Approximation Algorithms for the Incremental Knapsack Problem via Disjunctive Programming

Daniel Bienstock, Jay Sethuraman, Chun Ye

Department of Industrial Engineering and Operations Research
Columbia University

Abstract. In the *incremental knapsack problem* ($\mathcal{IK}$), we are given a knapsack whose capacity grows weakly as a function of time. There is a time horizon of T periods and the capacity of the knapsack is B_t in period t for $t = 1, \ldots, T$. We are also given a set S of N items to be placed in the knapsack. Item i has a value of v_i and a weight of w_i that is independent of the time period. At any time period t, the sum of the weights of the items in the knapsack cannot exceed the knapsack capacity B_t. Moreover, once an item is placed in the knapsack, it cannot be removed from the knapsack at a later time period. We seek to maximize the sum of (discounted) knapsack values over time subject to the capacity constraints. We first give a constant factor approximation algorithm for $\mathcal{IK}$, under mild restrictions on the growth rate of B_t (the constant factor depends on the growth rate). We then give a PTAS for $\mathcal{IIK}$, the special case of $\mathcal{IK}$ with no discounting, when $T = O(\sqrt{\log N})$.

1   Introduction

Traditional optimization problems often deal with a setting where the input parameters to the optimization problem are static. However, the static solution that we obtain may be inadequate for a system whose parameters—the inputs to our optimization problem—change over time. We consider one special case of this dynamic environment in which we have a maximization problem subject to certain capacity constraints. All of the inputs to the optimization problem are static except the capacities, which increase weakly over time. The goal is to find a sequence of compatible feasible solutions over time that maximizes a certain aggregate objective function. We will call such an optimization problem *an incremental optimization problem*. Unlike online and stochastic optimization problems, there is no uncertainty in the input parameters for the optimization.

In this paper we consider the *incremental knapsack problem*, which is a particular case of the incremental optimization problem. In the *discrete* incremental knapsack problem, we are given a knapsack whose capacity grows as a function of time. There is a time horizon of T periods and the capacity of the knapsack is B_t in period t for $t = 1, \ldots, T$. We are also given a set S of N items to be placed in the knapsack. Item i has a weight $w_i > 0$ that is independent of the time period, and a value at time t of the form $v_i \Delta_t$ where $v_i > 0$ and $\Delta_t > 0$ (this particular functional form will allow us to model discounting). At any time period t, the sum of the weights of the items in the knapsack cannot exceed the knapsack capacity B_t. Moreover, once an item is placed in the knapsack, it cannot be removed from the knapsack at a later time period. We are interested in maximizing the sum over the T time periods of the total value of the knapsack in each time period.

To put it formally, for $X \subseteq S$ define $V(X)$ to be $\sum_{i \in X} v_i$ and $W(X)$ to be $\sum_{i \in X} w_i$. Then we are interested in finding a feasible solution $F = \{S_1, S_2, \ldots, S_T\}$ and $S_1 \subseteq S_2 \ldots, S_T \subseteq S$, where S_t represents the subset of items in the knapsack in period t, that maximizes the quantity $\sum_{t=1}^{T} V(S_t) \Delta_t$ subject to the constraints $W(S_t) \leq B_t$ for $t = 1, \ldots, T$. The special case where $\Delta_t = 1$ for all t will be called *time-invariant*. For brevity, in what follows we will denote the incremental knapsack problem as $\mathcal{IK}$, and its time-invariant version as $\mathcal{IIK}$.

One can also consider a continuous version of the problem. Here we assume that we have a continuous parameter time parameter $s \in [0, S]$ for some $S > 0$. We are given a knapsack capacity function $B(s)$, weakly increasing with respect to s, and a set K of N items to be placed in the knapsack. Item i has a value of v_i and a weight of w_i, both time independent. At any time s, the sum of the weights of the items in the knapsack cannot exceed the knapsack capacity $B(s)$. Moreover, once an item is placed in the knapsack, it cannot be removed from the knapsack at a later time. We are interested in finding a feasible solution $F = \{K(s)\}_{s \in [0,S]}$ that maximizes the quantity $\int_{s=1}^{S} V(K(s)) ds$, where $V(K(s))$ is the total value of the items found in the knapsack at time s, under F. This problem can be approximated by partitioning $[0, S]$ into a finite set of intervals of length Δ_t, $t = 1, \ldots, T$. Under mild assumptions on the capacity function

$B(s)$, the approximation provided by this discretization can be made arbitrarily tight.

As the single period knapsack problem is already known to be NP-hard, we consider polynomial time approximation algorithms for $\mathcal{IK}$. For a maximization problem, a *k-approximation algorithm* (for some $k \leq 1$) is a polynomial time algorithm that guarantees, for all instances of the problem, a solution whose value is within k times the value of an optimal solution. Moreover, we say that the maximization problem has a (fully) polynomial time approximation scheme, or a PTAS (FPTAS respectively), if for every $0 \leq \epsilon < 1$, the algorithm guarantees, for all instances of the problem, a solution whose value is within $1 - \epsilon$ times the value of an optimal solution. Moreover, the algorithm runs in polynomial time in the size of the inputs (and ϵ) for every fixed ϵ.

1.1 Related Work

The special case of $\mathcal{IIK}$ where $v_i = w_i$ for all i has been examined in the literature. This problem is known as the *incremental subset sum problem*. Hartline [8] gave a 1/2-approximation algorithm for the incremental subset sum problem via dynamic programming. Sharp [9] gave a PTAS for the incremental subset sum problem that applies when T is a constant. This algorithm uses a variant of the dynamic programming algorithm for the standard (i.e., 1-period) knapsack problem, and it runs in time $O((\frac{VN}{\epsilon})^T)$, where $V = \max_i\{v_i\}$.

Further, it can be shown that the incremental subset sum is strongly NP-hard by a reduction from the 3-partition problem (proof provided in the Appendix).

Proposition 1. *The incremental subset sum problem is strongly NP hard.*

Consequently, the classic result of Garey and Johnson [6] rules out an FPTAS for the incremental subset sum problem (and hence for $\mathcal{IIK}$) both unless $P = NP$.

A well-studied problem related to $\mathcal{IIK}$ is the *generalized assignment problem* (GAP). In the generalized assignment problem, we are given a set of m knapsacks and N items, with knapsack j having a capacity b_j. Further, placing item i in knapsack j consumes w_{ij} units of capacity of knapsack j, and generates a value of v_{ij}. Notice that a variant of $\mathcal{IIK}$ where one is only allowed to pack, at each time t, an additional $B_{t+1} - B_t$ units, is a special case of the generalized assignment problem: Here, we would set $b_t = B_{t+1} - B_t$ and $w_{it} = w_i$ for all i and $v_{it} = (T - t + 1)v_i$ for all i and t. However, $\mathcal{IIK}$ is not a special case of GAP because in $\mathcal{IIK}$ we are allowed to pack more than $B_{t+1} - B_t$ units at time t, assuming the knapsack has spare capacity from earlier time periods. Approximation algorithms for the generalized assignment problem have been studied by [5],[10]. The best known constant factor algorithm is due to Fleischer et. al. [5] with an approximation ratio of $(1 - 1/e - \epsilon)$. They also showed that no approximation algorithm can beat the lower bound of $(1 - 1/e)$ unless $NP \subseteq DTIME(n^{O(\log \log(n))})$. Unfortunately, these results are not directly applicable to $\mathcal{IIK}$, because the knapsack capacities cannot be decomposed over time.

