum value over all vertices $v \in A$, we get the value of $\kappa$. For general $d$, we use an approximation algorithm for Set-Cover to find, for each vertex $v \in A$, a set $R'_v \in \mathcal{F}_v$ such that $f(R'_v) \leq f(R_v) \log d$ where $R_v$ is the optimal set. It can be seen that $\kappa' = \min_{v \in A} \frac{f(R'_v)}{f_v}$ has the property $\kappa' \leq \kappa \log d$.

## Proof of Lemma 23

The instance of weak validity problem is given by a vector $c \in \mathbb{Q}^n$ and rational numbers $\gamma$ and $\epsilon > 0$. We show that there is a reduction from general Weak Validity to Weak Validity with instances satisfying $c_i \geq 0$ for all $i \in [n]$.

Let $N = \{i \in [n] | c_i \leq 0\}$. Consider a vector $c'$ such that $c'_i = 0$ for $i \in N$ and $c_i$ otherwise and $\gamma' = \gamma - \sum_{i \in N} |c_i|$. If $x$ is in $S(K, \epsilon)$ then clearly $\bar{x}$ defined as $\bar{x}_i = -1$ if $i \in N$ and $x_i$ otherwise, is also in $S(K, \epsilon)$. If for some $x$ in $S(K, \epsilon)$, we have $(c')^T x \geq \gamma' - \epsilon$ then for $\bar{x} \in S(K, \epsilon)$, we have $c^T \bar{x} = \sum_{i \in N} |c_i| + (c')^T x \geq \gamma - \epsilon$. Also if for all $x \in S(K, -\epsilon)$, we have $(c')^T x \leq \gamma' + \epsilon$ then $c^T x \leq \sum_{i \in N} |c_i| + (c')^T x \leq \gamma + \epsilon$. This shows the reduction and hence we assume $c_i \geq 0$ in the instance of Weak Validity problem.

Let $OPT$ and $OPT'$ be the optimal value of Norm-P for $(f_1, \ldots, f_n) = (c_1, \ldots, c_n)$ and $(f_1, \ldots, f_n) = (Lc_1, \ldots, Lc_n)$ respectively ($L$ will be chosen later). Obviously $OPT' = L \cdot OPT$. Let the approximation algorithm for Norm-P return $\beta$ for instance $(f_1, \ldots, f_n) = (Lc_1, \ldots, Lc_n)$. Let $C = \sum_i c_i$. Therefore, $OPT' \leq \beta \leq OPT' + 2^{poly(n,m)}(LC)^\delta = L \cdot OPT + 2^{poly(n,m)}(LC)^\delta$ and hence $\beta/L \leq OPT + \frac{2^{poly(n,m)}(C)^\delta}{L^{1-\delta}}$. We set $L := \left(\frac{2^{poly(n,m)}(C)^\delta}{2\epsilon}\right)^{1/1-\delta}$ so that $\frac{2^{poly(n,m)}(C)^\delta}{L^{1-\delta}} = 2\epsilon$. Note that the number of bits to specify $L$ is polynomial in $\langle c \rangle, \langle \epsilon \rangle, n, m$, where $\langle c \rangle$, $\langle \epsilon \rangle$ denote the number of bits required to represent these quantities. Thus, $OPT \leq \beta/L \leq OPT + 2\epsilon$. Now if $\gamma + \epsilon \leq \beta/L$ then for the optimal solution $x^* \in K$, $c^T x^* = OPT \geq \frac{\beta}{L} - 2\epsilon \geq \gamma - \epsilon$. If $\gamma + \epsilon \geq \beta/L$ then for all $x$ in $K$ (and hence $S(K, -\epsilon)$), we have $c^T x \leq OPT \leq \beta/L \leq \gamma + \epsilon$. Since at least one of these two conditions must hold, the conditions of weak validity problem can be correctly asserted.