A special case of the generalized assignment problem where the items' weight and value are identical across knapsacks is known as the *multiple knapsack problem (MKP)*; for this problem, Chekuri and Khanna [4] developed a PTAS. Moreover, they also showed that two mild generalization of the MKP— $w_{ij} \in \{w_{i1}, w_{i2}\}$ and $v_{ij} = v_i$ or $v_{ij} \in \{v_{i1}, v_{i2}\}$ and $w_{ij} = w_i$— are APX hard, thus ruling out a PTAS for these generalizations, assuming $P \neq NP$. Again, neither the PTAS nor their hardness results are directly applicable to the $\mathcal{IIK}$.

1.2 Our Contributions

Our first result is a constant-factor approximation algorithm for $\mathcal{IK}$ under some mild assumptions on the growth rate of B_t. This algorithm rounds the solution to a polynomial-size linear programming relaxation to the problem, specifically, a disjunctive formulation (background and details, below). It is worth noting that, as is shown in Section 2, a standard formulation for the time-invariant incremental knapsack problem has an unbounded integrality gap—this is unlike the case for the standard knapsack problem. To the best of our knowledge, no constant factor approximation algorithm is known for $\mathcal{IK}$ before this work. The previous best algorithm is a general purpose approximation algorithm for incremental optimization problems due to Hartline and Sharp [7], which yield a $O(1/\log T)$ approximation ratio.

Our second result provides a PTAS for $\mathcal{IIK}$ and when $T = O(\sqrt{\log N})$. This approximation scheme involves a different disjunctive formulation that can be rounded to obtain the desired approximation. Specifically, we construct a disjunction over $O(N((1/\epsilon + T)^{O(\log(T/\epsilon)/\epsilon^2)}))$ LPs, each with NT variables and $O(NT)$ constraints. This improves on the result of Sharp [9]. Moreover, the analysis of the approach extends for certain Δ_t such as when $\Delta_t = e^{-rt}$ for some $r > 0$. This allows us to incorporate discounting. This PTAS also extends the earlier work of Bienstock and McClosky [2], [3] on the disjunctive approach for the single period knapsack problem.

Both of our algorithms rely on the classical approach of disjunctive programming [1]. Suppose we want to find an approximate solution to $\max\{w^T x : x \in P\}$ ($P \subseteq \mathbb{R}^n$, possibly non-convex), with approximation factor k. Moreover, the difficulty of the problem lies in that no good convex relaxation of P is known. In this case, we may still be able to leverage the idea of disjunctive programming to give us a good approximation guarantee. The idea is to find a set of polyhedra $Q^1, Q^2, \ldots Q^L$ in $\mathbb{R}^n$ such that $P \subseteq \cup_{i=1}^L Q^i$ and for each i we can compute, in polynomial time, $x^i \in P$ with $w^T x^i \geq k \max\{w^T x : x \in Q^i\}$. Taking $x^\star = \operatorname{argmax}_i\{w^T x^i\}$ yields a factor k approximate solution to the original optimization problem. As stated, this approach simply constitutes a case of enumeration (polynomially-bounded if L is polynomial). Further, $w^T x^\star \geq k \max\{w^T x : x \in \operatorname{conv}(\cup_i Q^i)\}$, and this last maximization problem can be formulated as a single linear program (polynomial-sized if L is), and, as will be the case below, we obtain an an approximation algorithm based on rounding.

The rest of the paper is organized as follows. In section 2, we show that the natural IP formulation of the time-invariant incremental knapsack problem

has an unbounded integrality gap. In section 3, we give the constant factor approximation algorithm for $\mathcal{IK}$. In section 4, we show a PTAS for $\mathcal{IIK}$. In section 5, we summarize our results and give suggestions for future directions.

2   LP Relaxation and Integrality Gap

To motivate the disjunctive approach, we will first show that the LP relaxation of a natural IP formulation for $\mathcal{IIK}$ has an unbounded integrality gap. This result implies that any constant factor approximation algorithm must do something more clever than simply solving the LP relaxation and rounding the fractional solutions to a feasible integral solution.

Let $x_{i,t} = 1$ if item i is placed in the knapsack at time t and 0 otherwise. We can formulate $\mathcal{IK}$ as the following binary integer program, whose feasible region will be denoted by P.

$$\text{IP} = \max \quad \sum_{t=1}^{T} \sum_{i=1}^{N} v_i \Delta_t x_{i,t} \tag{1}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} w_i x_{i,t} \leq B_t \quad \forall t$$

$$x_{i,t-1} \leq x_{i,t} \quad \forall i, \text{and } t = 2, 3, \ldots, T$$

$$x_{i,t-1} \in \{0, 1\} \quad \forall i, t.$$

Consider the case where $B_t = t$ for all t, $N = 1$ and $v_1 = w_1 = T$. Further assume $\Delta_t = 1$ for all t, i.e. we are consider the time-invariant case. Clearly (1) has value T in this instance, whereas the LP relaxation gives a value of $\frac{T(T+1)}{2}$, which implies that the integrality gap of $O(T)$ which is unbounded as $T \to \infty$.

A natural idea is to strengthen the LP relaxation by setting $x_{it} = 0$ if item i does not fit into the knapsack at time t. This strengthened LP relaxation, shown below, still has an unbounded integrality gap (as shown in the Appendix).

Theorem 2. *The following LP relaxation to $\mathcal{IIK}$ has an integrality gap that cannot be bounded by any constant.*

$$max \quad \sum_{t=1}^{T} \sum_{i=1}^{N} v_i x_{i,t}$$

$$s.t. \quad \sum_{i=1}^{N} w_i x_{i,t} \leq B_t \quad \forall t$$

$$x_{i,t-1} \leq x_{i,t} \quad \forall i, \text{ and } t = 2, 3, \ldots, T$$

$$x_{i,t} = 0 \quad \text{ for any } i, t \text{ such that } w_i > B_t$$

$$x_{i,t-1} \in [0, 1] \quad \forall i, t.$$

3 Constant Factor Algorithm

In this section we provide a constant-factor approximation algorithm for $\mathcal{IK}$ when the capacity function B_t is upper bounded by a polynomial function of t. To motivate our approach we introduce two definitions.

Definition 3. *Let $S > 0$ and $0 < \kappa \leq 1$. We say that a (S, κ)-split takes place at time t_κ if the following conditions happen:*

(i) $\sum_{t=1}^{t_\kappa - 1} \Delta_t \leq S \sum_{t=t_\kappa}^{T} \Delta_t$.
(ii) $B_T - B_{t_\kappa} \leq \kappa B_T$.

Remark 1. Consider the time-invariant case, i.e. $\Delta_t = 1$ for all t. When an (S, κ)-split takes place at t_κ, we have $t_\kappa < ST$, but the knapsack capacity at time t_κ is already (at least) a fraction $1 - \kappa$ of the final knapsack capacity, B_T.

Definition 4. *Let $0 < \kappa \leq 1$. Define $S(\kappa)$ to be the smallest value $S \geq 0$, such that there is a period t_κ where an (S, κ) split takes place.*

Remark 2. Note that $S(\kappa)$ decreases with κ. Also, suppose $\Delta_t = 1$ for all t and $B_t = \Theta(t^p)$ for some $p > 0$. Then we have that $S(\kappa) \approx \frac{(1-\kappa)^{1/p}}{1-(1-\kappa)^{1/p}}$. Even though this quantity converges to $+\infty$ as $p \to +\infty$, for *fixed* p it is bounded. This means that (by Definition 3) during the last $\frac{T}{S(\kappa)}$ time periods the knapsack only gains a fraction $\leq \kappa$ of the final capacity B_T.

Given $\kappa > 0$, there is an algorithm with approximation factor $\min\{\Omega(\frac{1-\kappa}{\kappa}), \frac{1}{S(\kappa)}\}$. Thus, roughly speaking, the quality of the approximation improves if both κ and $S(\kappa)$ are "small". Or, to put it differently, if the capacity function B_t is such that $S(\kappa)$ remains very large for $\kappa \approx 1$, then the quality of the approximation bound will suffer. As remark 2 shows, when B_t is polynomially bounded as a function of t, $S(\kappa)$ is bounded above, and so the ratio $\frac{1-\kappa}{\kappa S(\kappa)}$ remains bounded away from 0, i.e. we indeed obtain a constant-factor approximation algorithm. The class of cases where we do not attain a constant-factor approximation algorithms are those where $\lim_{\kappa \to 1} S(\kappa) = +\infty$. In such cases, the capacity function is attaining essentially all of its growth in an arbitrarily small final set of time periods.

In this paper, for brevity, we focus on the case $\kappa = 1/2$, and prove the following.

Theorem 5. *There is a polynomial-time algorithm for $\mathcal{IK}$ with approximation ratio $\min\left\{\frac{1}{9}, \frac{1}{6\max\{1, S(1/2)\}}\right\}$.*

In the Appendix we outline how to extend our approach to general $0 < \kappa < 1$.

Simplification. For simplicity, we assume, by perturbing the v_i if necessary that $v_i/w_i \neq v_j/w_j$ for all $i \neq j$. The perturbation changes the value of $\mathcal{IK}$ by an arbitrarily small amount.

3.1 The approximation algorithm

Our algorithm is based on running (up to) two polynomially bounded procedures given below, for each time period $1 \leq \bar{t} \leq T$: a "replicated knapsack" procedure, and an LP-rounding algorithm. Each run will produce a feasible (integral) solution to $\mathcal{IK}$; the algorithm will select a solution that attains the highest objective value. We will use the same notation as above, i.e. the variable $x_{i,t}$ is used to indicate whether item i is in the knapsack at time t.

To introduce the first procedure we need a definition.

Definition 6. *Let* $1 \leq \bar{t} \leq T$. *A 0-1 vector* $\bar{x}$ *is a* **replicated-knapsack solution** *at* $\bar{t}$ *for* $\mathcal{IK}$ *if the following conditions hold:*

$$\bar{x}_{i,t} = 0 \text{ for all } i \text{ and } t < \bar{t}, \quad \text{and} \quad \bar{x}_{i,t} = \bar{x}_{i,\bar{t}} \text{ for all } i \text{ and } t \geq \bar{t}.$$

Recall that the capacity function B_t is monotonely nondecreasing. Thus the replicated-knapsack solution at $\bar{t}$ is feasible for $\mathcal{IK}$ iff $\sum_i w_i x_{i,\bar{t}} \leq B_{\bar{t}}$, i.e. the solution replicates a feasible solution to the (single-period) knapsack with weights w_i and capacity $B_{\bar{t}}$. Further, the objective value of $\bar{x}$ equals $\left(\sum_{t=\bar{t}}^T \Delta_t\right) \sum_{i=1}^N v_i \bar{x}_{i,\bar{t}}$. This quantity can be approximated (arbitrarily closely) in polynomial time using several well-known methods.

In order to describe the second procedure, consider formulation (2)-(8), for which the time period $1 < \bar{t} < T$, the item index h, and a second time period $\check{t} \leq \bar{t}$ are inputs. Here, h is the most valuable item in the knapsack at time $\bar{t}$ and h was first placed in the knapsack at time $\check{t}$.

$$\boldsymbol{D(\bar{t}, \check{t}, h)}: \qquad \max \sum_{t=1}^T \sum_{i=1}^N v_i \Delta_t x_{i,t} \tag{2}$$

$$\text{s.t.} \quad \sum_{i=1}^N w_i x_{i,t} \leq B_t \quad \forall t \tag{3}$$

$$x_{i,t-1} \leq x_{i,t} \quad \forall i, \text{ and } t = 2, 3, \ldots, T \tag{4}$$

$$x_{i,t} \in \{0,1\} \quad \forall i, t \tag{5}$$

$$x_{h,\check{t}} = 1, \quad x_{h,\check{t}-1} = 0 \text{ (if } \check{t} > 0) \quad \text{and}$$
$$x_{i,\bar{t}} = 0, \quad \forall i \text{ with } v_i > v_h \tag{6}$$

$$\frac{1}{3} \sum_{i,t} v_i \Delta_t x_{i,t} \leq \sum_{t=\bar{t}}^T \sum_{i=1}^N v_i \Delta_t x_{i,t} \tag{7}$$

$$\frac{1}{3} \sum_{i,t} v_i \Delta_t x_{i,t} \leq \sum_{t=1}^{\bar{t}-1} \sum_{i=1}^N v_i \Delta_t x_{i,t}. \tag{8}$$

Other than constraints (6)-(8) this is the same formulation as in Section 2. Constraints (7) and (8) indicate that, in the time interval $[\bar{t}, T]$ (resp. $[1, \bar{t}-1]$) at least one-third of the total value is accrued. Using (4) and $\check{t} \leq \bar{t}$, constraint

(6) implies that $x_{h,\bar{t}} = 1$. Thus, (6) states that item h is placed in the knapsack at time $\breve{t}$, and that at time $\bar{t}$, item h is a highest-value item in the knapsack. We will now prove a number of results on this formulation.

Definition 7. *Let $D^R(\bar{t}, \breve{t}, h)$ denote the continuous relaxation of $D(\bar{t}, \breve{t}, h)$, i.e. that obtained by replacing (5) with $0 \leq x_{i,t} \leq 1$ for all i, t. Further, let $V^* = V^*(\bar{t}, \breve{t}, h)$ denote the value of $D^R(\bar{t}, \breve{t}, h)$.*

Lemma 8. *Suppose $\tilde{x} = \tilde{x}(\bar{t}, \breve{t}, h) \in \mathbb{R}^{N \times T}$ be an optimal solution to the linear program where:*

(i) The objective function is (2).
(ii) The constraints are (3)-(6), together with

$$\frac{1}{3}V^* \leq \sum_{t=\bar{t}}^{T} \sum_{i=1}^{N} v_i \Delta_t \tilde{x}_{i,t} \quad and \quad \frac{1}{3}V^* \leq \sum_{t=1}^{\bar{t}-1} \sum_{i=1}^{N} v_i \Delta_t \tilde{x}_{i,t}. \qquad (9)$$

i.e. we replace constraints (7) and (8) of $D^R(\bar{t}, \breve{t}, h)$ with (9). Then for any time period t there is at most one item i with $0 < \tilde{x}_{i,t} < 1$.

Remark. It can be shown that $\tilde{x}$ can be obtained by solving a single, polynomial-size linear program, rather than the two-LP procedure implied by Lemma 8.

Lemma 9. *Let $\tilde{x} = \tilde{x}(\bar{t}, \breve{t}, h)$ be as in Lemma 8 and let $\breve{x}$ be obtained by rounding down $\tilde{x}$. Then $\breve{x}$ is feasible for $D(\bar{t}, \breve{t}, h)$, and*

$$\sum_{i,t} v_i \Delta t \, \breve{x}_{i,t} \geq \frac{1}{6} \min\left\{ 1, \frac{\sum_{t=\bar{t}}^{T} \Delta_t}{\sum_{t=1}^{\bar{t}-1} \Delta_t} \right\} \sum_{i,t} v_i \Delta t \, \tilde{x}_{i,t}. \qquad (10)$$

Lemma 9, together with the replicated-knapsack construction given above, constitute the two cases that our algorithm will enumerate for each t. We will next show that the best solution of the replicated-knapsacks and the roundown constructions attain a factor $\frac{1}{6S}$ where $S = \max\{1, S(1/2)\}$ (recall Definition 3).