## Proof of Lemma 24

In the proof, for any vector $y$, recall that we use $||y||_\infty$ for $\max_i |y_i|$ and $||\hat{y} - y||$ for the Euclidean distance between $\hat{y}$ and $y$. We will frequently use the fact that the distance of a point $x_0$ from the hyperplane $w^T x + b = 0$ is equal to $\frac{|w^T x_0 + b|}{||w||}$.

Recall the definitions of Span Weak Membership, Cut Weak Membership and Densest Cut:

1. Given a weighted graph $G = (V, E)$ with weights $\hat{y}_e$ on the edges and $\delta > 0$,
   a. The goal in Span Weak Membership is to assert either $\hat{y} = (\hat{y}_e)_{e \in E}$ is in $S(K_s, \delta)$ or $\hat{y}$ is not in $S(K_s, -\delta)$ where

   $$K_s = \left\{ \sum_{e \in E^+(S)} y_e \leq 0 \quad \forall S \in 2^V \setminus \emptyset, ||y||_\infty \leq 1 \right\},$$

   b. The goal in Cut Weak Membership is to assert either $\hat{y} = (\hat{y}_e)_{e \in E}$ is in $S(K_c, \delta)$ or $\hat{y}$ is not in $S(K_c, -\delta)$ where

   $$K_c = \left\{ \sum_{e \in \delta(S)} y_e \leq 0 \quad \forall S \in 2^V \setminus \emptyset, ||y||_\infty \leq 1 \right\}.$$

   Note that in the Cut Weak membership, we have constraints $\sum_{e \in \delta(S)} y_e \leq 0$ instead of $\sum_{e \in E^+(S)} y_e \leq 0$ for all $S$.
2. In the Densest-Cut problem, given a graph $G = (V, E)$ and a positive rational $M$, the goal is to decide if there exist a set $S \subset V$ s.t. $\frac{|\delta(S)|}{|S||V \setminus S|} \geq M$.

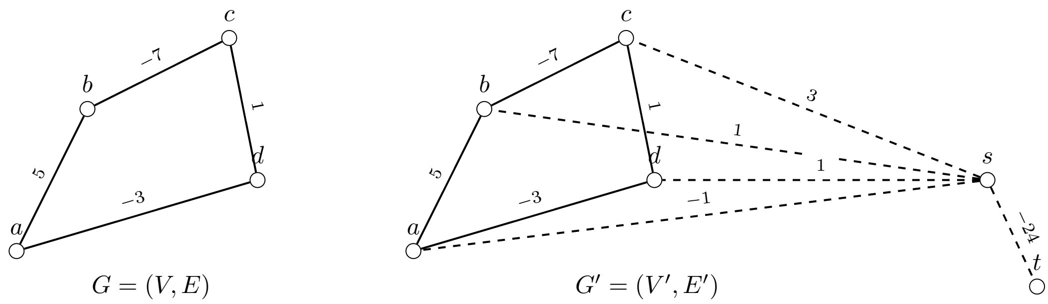


**Figure 2** Reduction from Cut-Weak-Membership to Span-Weak-Membership. The number shown on the edges in $E$ is the weight $y_e$, while on edges in $E'$ is product of $L = 48$ and weight $y'_e$.

The Densest-Cut is known to be NP-Hard [7]. Note that $\frac{|\delta(S)|}{|S||V \setminus S|}$ called the density of cut $(S, V \setminus S)$ can take values only from $\left\{\frac{r}{s(|V|-s)} | 1 \leq r \leq |E|, 1 \leq s \leq |V| - 1, r, s \in \mathbb{Z}_+\right\}$. Thus there are only polynomially many possible values of cut densities. We will use this fact in our proof.