3.2 Existence of approximation

In the remainder of this section, we assume that x^Z is an **optimal** solution to a given instance of $\mathcal{IK}$. We V^Z denote the value of the instance, i.e. $V^Z \doteq \sum_i v_i \Delta_t x_{it}^Z$.

Lemma 10. *Either there is a period $1 < t^{[3]} < T$ such that*

(a) $\frac{1}{3}V^Z \leq \sum_{t=t^{[3]}}^{T} \sum_i v_i \Delta_t x_{it}^Z$
(b) $\frac{1}{3}V^Z \leq \sum_{t=1}^{t^{[3]}-1} \sum_i v_i \Delta_t x_{it}^Z$

or there is a replicated-knapsack solution with value at least $\frac{V^Z}{3}$.

Using Lemma 10 we can assume that none of the T replicated-knapsack solutions (which we can approximate, in polynomial time, to any desirable constant factor) is within a factor of $1/3$ of V^Z, and thus, that a time period $1 < t^{[3]} < T$ satisfying (a) and (b) does exist. Using this fact, we will next consider the formulations $D(\bar{t}, \check{t}, h)$ discussed in the previous section, and prove that the bound obtained in Lemma 9 will yield a large enough constant factor for at least one choice of $\bar{t}$ and h.

Recall Definition 3. Let $S = S(1/2)$ and $t_{1/2}$ be a time period so that

$$\sum_{t=1}^{t_{1/2}-1} \Delta_t \leq S \sum_{t=t_{1/2}}^{T} \Delta_t \quad \text{and} \quad B_T - B_{t_{1/2}} \leq \frac{B_T}{2}. \tag{11}$$

Lemma 11. *Suppose that $t^{[3]} \leq t_{1/2}$. Let h^* be the most valuable item in the knapsack, under solution x^Z, at time $t^{[3]}$, and let $\check{t} \leq t^{[3]}$ be the time it was placed in the knapsack. Then rounding down the solution to $D^R(t^{[3]}, \check{t}, h^*)$ yields a feasible solution to problem $\mathcal{IK}$, of value at least*

$$\frac{1}{6 \max\{1, S\}} \sum_{i,t} v_i \Delta t \, x_{i,t}^Z.$$

Proof. By Lemma 9 applied to formulation $D(t^{[3]}, \check{t}, h^*)$, we will obtain a feasible solution to $\mathcal{IK}$ with value at least

$$\frac{1}{6} \min \left\{ 1, \frac{\sum_{t=t^{[3]}}^{T} \Delta_t}{\sum_{t=1}^{t^{[3]}-1} \Delta_t} \right\} \sum_{i,t} v_i \Delta t \, x_{i,t}^Z. \tag{12}$$

But we are assuming that $t^{[3]} \leq t_{1/2}$. This implies the desired bound. $\square$

We can now assume that $t_{1/2} < t^{[3]}$. We can show that in this case a replicated-knapsack solution has value at least $V^*/9$.

Lemma 12. *Suppose $t_{1/2} < t^{[3]}$. Then a replicated-knapsack solution has value at least $V^*/9$.*

4 A PTAS for $\mathcal{IIK}$ when $T = O(\sqrt{\log N})$

Now we are ready to present the PTAS for the time-invariant incremental knapsack problem when $T = O(\sqrt{\log N})$. This algorithm is easily extended to the case of $\mathcal{IK}$ with fixed T and monotonically nonincreasing Δ_t quantities. Consider an instance of $\mathcal{IIK}$ and let $\epsilon \in (0, 1)$. Without loss of generality, we can assume that the v_i's are integral. Fixing an optimal solution OPT, and let h be a the maximum valued item that is ever placed in the knapsack by OPT. Then it suffices to optimize over the set of items $S^h = \{i \in S | v_i \leq v_h\}$. We partition S^h into $K + 1$ subsets $X = \{S^{1,h}, S^{2,h}, \ldots, S^{K,h}, T^h\}$, where

$$S^{k,h} = \{j \in S, j \neq h : (1-\epsilon)^{k-1} v_h \geq v_j > (1-\epsilon)^k v_h\} \quad \text{for } k = 1, \ldots, K,$$

and

$$T^h = \{j \in S : (1 - \epsilon)^K v_h \geq v_j\}.$$

In order to attain the approximation ratio, we will choose K large enough so that $(1 - \epsilon)^K < \epsilon/T$ or equivalently, $K > \frac{\log(T/\epsilon)}{\epsilon}$.

Consider a modified instance of the problem where items have identical weights as the original instance and item i has a modified value of $v_i' = (1 - \epsilon)^{k-1}V$ if $i \in S^k$ and $v_i' = v_i$ otherwise. Let OPT_m denote an optimal solution to the modified instance of the problem. Let $V(SOL)$ and $V_m(SOL)$ be the objective value with respect to a solution SOL of the original instance and the modified instance respectively. As we did not change the item weights, OPT_m is a feasible solution to the original instance. Moreover,

$$V(OPT_m) \geq (1 - \epsilon)V_m(OPT_m) \geq (1 - \epsilon)V_m(OPT) \geq (1 - \epsilon)V(OPT),$$

where the first inequality follows from the fact that $v_i \geq (1-\epsilon)v_i'$ for every item i, the second inequality follows from the fact that OPT_m is an optimal solution to the modified instance, and the third inequality follows from the fact that $v_i \leq v_i'$ for every item i.

Now, since all items within each $S^{k,h}$ have equal value in the modified instance, it is clear that conditioning on the number of items chosen by OPT within each $S^{k,h}$, OPT would tend to choose the items within the same value class in the order of non-decreasing weight (breaking ties arbitrarily). Thus, it suffices to enumerate feasible solutions that can be described by a collection of K T-vectors $\{\sigma^1, \ldots, \sigma^K\}$, where $\sigma_t^k \in \{0, 1, \ldots, |S^{k,h}|\}$ denotes the number of items chosen from $S^{k,h}$ in time period t, in order to find an optimal solution. Nonetheless, the number of potential solutions that we have to enumerate would be exponential in N if we attempt to enumerate all possible configurations of $\{\sigma^1, \ldots, \sigma^K\}$. Consequently, we will only explicitly enumerate σ_t^k taking values from $\{0, 1, \ldots, \min(\lceil 1/\epsilon \rceil, |S^{k,h}|)\}$. For σ_t^k taking values larger than $J = \lceil 1/\epsilon \rceil$, we will instead let the feasible region of an LP capture these feasible points and try to let the LP choose an optimal solution for us and subsequently round this optimal solution. Lastly, since we don't know the most valuable item h taken by OPT in the original instance of the problem, we will have to guess such an item by enumeration.

Our disjunctive procedure is as follows. First, we guess the most valuable item $h \in S$ packed by an optimal solution. Subsequently, we only consider choosing items from S^h and round the values of the items in S^h to obtain the modified instance of the problem. We will then focus on solving the modified instance of the problem. Let k_i, $i = 1, 2, \ldots, |S^{k,h}|$, be the i-th lightest weight item in $S^{k,h}$ (break ties arbitrarily). Let $x_{k_i,t}$ be the variable indicating whether item k_i is placed in the knapsack in time period t. Let $\sigma = \{\sigma^1, \ldots, \sigma^K\} \in \{0, \ldots, J\}^{TK}$ and define the following polyhedron:

$$Q^{\sigma,h} = \{x \in [0,1]^{T|S^h|} : x_{h,T} = 1$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{|S^{k,h}|},t} = 0 \quad \forall (k,t) \text{ s.t. } \sigma_t^k = 0$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{\sigma_t^k},t} = 1, \ x_{k_{\sigma_t^k+1},t} = \ldots = x_{k_{|S^{k,h}|},t} = 0$$

$$\forall (k,t) \text{ s.t. } 1 \le \sigma_t^k < J \text{ and } \sigma_t^k < |S^{k,h}|$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{\sigma_t^k},t} = 1, \ \sum_{i=1}^{|S^{k,h}|} x_{k_i,t} \ge \sigma_t^k$$

$$\forall (k,t) \text{ s.t. } \sigma_t^k = J \text{ and } \sigma_t^k < |S^{k,h}|$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{|S^{k,h}|},t} = 1 \quad \forall (k,t) \text{ s.t. } \sigma_t^k \ge |S^{k,h}|$$

$$\sum_{k=1}^{K} \sum_{i=1}^{|S^{k,h}|} w_{k_i} x_{k_i,t} + \sum_{i \in T^h} w_i x_{i,t} \le B_t \quad \forall t$$

$$x_{k_i,t-1} \le x_{k_i,t} \quad \forall (k,i), \text{and } t = 2,3,\ldots,T$$

$$x_{i,t-1} \le x_{i,t} \quad \forall i \in T^h, \text{and } t = 2,3,\ldots,T\}.$$

Lemma 13. *For each fixed h, there are $O((1/\epsilon+T)^{O(\log(T/\epsilon)/\epsilon^2)})$ polyhedra $Q^{\sigma,h}$ in our disjunctive procedure.*

Since we have to enumerate our guess for the most valuable item h, we get the following corollary.

Corollary 14. *There are a total of $O(N(1/\epsilon+T)^{O(\log(T/\epsilon)/\epsilon^2)})$ LPs in our disjunctive procedure.*

Notice that when $T = O(\sqrt{\log T})$, then the number of LPs is polynomial in N for a fixed ϵ.

The following is our main result.

Theorem 15. *For every non-empty polyhedron $Q^{\sigma,h}$, there exists a polynomially computable point $x^{\sigma,h}$ feasible for $\mathcal{IIK}$, such that*

$$\sum_{t=1}^{T} \sum_{i \in S^h} v_i' x_{i,t}^{\sigma,h} \ge (1-\epsilon) \max\{\sum_{t=1}^{T} \sum_{i \in S^h} v_i' x_{i,t} : \ x \in Q^{\sigma,h}\}.$$

Theorem 16. *Let $x^\star = \arg\max_{Q^{\sigma,h} \ne \emptyset} \sum_{t=1}^{T} \sum_{i \in S^h} v_i' x_{i,t}^{\sigma,h}$, where $x^{\sigma,h}$ is defined in the previous theorem, then we have that*

$$\sum_{t=1}^{T} \sum_{i \in S^h} v_i' x_{i,t}^\star \ge (1-\epsilon) V_m(OPT_m) \ge (1-\epsilon)^2 V(OPT).$$

Proof. This is a direct consequence of the fact that $OPT_m \in \cup_{Q^{\sigma,h} \neq \emptyset} Q^{\sigma,h}$ as $\cup_{Q^{\sigma,h} \neq \emptyset} Q^{\sigma,h}$ covers P, and that

$$\sum_{t=1}^{T} \sum_{i \in S^h} v_i' x_{i,t}^{\star} \geq (1 - \epsilon) \max \{ \sum_{t=1}^{T} \sum_{i \in S^h} v_i' x_{i,t} : \ x \in \cup_{Q^{\sigma,h} \neq \emptyset} Q^{\sigma,h} \}.$$

5 Conclusion

In this work, we give a constant factor approximation algorithm for $\mathcal{IK}$ when the capacity function B_t is upper bounded by a polynomial function of t. We also give a PTAS for $\mathcal{IIK}$ when the time horizon $T = O(\sqrt{\log N})$, where N is the number of items. Our results generalize and improve on some of the earlier results of Hartline and Sharp for this problem. Our work leaves to the following open questions. First, is there a polynomial time algorithm for $\mathcal{IK}$ with a constant factor approximation ratio that makes no assumption on the growth rate of B_t? Second, is there a PTAS for $\mathcal{IIK}$ for an arbitrary time horizon T? It is also interesting to consider an incremental version of other combinatorial optimization problems.

References

1. Egon Balas. Disjunctive programs: Cutting planes from logical conditions, in o.l. mangasarian et al., eds.,. *Nonlinear Programming.*, 2:279–312, 1974.
2. Daniel Bienstock. Approximate formulations for 0-1 knapsack sets. *Oper. Res. Lett.*, 36(3):317–320, 2008.
3. Daniel Bienstock and Benjamin McClosky. Tightening simple mixed-integer sets with guaranteed bounds. *Math. Program.*, 133(1-2):337–363, 2012.
4. Chandra Chekuri and Sanjeev Khanna. A polynomial time approximation scheme for the multiple knapsack problem. *SIAM J. Comput.*, 35(3):713–728, 2005.
5. Lisa Fleischer, Michel X. Goemans, Vahab S. Mirrokni, and Maxim Sviridenko. Tight approximation algorithms for maximum general assignment problems. In *SODA*, pages 611–620. ACM Press, 2006.
6. Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness.* W. H. Freeman & Co., New York, NY, USA, 1990.
7. Jeff Hartline and Alexa Sharp. An incremental model for combinatorial maximization problems. In Carme Àlvarez and Maria J. Serna, editors, *WEA*, volume 4007 of *Lecture Notes in Computer Science*, pages 36–48. Springer, 2006.
8. Jeffrey Hartline. *Incremental Optimization.* PhD thesis, Cornell University, 2008.
9. Alexa Sharp. *Incremental Algorithms: Solving Problems in a Changing World.* PhD thesis, Cornell University, 2007.
10. David B. Shmoys and Éva Tardos. An approximation algorithm for the generalized assignment problem. *Math. Program.*, 62:461–474, 1993.

6   Appendix

Proof of Proposition 1

Proof. In the 3 partition problem, we're given a set S of $3m$ integers $a_1, \ldots a_{3m}$, and we want to decide whether S can be partitioned into m triples that all have the same sum B, where $B = (1/m) \sum_{i=1}^{3m} a_i$. It has been shown that 3-partition is strongly NP-hard even if all the integers are between $B/4$ and $B/2$. We will reduce any instance of the 3-partition problem to a corresponding instance of the incremental subset problem. Given a set S of $3m$ integers $a_1, \ldots, a_{3m}$. Let a_i be the weight/value of item i. And we have a knapsack whose capacity is B_t in period t for $t = 1, \ldots, m$. Notice that the partition can be done if and only if in every period the knapsack reaches its capacity, i.e. the incremental amount that we pack is B. Moreover, it can be shown inductively that since the value of the items are strictly in between $B/4$ and $B/2$, we would pack exactly 3 additional items in every period.