▶ **Lemma 26.** *There is a reduction from Cut Weak Membership to Span Weak Membership.*

**Proof.** Our goal in Cut Weak Membership, given a graph a $G = (V, E)$ with weights $\hat{y}_e$ on edges and $\delta > 0$, is to assert either $\hat{y} = (\hat{y}_e)_{e \in E}$ is in $S(K_c, \delta)$ or $\hat{y}$ is not in $S(K_c, -\delta)$. If the point $\hat{y}$ violates the constraint $||y||_\infty \leq 1$ of $K_c$ then it can be asserted that $\hat{y}$ is not in $S(K_c, -\delta)$. So we assume $||\hat{y}||_\infty \leq 1$. Given this assumption, we have $w(\delta_G(v)) \leq |E|$.

We construct an instance of Span-Weak-Membership (see Figure 2), i.e., graph $G' = (V', E')$, $\hat{y}'_e$ and $\delta'$ as follows (the values of $B$ and $L$ will be set later):
$V' = V \cup \{s, t\}$
$E' = E \cup \{\{s, t\}\} \cup \{\{v, s\}\} \quad \forall v \in V', v \neq \{s, t\}$

$$\hat{y}'_e = \begin{cases} \frac{y_e}{L} & \text{if } e \in E \\ -\frac{\frac{1}{2}w(\delta_G(v))}{L} & \text{if } e = \{v, s\}, v \neq t \\ -\frac{B}{L} & \text{if } e = \{s, t\}. \end{cases}$$

The value of $B$ is set to $2|E| + |V||E|$ so that $\sum_{e \in E^+_{G'}(S)} \hat{y}'_e \leq \frac{-B+|E|+1/2|V||E|}{L} \leq 0$ for all $S$ containing either $s$ or $t$. Further, $L = 2B$ so that $||\hat{y}'||_\infty = 1/2$ where $\hat{y}' = (\hat{y}'_e)_{e \in E'}$. Finally we choose $\delta' = \frac{1}{2} \min\left\{\frac{\sqrt{|E|}\delta}{2L\sqrt{|E'|}}, \frac{|E|+1/2|V||E|}{\sqrt{|E'|}L}, \frac{1}{2}\right\}$.

▷ Claim 27. For all $S \subseteq V$, $w(E^+_{G'}(S)) = \frac{w(\delta_G(S))}{2L}$.

Proof. This is because

$$L\, w(E^+_{G'}(S)) = L \sum_{e \in E^+_{G'}(S)} w'_e = w(E_G(S)) + w(\delta_G(S)) + \sum_{v \in S} -\frac{1}{2} w(\delta_G(v)),$$

and since $w(\delta_G(v))$ counts edges in $E_G(S)$ twice and edges in $\delta_G(S)$ once,

$$L\, w(E^+_{G'}(S)) = w(E_G(S)) + w(\delta_G(S)) - \frac{1}{2} \cdot (2w(E_G(S)) + w(\delta_G(S))) = \frac{w(\delta_G(S))}{2}. \quad \triangleleft$$

Suppose the algorithm for Span Weak Membership asserts that the point $\hat{y}'$ is in $S(K_s, \delta')$. If $\hat{y}'$ satisfies all the constraints $\sum_{e \in E^+_{G'}(S)} y_e \leq 0$ for all $S \in 2^V \setminus \emptyset$ then the point $\hat{y}$ must satisfy all the constraints $\sum_{e \in \delta_G(S)} y_e \leq 0$ for all $S \in 2^V \setminus \emptyset$ (because by the Claim 27

$w(\delta_G(S)) = 2L \cdot w(E^+_{G'}(S)))$ and hence $\hat{y} \in K_c$ . Thus $\hat{y} \in S(K_c, \delta)$. Now suppose $\hat{y}'$ violates a constraint $\sum_{e \in E^+_{G'}(R)} y_e \le 0$ for some $R \in 2^V \setminus \emptyset$. Since $\hat{y}' \in S(K_s, \delta')$, it is at most $\delta'$ distance away from the hyperplanes corresponding to the violated constraints. Therefore, we have $w(E^+_{G'}(R)) = \sum_{e \in E^+_{G'}(R)} \hat{y}'_e \le \delta' \sqrt{|E'|}$. By Claim 27, $w(\delta_G(R)) \le 2L\delta' \sqrt{|E'|}$. Therefore, the point $\hat{y}$ is at most $\frac{2L\delta' \sqrt{|E'|}}{\sqrt{|E|}}$ distance from $K_c$. Since $\delta' < \frac{\sqrt{|E|}\delta}{2L\sqrt{|E'|}}$, so $\hat{y} \in S(K_c, \delta)$.

Suppose the algorithm for Span Weak Membership problem asserts that the point $\hat{y}'$ is not in $S(K_s, -\delta')$. If $\hat{y}'$ violates a constraint $\sum_{e \in E^+_{G'}(S)} y_e \le 0$ for some $S \in 2^V \setminus \emptyset$ then the point $\hat{y}$ also violates $\sum_{e \in \delta_G(S)} y_e \le 0$ for $S$ (by Claim 27). Hence, it can be asserted that $\hat{y}$ is not in $S(K_c, -\delta)$. So now assume that $\hat{y}'$ satisfies all the constraints $\sum_{e \in E^+_{G'}(S)} y_e \le 0$ for all $S \in 2^V \setminus \emptyset$. Also, as shown earlier, $\hat{y}'$ satisfies the other constraints of $K_s$. Since $\hat{y}'$ is in $K_s$ but not in $S(K_s, -\delta')$, some $y \in S(\hat{y}', \delta')$ must have distance $< \delta'$ from some hyperplane of $K_s$. The distance of $\hat{y}'$ from the hyperplane $\sum_{e \in E^+_{G'}(S)} y_e = 0$ for $S$ containing $s$ or $t$ is at least $\frac{|-B+|E|+1/2|V||E'||}{\sqrt{|E'|L}} = \frac{|E|+1/2|V||E'|}{\sqrt{|E'|L}} > \delta'$. Also for any $y \in S(\hat{y}', \delta')$, we have $||y||_\infty - ||\hat{y}'||_\infty \le ||y - \hat{y}||_\infty \le ||y - \hat{y}'||$. So $||y||_\infty \le \delta + 1/2 \le 1$ for all $y \in S(\hat{y}', \delta')$. Therefore, it must be the case that distance of $\hat{y}'$ from the hyperplane $\sum_{e \in E^+_{G'}(S)} y_e = 0$ for some $S \in 2^V \setminus \emptyset$ is $< \delta'$. By Claim 27, the distance of the point $\hat{y}$ from the hyperplane $\sum_{e \in \delta_G(S)} y_e = 0$ is at most $\frac{2L\sqrt{|E'|}\delta'}{\sqrt{|E|}} < \delta$. Hence, $\hat{y}$ is not in $S(K_c, -\delta)$. ◀

Now we finish the proof by giving reduction from Densest-Cut to Cut Weak Membership.

▶ **Lemma 28.** *There is a reduction from Densest-Cut to Cut Weak Membership.*

**Proof.** In the Densest Cut problem, a graph $G = (V, E)$ and a positive rational $M$ are given and the goal is to determine if there exists a set $S \subset V$ s.t. the density of the cut $(S, V \setminus S)$ is at least $M$, i.e., $\frac{|\delta_G(S)|}{|S||V \setminus S|} \ge M$. Let $M = \frac{p}{q}$ for positive integers $p, q$. We set $L$ to $2\max\{M, |1 - M|\}$ (so that later, $||\hat{y}||_\infty = 1/2$) and $t$ to $\frac{1}{qL}$ .