Proof of Theorem 2

Proof. Fix a $k \geq 2$ and let T be a power of k. Consider a set of N items, where $v_i = w_i = k^i$ for $i = 1, \ldots, \log_k(T) = N$. The knapsack capacities follow the following pattern:

$$B_t = k^i \text{ if } T(1 - \frac{1}{k^{i-1}}) + 1 \leq t \leq T(1 - \frac{1}{k^i}) \text{ for } i = 1, \ldots, \log_k(T) \text{ and } B_T = B_{T-1}.$$

Since the LP can fractional pack the items, the knapsack attains its capacity in every time period. Moreover, since all items have weight equaling value, we have that the optimal value of the LP solution is (by evaluating the sum of the knapsack values over all time periods):

$$T + T \sum_{i=1}^{\log_k(T)} k^i T((1 - 1/k^i) - (1 - 1/k^{i-1}) = T(k-1)\log_k(T) + T = O(Tk\log_k(T)).$$

Let $t_i = T(1 - 1/k^{i-1}) + 1$ denote the first time when the knapsack capacity increases to k^i. Notice that any integer optimal solution would only pack additional items in time those periods, which means that we just need to figure out what to pack at those time periods in order to find an integer optimal solution. The only items that fit in the knapsack at time t_i are items 0 through i. If we decide to pack item i in period t_i, then the total revenue we get for packing i over times $t_i \leq t \leq t_{i+1} - 1$ is $T(1/k^i - 1/k^{i+1})k^{i+1} = T(k-1)$. Since we cannot pack any item before time t_i, the total revenue we get up to time $t_{i+1} - 1$ if we pack item i in time t_i would be $T(k - 1)$. For every $i > 1$, this is clearly suboptimal since we would get more revenue up to time $t_{i+1} - 1$ had we just packed item 1 in period 1 (since $kT(1 - 1/k^{i+1}) > kT(1 - 1/k) = T(k - 1)$ for $i > 1$). Hence, no integer optimal solution would pack item i at time t_i for every

$i > 1$.

If we do not pack item i at time t_i, then we can pack the first $i - 1$ items at time t_i for every $i > 1$ since $k \geq 2$. Hence, this is an optimal packing for time t_i for every $i > 1$ (and it is optimal to pack item 1 starting from period 1). Moreover, this solution respects the precedence constraints, which means that we have found an integer optimal solution. We evaluate the integer optimal solution by looking at how long each item has been placed in the knapsack:

$$kT + \sum_{i=2}^{N-1} k^i(T - t_{i+1} + 1) = kT + \sum_{i=2}^{N-1} k^i(T - T(1 - 1/k^i)) \simeq kT + T(\log_k T - 1)$$
$$= O(T \max(k, \log_k(T))).$$

Hence, the integrality gap is at least $\min(\log_k(T), k)$. For every k, we can choose T appropriately so that $\min(\log_k(T), k) = k$. Letting k go to infinity and we have the desired result.

Proof of of Lemma 8

Proof. Let $\hat{t}$ be a time period such that there exist items i, j with $\tilde{x}_{i,\hat{t}}$ and $\tilde{x}_{j,\hat{t}}$ both fractional. Without loss of generality assume that $v_i/w_i > v_j/w_j$. Let $t^1 = \min\{t : \tilde{x}_{j,t} > 0\}$, and $\epsilon = \tilde{x}_{j,t^1}$. Let $t^2 = \max\{t : \tilde{x}_{i,t} = \tilde{x}_{i,\hat{t}}\}$, and when $t^2 < T$ set $\delta = \tilde{x}_{i,t^2+1} - \tilde{x}_{i,\hat{t}}$ and otherwise set $\delta = 1 - \tilde{x}_{i,\hat{t}}$. Finally, write $\theta = \min\left\{\frac{w_i}{w_j}\delta, \epsilon\right\}$. Consider the vector y created by the following rule:

$$y_{j,t} = \tilde{x}_{j,t} - \theta, \quad \text{for } t^1 \leq t \leq t^2,$$
$$y_{i,t} = \tilde{x}_{i,t} + \frac{w_j}{w_i}\theta, \text{ for } t^1 \leq t \leq t^2,$$
$$y_{k,t} = \tilde{x}_{k,t}, \quad \text{for all remaining } k \text{ and } t.$$

Note that $i, j \neq h$ because constraint (6) of the formulation guarantees that $\tilde{x}_{h,t}$ is integral for all t. Thus, y is a feasible solution to $D^R(\bar{t}, \check{t}, h)$. But since $v_i/w_i > v_j/w_j$ (because we have perturbed data so that all values/weight ratios are distinct) the objective value attained by y is strictly larger than that of $\tilde{x}$, a contradiction.

Proof of of Lemma 9

Proof. For any period t, let $F(t) \doteq \{i : 0 < \tilde{x}_{i,t} < 1\}$. Suppose first that

$$\sum_{t=1}^{\bar{t}-1} \sum_{i \in F(t)} v_i \Delta_t \tilde{x}_{it} < \frac{1}{6}V^*.$$

By constraint (8) we therefore have

$$\sum_{t=1}^{\bar{t}-1} \sum_{i \notin F(t)} v_i \Delta_t \breve{x}_{it} > \frac{1}{6} V^*$$

and

$$\sum_{i,t} v_i \Delta t \, \breve{x}_{i,t} \geq \min \frac{1}{6} \left\{ 1, \frac{\sum_{t=\bar{t}}^{T} \Delta_t}{\sum_{t=1}^{\bar{t}-1} \Delta_t} \right\} \sum_{i,t} v_i \Delta t \, \tilde{x}_{i,t}. \tag{13}$$

(i.e. (10)) follows, which is the desired result. Thus, we instead assume that

$$\sum_{t=1}^{\bar{t}-1} \sum_{i \in F(t)} v_i \Delta_t \tilde{x}_{it} \geq \frac{1}{6} V^*.$$

In particular this means that $F(t) \neq \emptyset$ for at least one $1 \leq t \leq \bar{t} - 1$. But, for such t, if $i \in F(t)$ then $\tilde{x}_{i,\bar{t}} > 0$ and by (6) and (6),

$$v_i \leq v_h.$$

So, using Lemma 8,

$$\frac{1}{6} V^* < \sum_{t=1}^{\bar{t}-1} \sum_{i \in F(t)} v_i \Delta_t \tilde{x}_{it} \leq \sum_{t=1}^{\bar{t}-1} v_h \Delta_t \leq \frac{\sum_{t=1}^{\bar{t}-1} \Delta_t}{\sum_{t=\bar{t}}^{T} \Delta_t} \sum_{t=\bar{t}}^{T} \sum_{i} v_i \Delta t \, \breve{x}_{i,t}. \tag{14}$$

From this relationship (13) follows.

Proof of of Lemma 10

Proof. Define $t^{[3]}$ to be the largest period so that (a) holds. If $t^{[3]} = T$ then the replicated-knapsack solution at T has value at least $\frac{V^Z}{3}$. So we can assume $t^{[3]} < T$ (and similarly, that $1 < t^{[3]}$) and therefore

$$\sum_{t=1}^{t^{[3]}-1} \sum_{i} v_i \Delta_t x_{it}^Z < \frac{1}{3} V^Z.$$

Hence if (b) does not hold, then

$$\sum_{i} v_i \Delta_t x_{it^{[3]}}^Z > \frac{1}{3} V^Z,$$

and so the replicated-knapsack solution at $t^{[3]}$ has value at least $\frac{V^Z}{3}$.

Proof of of Lemma 12

Proof. We observe that if, under x^Z, some item i is added to the knapsack at time t, then, since x^Z is optimal and $v_i > 0$, it must be the case that it could not have been added any earlier (while keeping the remaining schedule fixed, otherwise). Under this assumption, the schedule is "pushed to the left" as much as possible.

Suppose first that no items are added to the knapsack, under x^Z, in the periods $t^{[3]}, \ldots, T$. Then by condition (a) of Lemma 10 we have that by replicating the knapsack solution at time $t^{[3]}$ we will have total value is at least $V^*/3$, as desired.