Given the graph $G = (V, E)$ and $M$, the instance of Cut Weak Membership is a complete graph $G' = (V, E')$ (so $|E'| = \frac{|V|(|V|-1)}{2}$), weight $\hat{y}_e$ on each edge $e \in E'$ such that $\hat{y}_e$ is $\frac{1-M}{L}$ if it existed in $E$ and $\frac{-M}{L}$ otherwise, and $\delta = \frac{1}{2}\min\{\frac{1}{2}, \frac{t}{\sqrt{|E'|}}\}$. Let $\hat{y} = (\hat{y}_e)_{e \in E'}$. This defines the polytope $K_c$ as in (10) for the instance of Cut Weak Membership.

It is easy to see that $w(\delta_{G'}(S)) = \frac{1}{L}(|\delta_G(S)| - M|S||V \setminus S|)$. Therefore, $\exists S \subset V$ s.t. $w(\delta_{G'}(S)) \ge 0 \Leftrightarrow \exists S \subset V$ s.t. $\frac{|\delta_G(S)|}{|S||V \setminus S|} \ge M$.

Since $M$ is equal to $\frac{p}{q}$ for some $p, q \in \mathbb{Z}_+$, therefore the weight of an edge is either $\frac{q-p}{qL}$ or $\frac{-p}{qL}$. So if a cut value $w(\delta_{G'}(S))$ is strictly positive for any $S$ then $w(\delta_{G'}(S))$ must be at least $\frac{1}{qL} = t$. Similarly, if $w(\delta_{G'}(S)) < 0$ then we have $w(\delta_{G'}(S)) \le -t$.

Now suppose an algorithm for Cut Weak Membership asserts $\hat{y}$ is in $S(K_c, \delta)$. Thus for all $S$, $w(\delta_{G'}(S)) = \sum_{e \in \delta_{G'}(S)} \hat{y}_e \le \sqrt{|E'|}\delta$. Since $\delta < \frac{t}{\sqrt{|E'|}}$, so it must be the case that for all $S$, $w(\delta_{G'}(S)) \le 0$. This implies that for all $S$, the cut density $\frac{|\delta_G(S)|}{|S||V \setminus S|} \le M$.

Suppose the algorithm for Cut Weak Membership asserts that $\hat{y}$ is not in $S(K_c, -\delta)$. If $\hat{y} \notin K_c$ (and since $||\hat{y}||_\infty \le 1$) then clearly there exists a set $S$ such that $w(\delta_{G'}(S)) > 0$. This implies there is a cut $(S, V \setminus S)$ with density $\frac{|\delta_G(S)|}{|S||V \setminus S|} > M$. Now assume $\hat{y}$ is in $K_c$. Now for any $y \in S(\hat{y}, \delta)$, we have $||y||_\infty - ||\hat{y}||_\infty \le ||y - \hat{y}||_\infty \le ||y - \hat{y}|| \le \delta$. So $||y||_\infty \le \delta + 1/2 < 1$. So there must exist a hyperplane $\sum_{e \in \delta_{G'}(S)} y_e = 0$ for some $S$ with at most $\delta$ distance from $\hat{y}$. Therefore, there exist a set $S$ with $0 \ge w(\delta_{G'}(S)) \ge -\sqrt{|E'|}\delta$. Since $\delta < \frac{t}{\sqrt{|E'|}}$, this means $w(\delta_{G'}(S)) = 0$ and hence density of cut $(S, V \setminus S)$ is $M$. Thus, if an algorithm for Cut Weak Membership asserts that $\hat{y}$ is not in $S(K_c, -\delta)$ then there exists a cut with density at least $M$.

Therefore, assuming an efficient algorithm for Cut Weak Membership, it can be determined if there exists a cut with density at least $M$ or all cuts have density at most $M$. However, the goal in Densest Cut is to determine if there is a cut with density $\geq M$ or all cuts have density strictly less than $M$. But since the density can take only polynomial number of values $\left\{ \frac{r}{s(|V|-s)} | 1 \leq r \leq |E|, 1 \leq s \leq |V| - 1, r, s \in \mathbb{Z}_+ \right\}$ (as noted before), by using at most two oracle calls to the Cut-Weak-Membership problem we can solve the original problem. ◀