Thus, let t_1 be the first period $\geq t^{[3]}$ such that x^Z adds an item is added to the knapsack at that period. By the "pushed to the left" analysis, we have that either $t_1 = t^{[3]}$, or

$$t_1 > t^{[3]}, \quad \text{and}$$
$$\sum_i w_i x_{i,t_1} > B_{t^{[3]}}. \tag{15}$$

Let t_2 be the first time period after t_1 where x^Z adds an item knapsack. If no such period exists, let $t_2 = T$. Finally let $A(1)$ be the set of items added by x^Z to the knapsack in period t_1, and $A(2)$ the set of items added in periods t_2 through T. Now the sum

$$\sum_{t=t^{[3]}}^{T} \sum_i v_i \Delta_t x_{it}^Z$$

which by condition (a) of Lemma 10 is at least $V^*/3$, can be split into three terms, some of which may be empty:

(a) $(\sum_{t=t^{[3]}}^{T} \Delta_t)(\sum_i v_i x_{i,t^{[3]}})$
(b) $(\sum_{t=t_1}^{T} \Delta_t)(\sum_{i \in A(1)} v_i)$
(c) $\sum_{t=t_2}^{T} \Delta t \sum_{i \in A(2)} v_i x_{i,t}$.

So the largest of these three terms has value at least $V^*/9$. If it is (a) or (b) then we have that a replicated-knapsack solution of value at least $V^*/9$ and we are done. If it is (c), then note that by (15) (and the pushed to the left condition)

$$\sum_{i \in A(2)} w_i \leq B_T - B_{t^{[3]}} \leq B_T - B_{t_{1/2}} \leq B_{t_{1/2}}, \tag{16}$$

where the inequalities follow because $t_{1/2} < t^{[3]}$ and B_t is nondecreasing. It follows that we obtain a feasible solution to $\mathcal{IK}$ by placing in the knapsack the set $A(2)$, at time $t_{1/2}$ (and no items added at any other time). This solution is feasible by (16), and its value is at least the quantity in (c) and so at least $V^*/9$. This is the same as saying that the replicated-knapsack solution at $t_{1/2}$ has value at least $V^*/9$.

Extension of Lemma 5 to general $0 < \kappa < 1$

Here we outline how to extend Lemma 5 to obtain, for general $0 < \kappa < 1$, an approximation algorithm with ratio $\min\{\Omega(\frac{1-\kappa}{\kappa}), \frac{1}{S(\kappa)}\}$. The second case in the "min" corresponds to the case where $t^{[3]} < t_\kappa$, and it follows by an analysis very similar to that of Lemma 11 (the reader will notice that in that proof, the fact that $\kappa = 1/2$ was not actually used).

Now we consider the case $t^{[3]} \geq t_\kappa$. First, we note that if $\kappa < 1/2$ is such that an (S, κ)-split takes place at time t, then an $(S, 1/2)$-split takes place at t, as well (from (ii)) in Definition 3. So we will assume $1/2 < \kappa < 1$.

Now suppose an (S, κ)-split takes place at time t_κ. The idea is to split the time interval $[t_\kappa, T]$ into intervals $[t^0, t^1], [t^1, t^2], \ldots [t^{m-1}, t^m]$, where $t_\kappa = t^0 \leq t^1 \leq \ldots t^m = T$, $m = O(\kappa/(1-\kappa))$, and the capacity increase experienced in each interval $[t^i, t^{i+1}]$ is $O(B_T/(1-\kappa))$. Then, at the boundary between successive intervals we apply an analysis similar to that used to prove Lemma 12 to consider (1) two possible replicated-knapsacks (as in cases (a), (b) of the proof of Lemma 12), or (2) a solution similar to that in case (c) of Lemma 12, which in this case will use (by time T) total capacity at most $(1 - \kappa)B_T$, and thus can be used to lower bound the replicated-knapsack solution at time t_κ. Altogether, therefore, we obtain $O(3\kappa/(1-\kappa))$ replicated-knapsack solutions, and thus the best attains an approximation factor

Proof of Lemma 13

Proof. For a fixed $k \in \{1, \ldots, K\}$, we have that $\sigma_1^k \leq \sigma_2^k \leq \ldots \leq \sigma_T^k$. If $\sigma_T^k = m$, then since the σ_i^ks are integers, there are at most $\binom{m+T-1}{m}$ feasible T-tuples $(\sigma_1^k, \sigma_2^k, \ldots, \sigma_T^k)$. Since $0 \leq m \leq J$, we have that $\binom{m+T-1}{m} \leq (J+T)^J$. Consequently, there are at most $\sum_{m=1}^{J} \binom{m+T-1}{m} \leq J(J+T)^J$ feasible T-tuples $(\sigma_1^k, \sigma_2^k, \ldots, \sigma_T^k)$. Thus, there are at most $(J(J+T)^J)^K = O((1/\epsilon + T)^{O(\log(T/\epsilon)/\epsilon^2)})$ in the disjunctive procedure. $\square$

Proof of Theorem 15

Proof. Let $\bar{x}$ be an optimal solution of $\max\{\sum_{t=1}^{T} \sum_{i \in S^h} v_i' x_{i,t} : x \in Q^{\sigma,h}\}$. In order to prove the theorem, we will show the validity of the inequality (see Lemma 18)

$$\sum_{t=1}^{T} \sum_{i=1}^{|S^{k,h}|} v_i' x_{k_i,t}^{\sigma,h} \geq (1-\epsilon) \sum_{t=1}^{T} \sum_{i=1}^{|S^{k,h}|} v_i' \bar{x}_{k_i,t}, \tag{17}$$

for every $S^{k,h}$ and that of (see Lemma 19)

$$\sum_{t=1}^{T} \sum_{i \in T^h} v_i' x_{i,t}^{\sigma,h} \geq \sum_{t=1}^{T} \sum_{i \in T^h} v_i' \bar{x}_{i,t} - \epsilon v_h. \tag{18}$$

The two inequalities imply:

$$\sum_{t=1}^{T} \sum_{i \in S^h} v_i' x_{i,t}^{\sigma,h} = v_h + \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{i=1}^{|S^{k,h}|} v_i' x_{k_i,t}^{\sigma,h} + \sum_{t=1}^{T} \sum_{i \in T^h} v_i' x_{i,t}^{\sigma,h}$$

$$\geq v_h + (1-\epsilon) \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{i=1}^{|S^{k,h}|} v_i' \bar{x}_{k_i,t} + \sum_{t=1}^{T} \sum_{i \in T^h} v_i' \bar{x}_{i,t} - \epsilon v_h$$

$$\geq \sum_{t=1}^{T} \sum_{i \in S^h} v_i' \bar{x}_{i,t}.$$

Next, we give proofs for equations (17) and (18). For every nonempty polyhedron $Q^{\sigma,h}$, we begin by showing (17) holds for every $S^{k,h}$ with the help of the following auxiliary LP.

$$\max \quad \sum_{t=1}^{T} \sum_{i=1}^{|S^{k,h}|} v_{k_i}' x_{k_i,t}$$

$$\text{s.t.} \quad \sum_{i \in T^h} w_{k_i} x_{k_i,t} \leq \sum_{i=1}^{|S^{k,h}|} w_{k_i} \bar{x}_{k_i,t} \quad \forall t$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{|S^{k,h}|},t} = 0 \quad \forall(k,t) \text{ s.t. } \sigma_t^k = 0$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{\sigma_t^k},t} = 1,$$

$$x_{k_{\sigma_t^k}+1,t} = \ldots = x_{k_{|S^{k,h}|},t} = 0 \quad \forall(k,t) \text{ s.t. } 1 \leq \sigma_t^k < J \text{ and } \sigma_t^k < |S^{k,h}|$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{\sigma_t^k},t} = 1, \quad \sum_{i=1}^{|S^{k,h}|} x_{k_i,t} \geq \sigma_t^k \quad \forall(k,t) \text{ s.t. } \sigma_t^k = J \text{ and } \sigma_t^k < |S^{k,h}|$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{|S^{k,h}|},t} = 1 \quad \forall(k,t) \text{ s.t. } \sigma_t^k \geq |S^{k,h}|$$

$$x_{k_i,t-1} \leq x_{k_i,t} \quad \forall k_i, \text{ and } t = 2, 3, \ldots, T$$

$$x_{k_i,t-1} \in [0,1] \quad \forall k_i, t.$$

Lemma 17. *For every $S^{k,h}$, there exists an optimal solution to the auxiliary LP that contains at most one fractional variable $x_{k_i,t}$ in each time period t.*

Proof. The claim is true when $\sigma_t^k < J$ and $\sigma_t^k < |S^{k,h}|$ or when $\sigma_t^k \geq |S^{k,h}|$ as there are no fractional variables in both cases. Hence, the only case left is when $\sigma_t^k = J$ and $\sigma_t^k \geq |S^{k,h}|$.

Let $t^\star$ be the first (smallest) period in which we are in the case $\sigma_t^k = J$ and $\sigma_t^k \geq |S^{k,h}|$. Ignoring the precedence constraints for a moment, then

the auxiliary LP can be broken up into $T - t^{\star} + 1$ single period LPs of the following form, one for each $t \geq t^{\star}$.

$$LP_t = \max \sum_{i=1}^{|S^{k,h}|} v_i' x_{k_i,t}$$

$$\text{s.t.} \quad \sum_{i=1}^{|S^{k,h}|} w_i x_{k_i,t} \leq \sum_{i=1}^{|S^{k,h}|} w_i \bar{x}_{k_i,t} \tag{19}$$

$$x_{k_1,t} = x_{k_2,t} = \ldots = x_{k_{\sigma_t^k},t} = 1, \quad \sum_{i=1}^{|S^{k,h}|} x_{k_i,t} \geq \sigma_t^k \quad \forall (k,t) \text{ s.t. } \sigma_t^k = J \text{ and } \sigma_t^k < |S^{k,h}| \tag{20}$$

$$x_{k_i,t-1} \in [0,1] \quad \forall k_i.$$

Notice that in the modified instance of the problem, all items have the same value within a value class $S^{k,h}$. Hence, an optimal solution to LP_t is simply to pack the items in the order of their weight, starting from the smallest weight item first. Moreover, notice that this set of optimal solutions satisfy the precedence constraints.

Lemma 18. *Let $\bar{x}$ be an optimal solution to the optimization problem over a non-empty $Q^{\sigma,h}$ for some $\sigma \in \{0, \ldots, J\}^{TK}$ and $h \in S$, then there exists an integer feasible solution $x^{\sigma,h}$ to the auxiliary LP such that*

$$\sum_{t=1}^{T} \sum_{i=1}^{|S^{k,h}|} v_i' x_{k_i,t}^{\sigma,h} \geq (1 - \epsilon) \sum_{t=1}^{T} \sum_{i=1}^{|S^{k,h}|} v_i' \bar{x}_{k_i,t}.$$

Proof. Without lost of generality, let $\hat{x}$ be the optimal solution to the auxiliary LP found using lemma 17. For time periods t where $\sigma_t^k < J$ and $\sigma_t^k < |S^{k,h}|$ or when $\sigma_t^k \geq |S^{k,h}|$, we don't need to round the variables $\hat{x}_{k_i,t}$ since they are already integral. Hence, we set $x_{k_i,t}^{\sigma,h} = \hat{x}_{k_i,t}$ for all the variables in this period, which implies that

$$\sum_{i=1}^{|S^{k,h}|} v_i' x_{k_i,t}^{\sigma,h} = \sum_{i=1}^{|S^{k,h}|} v_i' \hat{x}_{k_i,t} \geq \sum_{i=1}^{|S^{k,h}|} v_i' \bar{x}_{k_i,t}$$

for such a period t.

For time periods t where $\sigma_t^k = J$ and $\sigma_t^k \geq |S^{k,h}|$, by lemma 17, there is at most one fractional $\hat{x}_{k_i,t}$ in such a time period. Consequently, we round down this fractional variable while keeping others the same (or equivalently, setting $x_{k_i,t}^{\sigma,h} = 0$ for this variable and setting $x_{k_i,t}^{\sigma,h} = \hat{x}_{k_i,t}$). Since all the variables have the same value within a value class $S^{k,h}$, we have that

$$\frac{\sum_{i=1}^{|S^{k,h}|} v_i' \hat{x}_{k_i,t} - \sum_{i=1}^{|S^{k,h}|} v_i' x_{k_i,t}^{\sigma,h}}{\sum_{i=1}^{|S^{k,h}|} v_i' \hat{x}_{k_i,t}} \leq \frac{1}{J} < \epsilon,$$

which implies that

$$\sum_{i=1}^{|S^{k,h}|} v_i' x_{k_i,t}^{\sigma,h} \geq (1-\epsilon) \sum_{i=1}^{|S^{k,h}|} v_i' \hat{x}_{k_i,t} \geq (1-\epsilon) \sum_{i=1}^{|S^{k,h}|} v_i' \bar{x}_{k_i,t}.$$

Summing up the above inequalities over all time periods gives us the desired result.

Lemma 19. *Let $\bar{x}$ be an optimal solution to the optimization problem over a non-empty $Q^{\sigma,h}$ for some $\sigma \in \{0, \ldots, J\}^{TK}$ and $h \in S$, then there exists an integer feasible solution $x^{\sigma,h}$ to the auxiliary LP such that*

$$\sum_{t=1}^{T} \sum_{i \in T^h} v_i' x_{i,t}^{\sigma,h} \geq \sum_{t=1}^{T} \sum_{i \in T^h} v_i' \bar{x}_{i,t} - \epsilon v_h.$$

Proof. Consider the following auxiliary LP:

$$\max \quad \sum_{t=1}^{T} \sum_{i \in T^h} v_i' x_{i,t}$$

$$\text{s.t.} \quad \sum_{i \in T^h} w_i x_{i,t} \leq \sum_{i \in T^h} w_i \bar{x}_{i,t} \quad \forall t$$

$$x_{i,t-1} \leq x_{i,t} \quad \forall i \in T^h, \text{and } t = 2,3,\ldots,T$$

$$x_{i,t-1} \in [0,1] \quad \forall i,t.$$

An optimal solution of the LP above would be to greedily pack items in the order of non-increasing value to weight ratio. Let $\hat{x}$ be such an optimal solution, then it is clear that $\hat{x}$ has at most one fractional variable in each time period. We round down such a fractional variable in each time period to 0 to obtain an integer solution $x^{\sigma,h}$. Consequently, we have that

$$\sum_{t=1}^{T} \sum_{i \in T^h} v_i' \hat{x}_{i,t} - \sum_{t=1}^{T} \sum_{i \in T^h} v_i' x_{i,t}^{\sigma} \leq \frac{\epsilon v_h}{T} \sum_{t=1}^{T} 1 = \epsilon v_h,$$

where the first inequality follows from the fact that every item in T^h has value weakly less than $\epsilon v_h/T$. Rearrange the terms gives us that:

$$\sum_{t=1}^{T} \sum_{i \in T^h} v_i' x_{i,t}^{\sigma} \geq \sum_{t=1}^{T} \sum_{i \in T^h} v_i' \hat{x}_{i,t} - \epsilon v_h \geq \sum_{t=1}^{T} \sum_{i \in T^h} v_i' \bar{x}_{i,t} - \epsilon v_h.$$

Fri.Nov.15.114345.201